ENGINEERING THERMO-RESPONSIVE AFFINITY LIGANDS

FOR GLYCOPROTEIN PURIFICATION BY

AFFINITY PRECIPITATION

A Dissertation
Presented to
The Academic Faculty

by

Lindsay G. Arnold

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Chemical Engineering in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology
May 2015

ENGINEERING THERMO-RESPONSIVE AFFINITY LIGANDS

FOR GLYCOPROTEIN PURIFICATION BY

AFFINITY PRECIPITATION

Approved by:

Dr. Rachel Chen, Advisor
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Dr. Julie Champion
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Dr. Mark Prausnitz
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Dr. Sven Behrens
School of Chemical and Biomolecular
Engineering
Georgia Institute of Technology

Dr. Al Merrill
School of Biology
Georgia Institute of Technology

Date Approved: December 1, 2014

To my fostering family and my faithful friends

PREFACE

The research presented in this dissertation was written with either intent to publish or is already published in a scientific journal. The previously published manuscripts were modified to meet formatting guidelines. This dissertation features and introduction on glycoprotein purification, current methods and technology (Chapter 1), the experimental results of this research project with discussion and methods (Chapters 2-5), and conclusions of this dissertation along with recommendations on directions for future studies (Chapter 6).

ACKNOWLEDGEMENTS

It is disproportionate to only have one page of acknowledgements in this document, as the help, support, and guidance I've received accounts for the majority of this success.

First, I'd like to thank my thesis advisor, Dr. Rachel Chen. Dr. Chen's guidance, mentoring, instruction, and insight were invaluable throughout every step of this arduous process. She has taught me to be a diligent and attentive scientist. I would also like to thank Dr. Hyun-Dong Shin who was a daily resource of experience and knowledge and who encouraged me to be inquisitive of all things in all aspects of life. Drs. Chen and Shin shaped me in more ways than I can describe. Special thanks to my committee, as well: Dr. Sven Beherns, Dr. Julie Champion, Dr. Al Merrill, and Dr. Mark Prausnitz. My committee's support, commitment, and enthusiasm regarding my work was beneficial and critical to the cultivation of my final thesis.

I would also like to extend endless gratitude to my supportive family. My parents and sister, Katie, have always been ready with words of love and encouragement, and have been models of hard work and dedication. Special thanks to my boyfriend, Barrett, for his unwavering confidence in me and continuous support.

Finally, thank you to my grad student colleagues, who have made grad student life not just manageable, but quite enjoyable. Thanks to past Chen group members for entertainment as well as education. To my study group (with the appropriate acronym MATLAB ME), I wish continued success in everything you put your minds to. Thank you for the constant encouragement, corroboration, and validation I needed to get me this far.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

BSA	bovine serum albumin
CHO	chinese hamster ovary
ELP	elastin like protein
EPO	erythropoietin
FBS	fetal bovine serum
GFP	green fluorescent protein
GGGS	glycine-glycine-glycine-serine
Gly	glycine
His-tag	6 histidine affinity tag
HRP	horseradish peroxidase
HSA	human serum albumin
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	inverse temperature cycling
K_d	dissociation constant
kDa	kilo Daltons (moleculare weight unit)
LB	luria broth
lpp	Brauns lipoprotein
MLAC	multi-lectin affinity chromatography
MLAP	multi-lectin affinity precipitation
MW	molecular weight
NanH	neuraminidase enzyme

OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pro	proline
RSL	lectin from *Ralstonia solanacearum*
RSL-ELPXX	lectin fused to elastin like protein, XX designates number of repeats
SA	sialic acid
SBP	soybean peroxidase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	secretion pathway
Tat	twin arginine translocation
T_t	phase transition temperature
Val	valine
VCNA	lectin from *Vibrio cholerae*
VCNA-ELP40	lectin fused to elastin like protein

SUMMARY

Effective methods for isolation and purification of glycoproteins are of increasing significance to the rapidly growing biopharmaceutical and diagnostic industry. Glycoproteins represent the majority of therapeutic proteins on the market and are effectively used to treat immune disorders, infections, cancers, and other diseases. Targeting these glycoproteins is also critical to an emerging field of glycoproteomics aimed to understand structure-function relationships of glycans. Architecturally, these glycoproteins are proteins with covalently linked oligosaccharide chains of varying monosaccharide composition. Affinity chromatography has proven to be an excellent method of glycoprotein purification at the bench scale. However, chromatography in large scale production has its drawbacks. Column fowling, flowrate limitations, and diffusional constraints collectively hinder the effectiveness of the method. An alternative proposed in this dissertation is the use of affinity precipitation as a purification technique. The three main objectives are 1) develop and produce dual-functional, thermo-responsive affinity ligands from a biological host, 2) characterize and optimize the accompanying affinity precipitation method, and 3) apply the ligand and process to relevant, unmodified glycoproteins.

The design of the thermo-responsive affinity construct was comprised of two main functional domains. The binding capability was achieved by selection of small ligands with affinity to a specific monosaccharide moiety. Two different lectins, or sugar binding proteins, were used in the fusion design: a fucose binding lectin from *Ralstonia solanacearum*, and a sialic acid binding lectin from *Vibrio cholera*. The thermo-

responsive functionality was obtained by use of an elastin-like peptide (ELP), which confers inverse solubility relationship properties to the fusion construct. A small library of varying ELP chain lengths were designed to find the optimal size fusion for both production and function. These dual functional ligands were cloned and expressed in the microbial host, *E. coli*. Furthermore, secretion of these constructs was achieved by employing the Tat secretion pathway in combination with an outer membrane lipoprotein deletion mutant with a leaky periplasm phenotype. This secretory mechanism allows for easy isolation, avoidance of inclusion bodies, and no additional protease inhibitors.

After successful production, the ligands were tested to confirm that dual functionality was preserved in fusion form. Once binding conditions and precipitation properties were ascertained, the purification ability was tested on model glycoproteins. Experimentation was carried out monitoring the purification yield, purity, and retained activity of the target enzymes. High contaminant solutions, such as cell lysates, were spiked with the model glycoproteins to mimic crude protein solutions. The purification ability of the constructs in these models was observed.

The method was then implemented on two relevant glycoprotein applications: 1) purification of soybean peroxidase from a crude protein extract and 2) targeting the therapeutic protein erythropoietin from albumin rich, used CHO cell media. By implementation of the fucose targeting fusion construct, the unmodified soybean peroxidase is isolated from a natural crude extract from the soybean hull, a by-product of the soybean industry. The affinity precipitation method parameters were optimized with respect to ratios, temperatures, recycle, and elution buffers to achieve successful isolation

of the low abundance enzyme. Under the optimized conditions, >95% recovery yield and a purification of 22.7 fold of an active, pure product was attainable.

The purification of erythropoietin led to additional experimentation with high-abundant glycoprotein solutions, as well as expansion of the affinity ligand platform. The concept of multi-lectin affinity precipitation, using the fucose and sialic acid binding lection sequentially, was introduced and tested for purification capability. An industrially relevant scheme involving isolation of the erythropoietin from used CHO cell media allowed for an achievable yield of about 60%, with a resulting albumin depletion of about 85%.

In addition to development of a pair of novel thermo-responsive affinity ligands for glycoprotein purification, this dissertation provides insight on possible improvements and future directions with respect to the thermo-responsive affinity ligand platform. This unique concept employs novel lectin fusions to target valuable glycoproteins using a method avoiding the major drawbacks associated with chromatography.

CHAPTER I

INTRODUCTION

1.1 Glycoprotein Purification Applications

Glycoproteins are conjugated proteins containing covalently bonded carbohydrates to the amino acid backbone. The study of glycoconjugates has quickly expanded to cover almost all biochemical processes in living organisms, and glycosylation is regarded as the most frequent and most diverse post translational modification[1–3]. Specifically, glycoproteins have been found as essential factors in processes such as immune defense, fertilization, cell adhesion, and inflammation[4]. More than 50% of eukaryotic proteins are glycosylated[5]. Furthermore, glycoproteins are often used as biomarkers for certain diseases, such as rheumatoid arthritis, cystic fibrosis, or various cancers[6–8], and the field of glycomics, the study of these carbohydrates compositions, has emerged. With this range of study, it is clearly beneficial in many areas to isolate and target these glycoproteins, both for further characterization as well as glycoprofiling and activity determination. Thus, many methods of glycoprotein purification have been experimentally pursued to purify a range of different glycoproteins for different applications.

1.1.1 Therapeutic Protein Purification

Over 70% of the therapeutic protein drugs on the market are classified as glycoproteins[9]. These therapeutic proteins are the fastest growing market in the

pharmaceutical industry, and include several categories of drugs, such as monoclonal antibodies, enzymes, blood factors, and hormones[10]. It has even been estimated over 30% of new drugs licensed throughout the next ten years will be antibodies and antibody derivatives[4,11]. While these protein therapeutics generate a 30 billion dollar industry, these high demand drugs are expensive to produce. The high cost, however, is primarily traced to downstream purification procedures, comprising 50-80% of the total manufacturing costs[12].

Typically, the downstream process is comprised of three mains phases of preparation: clarification, purification, and formulation[13]. While clarification and formulation are required steps in all fermentation based production processes, the purification phase is specific to the desired target. In the case of protein therapeutics and most drug products, purification is the bottleneck in terms of cost, energy, complexity, and specification requirements[14,15]. The purification technology has yet to meet the demand of these drugs, but the currently implemented standard is chromatography. While various types of chromatography are often employed, such as size exclusion chromatography or anion exchange chromatography, the most relied upon methodology for targeting specific glycoproteins is affinity chromatography[16,17].

1.1.1.1 Purification Methods

Affinity chromatography has proven to be the most ideal to date, as it utilizes high specificity and attempts to avoid denaturing the target therapeutic protein by exposure to low pH or high salt conditions. To obtain this high specificity, it exploits the use of affinity ligands. These affinity ligands are typically proteins that have a high binding specificity to the target. These affinity ligands are immobilized in a column, allowing the

clarified crude protein solution to be flushed through the column while the target remains bound to the ligand. The target is then eluted from the column using an elution buffer. The use of these affinity ligands allows for the most ideal targeting of a specific product. In current industrial purification schemes, Protein A and Protein G are by far the most widely used ligands.[12,18] Either full form or truncated, with association constants in the 10^8 M^{-1} range, these have been the gold standard for targeting immunoglobulin G species and other monoclonal antibodies.

1.1.1.2 Importance of Glycan Recognition

However, isolation of a target glycoprotein is not the whole story. As specific isoforms of glycans have been noted to enhance protein stability, extend therapeutic half-life, modulate biological activity, and increase efficacy, it is imperative the final drug product contain the correct isoform and composition of glycan[19,20]. Different isoforms can be produced by variation in culture conditions. For example, osmolality levels as well as culture duration contribute to increased levels of mannose-5 glycoform of antibodies in Chinese hamster ovary (CHO) cell production[21]. This leads reduced efficacy of the antibodies, as mannose-5 glycoforms are found in very low amounts in the body.

Glycoproteins can range from 1-90% in carbohydrate composition[3]. There are 8 common glycan residues, each of which could be further derivatized. Furthermore, the glycosidic linkages can be formed between any of the hydroxyl groups, leading to numerous configurations and linkages. For example, four different monosaccharides could link a variety of ways and form 35,560 unique tetrasaccharides[22]. Unlike proteins and nucleic acids, glycan chains can be branched as well as linear. This complex

microheterogeneity leads to many glycoforms present in a solution, and must be addressed when producing a pure and efficacious drug product. An example of this increasing complexity is shown below by illustrating increased branching and core additions.
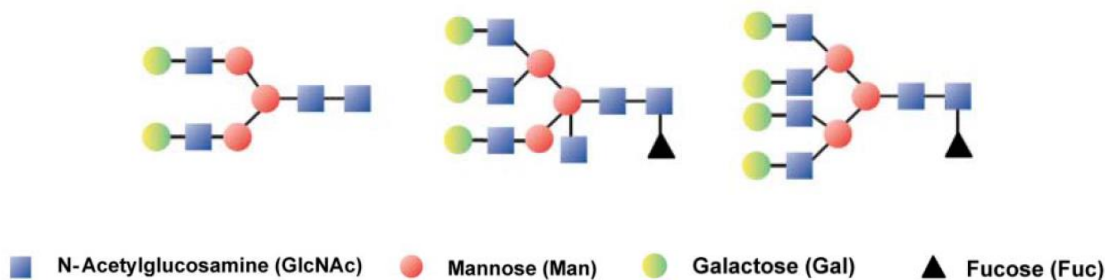
Figure 1.1 Example glycans. Adapted from Dalpathado et. al.[23]

The purification process must, in turn, target a single glycoform. Because of this additional complexity, affinity chromatography is not sufficient as a single method of purification. As described earlier, affinity chromatography procedures currently used industrially purify by exploiting the binding properties between ligand and polypeptide backbone. However, this does not discern amongst glycoforms. Additional purification methods, such as anion exchange and size exclusion chromatography, must be implemented to yield a single isoform of the desired therapeutic product.

1.1.2 Crude Enzyme Purification

Many clinically and industrially useful glycoproteins are extracted from natural sources. In order to foster a "green" economy, biosciences must find new ways to target and apply natural sources to more products and services[24,25]. Purification techniques are

4

required for identifying new and useful glycoproteins from crude sources from plants, animal serums, fungi, and bacteria. Specifically, such useful glycoproteins include peroxidases, which are used for a myriad of industrial or laboratory applications.

One of the more commonly used peroxidases is one isolated from horseradish roots. It is the most widely studied member of the peroxidase enzyme family EC 1.11.1.7, and has been applied to many biosensing and diagnostic applications[26]. As a biotechnological tool, horseradish peroxidase (HRP) is invaluable in ELISA assays and bioremediation, so much effort has been put into extraction and purification of the valuable glycoprotein. Early purification methods still prevail; processes include ammonium sulfate precipitation followed by chromatographic methods, such as ion-exchange, size exclusion, or affinity chromatography. This combination of techniques was used in the 1960's and is used today[27]. Although column chromatography technology has improved, it often requires very costly set-ups, as addressed in section 1.2.1. Other methods have been pursued, such as aqueous two-phase partitioning, with varying results[28]. However, as the buffer in which the HRP is stored can increase the sensitivity of the HRP thermal stability, caution must be taken with the solvents chosen[29]. These alternative solvent methods also require large amounts of water that is wasted, rendering them relatively inefficient. In order to purify HRP on the order of magnitude required for large applications, such as wastewater treatment, other alternatives must be considered[30]. Also, the relatively poor stability of this enzyme limits the purification methods that can be applied.

Where HRP has a long standing reputation as a useful tool, a less traditional enzyme that emerged in the 1990s as a stable and worthy competitor is soybean

peroxidase (SBP)[31,32]. SBP is a peroxidase isolated from the hull, or seed coat, of soybeans. As the hulls are an inexpensive by-product of soybean processing, this enzyme already lends itself credence in decreased cost of raw material[33]. This soybean peroxidase (SBP) has shown promising thermal and pH stability where HRP has not.[34,35] With a large pH range of 2.4-12 and a melting temperature of 90.5°C, SBP has the potential to bypass the current popularity of HRP to become a very useful enzyme[36]. This stable enzyme can easily be applied as a catalyst for phenolic resin synthesis, an indicator for food processing, a diagnostic component, an additive for bioremediation, or a bleach reagent for paper processing[36–39].

1.1.3 Monosaccharide Significance

Glycans are typically comprised of eight common monosaccharides, shown below in Table 1.1. All are very similar in composition. In the case of glucose, mannose, and galactose, the only different is the confirmation of a hydroxyl group. Alterations of these monosaccharides in a glycan directly affect the metabolic fate of the glycoprotein, as well as possible structure-function relationships[3].

As glycans are involved in cell-cell recognition, a change in monosaccharide composition also relates to adhesion and pathogenicity[41]. Additionally, "changes in glycosylation are often a hallmark of disease states" [42], so identification, progression, and treatment often directly link to the presence and abundance of certain monosaccharides. By designing a system that targets a specific monosaccharide, not only is there enrichment of low abundance glycoproteins, but there is fractionation toward a specific characteristic monosaccharide, thus eliminating some undesired glycoforms. This

method aids in biomarker discovery and therapeutic protein purification. Specifically, two lesser abundant carbohydrate moieties, fucose and sialic acid, were used as target monosaccharides.

Table 1.1 The most common monosaccharides found in glycans and their size[40].

Common Monosaccharides	Abbreviation	Residue Mass (g/mol)
Glucose	Glc	180.16
Galactose	Gal	180.16
Mannose	Man	180.16
Fucose	Fuc	164.16
Acetylgalactosamine	GalNAc	221.21
Acetylglucosamine	GlcNAc	221.21
Acetylneuraminic acid	NeuNAc	309.27
Xylose	Xyl	150.13

1.1.3.1 Fucose

Fucose is a deoxy hexose monosaccharide present in the oligosaccharide chains of many glycans and glycolipids. Though it's not the most abundant oligosaccharide, its function and presence are important in adhesion, host-microbe interactions, signaling, inflammation, and many ontogenic events[6,43,44]. Terminal glycan fucosylation is common and can confer unique properties for cell differentiation. Aberrant fucosylation is also linked with as many as 18 diseases. Fucose has become an important biomarker for many conditions, such as certain cancers, liver disease, rheumatoid arthritis, Alzheimers

7

disease, cystic fibrosis, and asthma[45,46]. In antibodies, the structure of glycans N-linked to immunoglobulin G (IgG) molecules is often core-fucosylated, typically via α1,3 or α1,6 linkages[47,48]. Sometimes, antibodies should be isolated based on their fucosylation. For example, Antibody Dependent Cellular Cytotoxicity is increased 100 fold when IgG4 is non-fucosylated, so removal of fucosylated IgG4 is a necessary step in pharmaceutical production[11].

1.1.3.2 Sialic acid

Sialic acid, or N-acetylneuraminic acid, is a larger monosaccharide comprised of a 9 carbon backbone and an acetyl group that confers a negative charge to the molecule. Even as a less abundant monosaccharide, it is of great physiological and pathological importance. It is typically found at the terminal position of glycans on animal glycoconjugates and is thought to have appeared late in evolution[49]. It is also the most diversely substituted natural sugar[50]. Sialic acid, like fucose, contributes to the high viscosity of mucins, is often found in membranes, and is often involved in pathogenic recognition. However, because it's a unique target monosaccharide, it is also used in anti-recognition effects resulting in the masking of penultimate sugars[51]. Unfortunately, over sialylated malignancies are often masked as well, keeping them protected from cellular defense systems. In vivo, sialic acid is also used for cell-cell recognition. For example, red blood cells having lost 10-20% of their surface sialic acids are targeted, degraded, and expelled from the body.

Because of the importance of sialic acid residues, extensive studies of sialylated glycoproteins and their glycoforms have been researched. It's been shown that additional sialic acid residues increase the negative net charge and improve the pharmacokinetic

properties of certain glycoproteins[20,52]. Specifically, analysis of the therapeutic protein erythropoietin (EPO) illustrates the role sialic acid plays in biosynthesis, secretion, expression, and, mainly, biological activity. Removal of the terminal sialic acids result in complete loss of in vivo activity, but increase in vitro biological activity[52].

1.2 Affinity Precipitation

To attempt to circumvent the pitfalls associated with affinity chromatography, affinity precipitation has become an increasingly investigated technique. As the name suggests, affinity precipitation exploits the same target specific ligands, but avoids ligand immobilization and chromatographic methods. Affinity precipitation is achieved by utilization of the affinity ligand bound to a stimulus responsive polymer. As the binding step of the ligand to the target protein is done in aqueous conditions, diffusional limitations and steric hindrance are minimized. The ligand bound target complex is then precipitated out of solution via the environmentally stimulated polymer. This yields a precipitate containing the ligand, target, and some non-specific precipitates. The soluble contaminants are removed, and the precipitate is resolubilized. The reversibility of precipitation makes the tool recyclable, reducing the cost of materials. It is also very scalable, especially when compared to affinity chromatography, which makes it a viable candidate for large scale processes.

1.2.1 Affinity Chromatography Disadvantages

Affinity chromatography has proved an excellent method of glycoprotein purification at bench scale processes. However, chromatography in large scale production has its drawbacks. Column fowling, flowrate limitations, and diffusional constraints collectively hinder the effectiveness of the purification process[17,53]. The biological component, such as Protein A, is expensive, sensitive to cleaning solutions, and sometimes leaks from the matrix into the product, requiring further purification steps downstream[18]. Furthermore, high water usage in flow through, multiple washing steps, and elution steps augment the cost and waste of the process[54]. Because of these shortcomings, affinity chromatography proves to be a very costly and insufficient method of purification when scaled up. Alternative methods of purification, preferably avoiding chromatographic methods, are currently in investigation. Affinity precipitation is a promising substitute.

1.2.2 Alternative Precipitation Methods

Precipitation as a method of purification is not a new concept, especially in relation to proteins. As misfolding and aggregating proteins is relatively simple to do, it's long been experimented with as a method of purification. A common example is ammonium sulfate precipitation, which is widely used and is easy to scale up[53]. Also, solid-liquid separation is straight forward and well-understood, which makes precipitation quite desirable, especially at the industrial scale. As different protein have differing ionic strengths, the salting-out method of altering the ionic properties of the solvent has been a quick and simple process. Sequential precipitation has even been employed to separate desired proteins from high contaminant solutions. Other

precipitation methods, such as acetone and caprylic acid precipitation, have also been explored. However, the main pitfall of all these methods is the lack of specificity. They are not always amenable for efficiently targeting a specific isoform or glycoform.

In an effort to develop a more specific purification system, affinity macroligands were synthesized. These affinity macroligands had dual functionality: an affinity domain and a stimulus responsive region. The first attempt at this was a copolymer of acrylamide, N-acryloyl-p-aminobenzoic acid, and N-acryolyl-m-aminobenzamidine[55,56]. The benzamidine units also acted as affinity ligands to target enzymes. However, the environmental stimulus required for purification was a decrease in pH to 4, which often hindered the activity of the target enzyme.

Another attempt was the design of a Eudragit S-100 responsive ligand. Eudragit S-100 is a copolymer of methacrylic acid and methyl methacrylate and also has pH dependent solubility[57,58]. This ligand was chemically fused to an affinity ligand, specifically the antigen of the IgG monoclonal antibody, and precipitated by lowering the pH to 4.8. Centrifugation was used to separate the pellet, and the purified product was analyzed. This procedure yielded a purification of 68.4%, with antibody activity lost due to inactivation.

1.2.3 Current Affinity Precipitation Methods

Affinity precipitation also improves the overall process of therapeutic protein separations by enhancing process integration, which can potentially save on energy, materials, and time. After fermentation, clarification of the fermentation mixture is necessary to avoid further fouling of the equipment, as well as to achieve a practical

viscosity for the other downstream processing steps. After initial clarification come multiple isolation steps, including affinity and size exclusion chromatography. Finally, consolidation steps concentrate the target protein from the high dilution steps involved in chromatography down to a highly concentrated pure batch of drug substance. The many steps involved in the purification and consolidation process each allow for potential loss of product, lowering the final yield of pure protein. However, by utilizing affinity precipitation, many of these steps can be consolidated into a single step procedure. Clarification can be kept to a minimum, as the affinity ligand can be added to a crude solution. The purification of the technique is inherent, and though other purification steps may be needed for virus clearance, it eliminates the need for multiple chromatographic steps. Also, as the product of affinity precipitation is initially a solid precipitate, resuspension in a smaller amount of liquid for solubilization dramatically aids in the consolidation step. Thus, affinity precipitation aids in all three major sections or therapeutic protein purification. Consolidation of this process not only has the potential to save time and money, but often rapid purification aids in maintaining stabilization, improving the active yield.

1.2.4 Elastin-Like Proteins (ELPs)

One particularly useful biopolymer used in affinity precipitation is Elastin-Like Proteins (ELPs). Originally derived from mammalian elastin, ELPs are composed of a repeating pentapeptide sequence Val-Pro-Gly-Xaa-Gly, where the fourth position, Xaa, is a guest residue that can be substituted by any naturally occurring amino acid except proline[59,60]. These biopolymers exhibit a reversible thermal response at a tunable

transition temperature Tt, above which the polymer becomes insoluble[61–64]. This inverse phase transition behavior can be manipulated and applied to many areas in biotechnology, such as non-chromatographic purification, drug delivery, tissue engineering, responsive hydrogels, and biosensing[65,66].

The responsiveness of an ELP peptide is affected, at the molecular level, by the hydrophobicity of the guest residue and the total length of the peptide. Alteration of the guest residue has been studied extensively, with variations in both the guest amino acids and fractions at which they appear in the total peptide. This deviation manifests as a very distinct, almost predictable change in transition temperature. This resulting change in transition temperature has even been used as a measure of the hydrophobicity of residues respectively[59,61]. This hydrophobicity scale is formed on the basis of an increase in transition temperature when the pentapeptide contains a more hydrophobic residue, where, conversely, a more hydrophilic residue yields a lower transition temperature. This tunable nature is also based on the length of the resulting peptide. An extended chain of pentapeptide repeats in the sequence decreases the transition temperature.

After these intrinsic factors have been specified, the phase behavior then becomes a dependent on other solution conditions, such as pH, salt concentration, and concentration of the peptide in solution[65]. Predictive models on this variation in phase transition have been done for various constructs of ELPs, and the behavior at variable conditions is fairly well understood.

As suggested by the relation to hydrophobicity, the ELP solubility is related to the secondary structure of the protein and the hydration state[67–70]. Below the transition temperature, the protein is relatively disordered, but at higher temperatures, it becomes

more compact, due to an increased number of intramolecular hydrogen bonds[64]. At this lower hydration state, the protein precipitates, albeit differently than that of standard protein denaturation. By lowering the temperature again, the entropically driven process unfolds, rehydrates, and becomes soluble once more. Fusion of ELPs has been implemented in an effort to use this unique inverse transition state for protein purification purposes.

1.2.4.1 ELP Tags

At the transcriptional level, ELP biopolymer has been experimentally fused to the gene corresponding to the protein of interest, typically on a plasmid. This ELP tag confers the phase transition properties to the target protein, so the target can be isolated by environmental stimulus. Both soluble and insoluble proteins, such as thioredoxin and tendamistat respectively, have been fused with ELP tags in microbiological hosts and efficiently harvested from cultures[71–74]. This demonstrates the versatility of the tags in relation to aiding in solubility as well as purification. The fusion proteins display retained activity inspite of the ELP fusion, suggesting functionality is not lost. Fusion order has been investigated, and has been noted to sometimes effect the expression of the fusion protein[75]. Expression of these fusion proteins in E. coli has been further optimized to increase production of these larger constructs[73,76]. This fusion method is not limited to microbes, but as also been investigated in molecular farming techniques. An ELP tag on biologically active glycoprotein 130, GFP, erythropoietin, and interleukin-10, amongst others, have been successfully expressed in transgenic tobacco plants[77–79]. The ELP tag allows for an easy purification step, allowing for dramatic reduction of undesirable phenolics and toxic alkaloids and enrichment of low yield target in a single process.

However, in most practical applications, a purified protein containing an additional tag is undesirable. The ELP must be removed from the desired product. However, many chemical cleaving mechanisms require harsh chemicals, compromising the folding and activity of the desired protein in the process. Proteolytic methods are costly, as addition enzymes are required. So in order to address this obstacle, a few solutions have been introduced.

1.2.4.2 ELP Intein Fusions

The first solution to address the removal of the ELP tag from the target protein was implementation of a self-cleaving intein fusion between the ELP fusion and the protein of interest. An intein is a protein containing an endonuclease and splicing sequence that is self-cleaving upon pH variation[80]. Mini inteins have been developed and used as fusions for the purpose of affinity tag addition and removal. Target proteins have target-intein-tag domain conformations, where the affinity tag is a moiety that can be targeted via affinity chromatography, such as maltose binding protein[81,82]. However, as it is desirable to shift away from affinity chromatography, this affinity tag can be replaced with other mechanisms of purification, such as an ELP functional domain. Dr. Wood's lab, amongst others, has designed a variety of these intein-ELP fusion proteins, in both microbial and mammalian hosts[72,83–86]. While functional and very efficient, it is still required that the protein of interest are recombinant and transcriptionally fused to these domains, which is not always a viable route. This requires additional cloning, plasmid maintenance, and expression optimization. Furthermore, this approach cannot be used with new or natural crude proteins; the approach must be catered to a specific

recombinant target. Other approaches must be investigated in increase the range of applicability.

<u>1.2.4.3 ELP Fused to Affinity Tags</u>

Instead of ELP fusion to the protein of interest, a different approach has evolved that expands the applicability of the ELP purification scheme. An ELP fusion to an affinity tag still exploits the inverse temperature properties of the ELP while maintaining the specificity achieved by using an affinity tag[87–89]. This technique is similar to the Eudragit S-100 technique; however, the ELP functionality allows for different stimuli to facilitate precipitation. Instead of requiring a low pH, which often denatures sensitive therapeutic glycoproteins, a change in salt concentration, temperature, or combination of the two is the initiating factor for precipitation.

| Binding | Precipitation | Ligand Release |
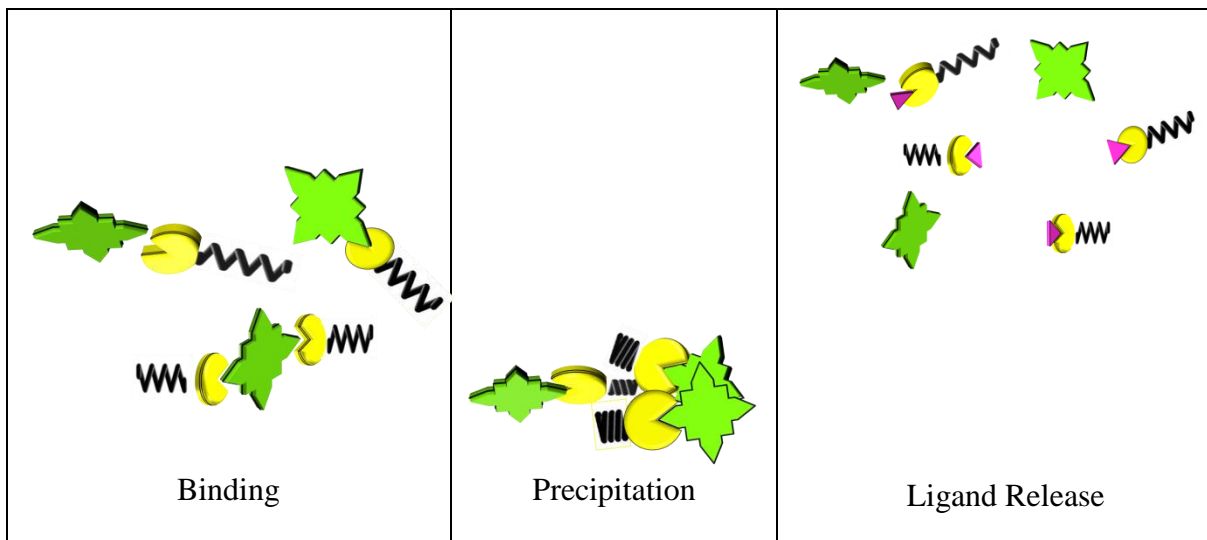
Figure 1.2 Schematic of purification method using an ELP-affinity construct.

Using this method, simplified steps and solid liquid separation are used to target a specific moiety of a glycoprotein of interest, as depicted below in Figure 1.1. Exploiting

the affinity end, the ELP-affinity fusion construct binds to the target. The precipitation of the ELP is then achieved by some external stimuli, transitioning the protein of interest to the solid state. Finally, the ELP-affinity ligand and target protein are reimbursed in a new solution and solubilized. After release, the ELP-affinity ligand is again isolated, leaving a solution containing only the target protein.

Furthermore, instead of an additional fusion step, the entire construct is produced in a bacterial host, so no additional chemical steps or lost ligand are required. While this is more of an economic advantage, it's also a benefit in consolidation of the process: where production and isolation can be done without any extra steps.

1.3 Lectins

Sugar binding proteins, called lectins, have emerged as vital tools for identification, purification, and structural analysis of complex carbohydrates and glycans. More than 300 lectins have been discovered, most of which are plant-based in origin[90]. However, lectins have been discovered in animals, microorganisms, and viruses as well[91,92]. Lectins function biologically as cell-surface receptors, cell adhesion, and glycan synthesis control[22,93]. They are thought to play an important role in the immune system as well, particularly in pathogen recognition. Typically multimeric in structure, their superb binding capacity (10^{-6} M) has made lectins key players in glycomics and protein purification[94]. Although non-enzymatic in function, they recognize mono- and oligosaccharides with varying specificities.

1.3.1 Lectin Technology

Because of their unique sugar targeting properties, lectins have been employed in various constructs of biomolecular analytical methods. Specifically, lectins are used as tools to target specific glycan compositions. Lectinosorbent assays, lectinoblotting, and biotinylated lectins have all been used for characterization and quantitation of glycoconjugates[95,96]. Lectin microarrays have aided tremendously in glyco-profiling, and have given rise to new high-throughput techniques[90,97]. However, for glycoprotein enrichment and isolation, it is lectin affinity chromatography that has emerged as a key player in glycomics[98]. Lectin affinity chromatography has been successfully utilized for biomarker detection, diagnostics, and characterization, as well as larger scale downstream processing[15,98,99]. Immobilization of lectins on a column allow for higher throughput of sample under benign conditions, enabling capture of glycoproteins containing specific sugar moieties.

1.3.2 Multi-lectin Affinity

In addition to lectin affinity chromatography, a novel technique of interest is the implementation of multiple types of lectins immobilized on a column, called multi-lectin affinity chromatography (M-LAC). By combining the affinity of different lectins to different respective carbohydrates, the target can be defined by more than a single monosaccharide composition. In fact, this method allows for screening of glycoproteins with known glycosylation patterns. In some cases, as many as 5 different lectins were chosen to selectively capture different glycoproteins in a biological sample[100]. Glycoproteins are fractionated based on their total glycan compositions, and other high

18

abundance glycoproteins can be depleted[99,101,102]. Use of this technique allows more explicit detection of varying isoforms for the same glycoprotein of interest. Furthermore, this creates a customizable platform for targeting different glycoproteins.

1.3.3 Bacterial Lectins

While plant lectins are often utilized for their interactions with complex glycans or internal sugar sequences, bacterial lectins generally bind to terminal monosaccharides. Often referred to as adhesion proteins, these lectins are used for attachment to host cells and facilitate infection[103]. Although some lectins are promiscuous in binding, most are specific as to linkage of a monosaccharide. Bacterial lectins, as already a product of bacterial organisms, are more likely to be compatible with recombinant expression in E. coli, due to similar codon and protein usage. With minimal codon optimization, bacterial lectins can be easily duplicated, as no posttranslational modification is necessary. However, as many plant lectins are now commercially available, further development and utilization of bacterial lectins has not been a large focus. By producing these lectins in E. coli¸ higher yields can be obtained in a shorter amount of time than lectin production in fungi and plants, making bacterial lectins a cost effective alternative[104]. Bacterial lectins are also generally much smaller than higher organism lectins, which again lends credence to compatibility of expression in *E. coli*. Furthermore, the small size makes these lectins optimal candidates for fusion proteins, as smaller, less-bulky proteins are less likely to interfere with the other functional domain. Although these lectins are less popular in the current field of lectinology, I chose to pursue fusion and expression of a bacterial lectin, and I hope this research stimulates further exploration and manipulation of bacterial lectins in the future.

1.3.3.1 Ralstonia solanacearum Lectin (RSL)

Based on the importance of fucose and the many prospects associated with isolation based on fucosylation, fucose-binding lectins were investigated. There are many fucose binding lectins used in monosaccharide identification. A particularly suited lectin is one found in the bacteria *Ralstonia solanacearum* (RSL). As this lectin originates in bacteria, expression in E.coli is more favorable than the expression of a protein natively found in a plant or animal. The RSL lectin is also relatively small at only 90 amino acids and 10 kD[105]. It does not require additional metal ions for binding, such as calcium or magnesium. As with many lectins, it forms a homo-oligomer, specifically a homotrimer for optimal binding. It binds preferentially to L-fucose, and preferentially still toward α1,6 and α1,2 linkages of fucose to an oligosaccharide chain[106]. Putatively, this lectin is involved in the infectivity of *R. solanacaearum*, which known to cause lethal wilt in plants. It has 6 binding sites for fucose; 3 are within a monomer (intramonomer), while 3 are sandwiched between two of the three monomers (intermonomer), as pictured in Figure 1.3. This lectin has an affinity for fucosylated oligosaccharaides in the 10^{-7} M range, and an affinity for free fucose in the 10^{-6} M range. The binding of fucose is primarily a result of hydrogen bonding of amino acids to the hydroxyl groups, but Van der Waals forces have been characterized as well.
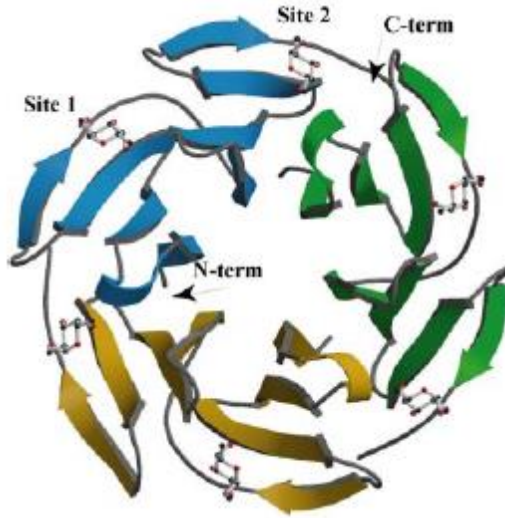
Figure 1.3 RSL structure[107]

1.3.3.2 *Vibrio cholera* Neuraminidase (VCNA)

As mentioned in section 1.1.3, often the presence of certain monosaccharides directly affects the pathogenicity of certain bacteria. In the case of *V. cholera*, the organism has evolved to remove the abundance of sialic acids present from high order gangliosides to reveal the real receptor, GM1[108,109]. The removal of sialic acid from surrounding mucin layers is credited to a secreted neuraminidase enzyme. This *Vibrio cholera* neuraminidase (VCNA) is comprised of 3 distinct domains: a neuraminidase catalytic domain, and two flanking lectin regions[110]. Of these two lectin domains, the N-terminal lectin has been found to bind with high affinity to sialic acid (K_d of 30 μM) and similar affinity to sialic acid containing substrates, α-2,3-sialyllactose and α-2,6-sialyllactose[108]. This 21 kD binding site is efficient in targeting terminal sialic acid moieties and does not require the addition of metal ions for binding. As pictured in

Figure 1.4, the binding is primarily hydrogen binding with some water molecule mediated interactions.
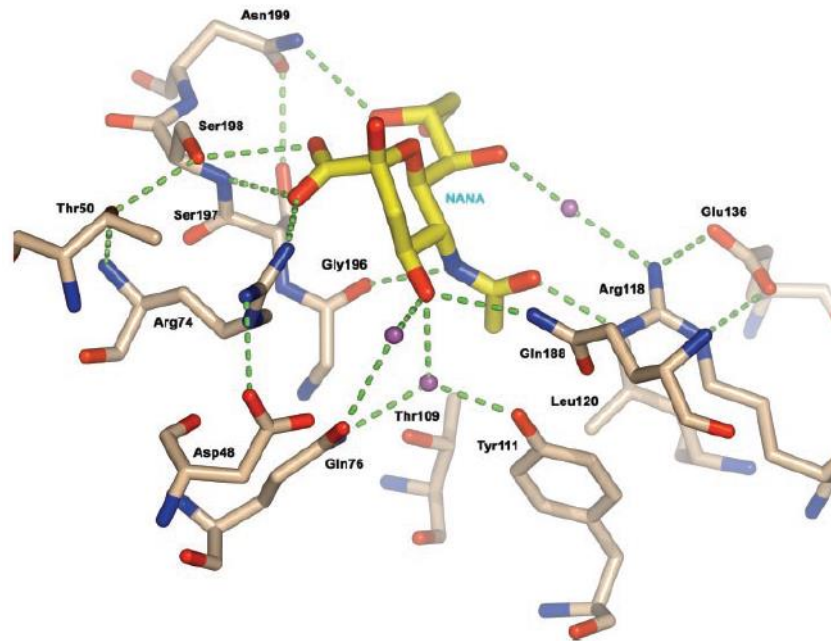
Figure 1.4 VCNA binding pocket with associated sialic acid[108].

1.4 Host Secretion

In biotechnology, an essential tool for recombinant protein production is the use of *Escherichia coli*. *E. coli* is one of the most popular hosts for protein expression as it is well characterized, has many inherent known and available expression systems, and is relatively easy to manipulate[111]. However, some major drawbacks are commonly associated with expression in *E. coli*. Over expressed proteins often form inclusion bodies, which then require extra steps for denaturing and refolding and frequently results

in a loss of activity. Certain protein cannot be actively expressed at all in *E. coli*, as complex disulfide bonds cannot be formed in the reducing nature of the cytoplasm[112]. Conversely, should the protein be translocated to the periplasm, proteolytic degradation will likely occur, again reducing total yield of desired protein. Ideally, extracellular secretion of a recombinant protein would suggest greatest yield, as well as avoid the need for cell lysis for protein isolation. This would be ideal in ease of production, which logically leads to a reduced cost of the process as a whole. However, *E. coli* does not have natural, efficient extracellular secretion mechanisms. This obstacle must be addressed to proficiently express the desired recombinant protein and resourcefully purify it without using sonication, addition of protease inhibitors, or chromatographic processes. Multiple molecular engineering approaches have been taken to try to integrate bacterial secretion pathways into *E. coli.* Combinations of translocation pathways and increased outer membrane permeability have led to some promising options.

1.4.1 Tat Pathway

Several prokaryotic organisms express secretory proteins by way of signal sequence. A popular design of signal sequence is one harboring a twin-arginine motif, designated as the twin-arginine translocation (Tat) pathway. This characteristic motif is typically seen as an N-terminal signal sequence comprised of S-R-R-x-F-L-K residues[113]. Another distinct characteristic of this pathway is the ability to translocate fully folded proteins across a membrane. This is dissimilar from other typical mechanisms that, while still translocating post-translationally, delay folding until after the transmembrane movement. This is beneficial for proper folding in the cytoplasm. Additionally, a large

range of protein sizes are amenable with the Tat dependent pathway. Tat substrates encompass roughly 10-100 kDa proteins without inciting addition cytoplasmic leakage[114]. However, this signal sequence induced translocation merely moves the desired protein across the inner membrane. Collection of the target protein in the periplasm still leads to inclusion bodies and requires additional cell processing. To complete the process of complete extracellular secretion, additional steps must be engineered. A mutant of *E. coli* strain, designated E609Y, has been engineered with a deletion of the coding for Braun's lipoprotein, resulting in a leaky periplasm phenotype[115]. Combining this *lpp* deletion mutant with the Tat dependent pathway, recombinant proteins are translocated to the periplasm and leaked out into the extracellular milieu. This has resulted in relatively high extracellular expression without the requirement of cell lysis. In the proposed work, the fusion proteins will be produced using a well characterized *E. coli* expression system. Since these fusions are thermal responsive ligands and are used in affinity precipitation, the cost of production must be low enough for large scale industrial production. This could be achieved by implementing this extracellular secretion production pathway.

1.4.2 Sec Pathway

Another secretion pathway amenable to bacterial protein production in E. coli is the Sec pathway. Similarly to the Tat-dependent pathway, an N-terminal signal sequence is targeted after translation. These signal sequences are characteristically comprised of a hydrophobic core of ten or more residues flanked by positively charged terminal regions[116]. Once targeted, the protein is translocated, and continues on to either its native folded state or outer membrane integration[117]. Cleavage is not required for complete

protein secretion. At least 11 proteins are confirmed to be involved with this secretion mechanism, which requires extra metabolic burden on the cell compared to that of the Tat pathway, which only requires 3^{113}. However, this method of secretion has been applied abundantly to a wide range of biotechnological applications with great success.

1.5 Project Objectives

The overall goal of this project is to engineer a recombinant thermal responsive affinity ligand by combining the carbohydrate binding properties of a lectin with the thermal responsive properties of elastin to form lectin-ELP fusion proteins. These fusion proteins will be used to develop a novel thermal responsive affinity precipitation method for isolation of glycoproteins from high contaminant solutions. I propose to engineer several fusion proteins by manipulating the size of the ELP fusion, which results in changes in the precipitation temperature. Furthermore, fusion proteins with different lectin domains will be engineered, so as to target a variety of monosaccharides found on desirable glycoproteins. These fusion proteins will be produced from an *E. coli* expression system capable of extracellular secretion. The primary application will be glycoprotein purification processes for valuable peroxidases and therapeutic proteins.

1.5.1 Construction and Production

The synthetic gene encoding these fusion proteins will be created by a transcriptional fusion of the lectin to an ELP. By cloning and expressing the constructs in a microbial host, I will achieve efficient production. Furthermore, when created in

tandem with extracellular secretion, I will illustrate ease of production and isolation from a commonly used biological model, *E. coli*, making this method a cost effective, integrated alternative to current separations processes.

By studying the secretion levels of a range of different sized fusion proteins, the benefits and limitations of this previously developed secretion pathway for *E. coli* will be evaluated as a method of large scale protein production for the construct. Different secretion pathways will be investigated, as well as a variety of culture conditions to optimize expression.

1.5.2 Ligand Characterization

The produced fusion protein will be tested for functionality and response based on process conditions such as temperature, pH, and salt concentration. The transition temperature will be characterized based on these parameters. The transition temperature will be varied to extend the range of application of the construct to high and low temperature situations. The trimer formation of the lectin portion as a function of time, temperature, and fucose concentration will be experimentally determined, as well as the effect this trimerization has on lectin binding. The binding parameters will be experimentally determined.

1.5.3 Affinity Precipitation Purification Method Optimization

Using a fucose-binding lectin as a model affinity ligand, a method of fucosylated glycoprotein purification will be developed. Horseradish peroxidase, a fucosylated enzyme, will be used as a model protein. Using a sialic acid-binding lectin as a second

fusion construct, a method of sialylated glycoprotein purification will be developed. Fetuin, a sialylated glycoprotein, will be used as a model protein. Parameters relevant to that purification process such as temperature, target protein concentration, affinity ligand concentration, precipitation time, incubation time, and pH will be investigated. The robustness of this new affinity precipitation method will be tested to constraints such as low detection limits and high contaminant concentrations. The recyclability of the fusion construct will also be studied. The two constructs will then be used in tandem to create a multi-lectin affinity precipitation technique for glycoproteins containing both fucose and sialic acid.

1.5.4 Industrial Application

The advantages of this bi-functional protein will be extended to purification of an industrially desirable glycoproteins. Specifically, soybean peroxidase and erythropoietin will be isolated from other glycopeptides present in a crude biological sample. This will allow rapid isolation of targeted glycopeptides for further identification. This is a much more focused approach of analysis, potentially leading to a rapid method for biomarker identification for potential medical detection applications.

1.6 References

1. Wiederschain, G. Y. Glycobiology: progress, problems, and perspectives. *Biochem. Biokhimii͡a* **78,** 679–96 (2013).

2. Breen, K. C., Coughlan, C. M. & Hayes, F. D. The role of glycoproteins in neural development function, and disease. *Mol. Neurobiol.* **16,** 163–220 (1998).

3. Voet, D. & Voet, J. *Biochemistry*. (Wiley, 2004).

4. Reichelt, S. *et al.* Amino-functionalized monolithic spin-type columns for high-throughput lectin affinity chromatography of glycoproteins. *Analyst* **137,** 2600–7 (2012).

5. Ito, S., Hayama, K. & Hirabayashi, J. Enrichment Strategies for Glycopeptides. *Methods Mol. Biol.* **534,** 195–203 (2009).

6. Becker, D. J. & Lowe, J. B. Fucose: biosynthesis and biological function in mammals. *Glycobiology* **13,** 41R–53R (2003).

7. Taniguchi, N. *et al.* Review A glycomic approach to the identification and characterization of glycoprotein function in cells transfected with glycosyltransferase genes. 239–247 (2001).

8. Ghazarian, H., Idoni, B. & Oppenheimer, S. B. A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem.* **113,** 236–47 (2011).

9. Li, H. & d'Anjou, M. Pharmacological significance of glycosylation in therapeutic proteins. *Curr. Opin. Biotechnol.* **20,** 678–84 (2009).

10. Datta, P., Linhardt, R. J. & Sharfstein, S. T. An 'omics approach towards CHO cell engineering. *Biotechnol. Bioeng.* **110,** 1255–71 (2013).

11. Beck, A. *et al.* Trends in Glycosylation , Glycoanalysis and Glycoengineering of Thera- peutic Antibodies and Fc-Fusion Proteins. 482–501 (2008).

12. Roque, a C. a, Lowe, C. R. & Taipa, M. A. Antibodies and genetically engineered related molecules: production and purification. *Biotechnol. Prog.* **20,** 639–54 (2004).

13. Wang, F. *et al.* A novel "pipeline" system for downstream preparation of therapeutic monoclonal antibodies. *Biotechnol. Lett.* **35,** 1411–8 (2013).

14. Polykarpou, E. M., Dalby, P. a & Papageorgiou, L. G. Optimal synthesis of chromatographic trains for downstream protein processing. *Biotechnol. Prog.* **27,** 1653–60 (2011).

15. Opitz, L., Salaklang, J., Büttner, H., Reichl, U. & Wolff, M. W. Lectin-affinity chromatography for downstream processing of MDCK cell culture derived human influenza A viruses. *Vaccine* **25,** 939–47 (2007).

16. Burnouf, T. & Radosevich, M. Affinity chromatography in the industrial purification of plasma proteins for therapeutic use. *J. Biochem. Biophys. Methods* **49,** 575–86 (2001).

17. Lightfoot, E. N. & Cockrem, M. C. M. Complex Fitness Diagrams: Downstream Processing of Biologicals. *Sep. Sci. Technol.* **48,** 1753–1757 (2013).

18. Gottschalk, B. U. Downstream Processing of Monoclonal Antibodies : from High Dilution to High Purity. *BioPharmInternational.com* 1–17 (2010).

19. Shriver, Z., Raguram, S. & Sasisekharan, R. Glycomics: a pathway to a class of new and improved therapeutics. *Nat. Rev. Drug Discov.* **3,** 863–73 (2004).

20. Sethuraman, N. & Stadheim, T. a. Challenges in therapeutic glycoprotein production. *Curr. Opin. Biotechnol.* **17,** 341–6 (2006).

21. Pacis, E., Yu, M., Autsen, J., Bayer, R. & Li, F. Effects of cell culture conditions on antibody N-linked glycosylation-what affects high mannose 5 glycoform. *Biotechnol. Bioeng.* **108,** 2348–2358 (2011).

22. Mody, R., Joshi, S. & Chaney, W. Use of lectins as diagnostic and therapeutic tools for cancer. *J. Pharmacol. Toxicol. Methods* **33,** 1–10 (1995).

23. Dalpathado, D. S. & Desaire, H. Glycopeptide analysis by mass spectrometry. *Analyst* **133,** 731–8 (2008).

24. Lentz, T. D. & Akridge, J. T. Economic Evaluation of Alternative Supply Chains For Soybean Peroxidase. *J. Food Distrib. Res.* **28,** 28–41 (1997).

25. Hailu, G., Weersink, A. & Cahlík, F. Examining the Prospects for Commercialization of Soybean Peroxidase. *AgBioForum* **13,** 263–273 (2010).

26. Chattopadhyay, K. & Mazumdar, S. Structural and Conformational Stability of Horseradish Peroxidase: Effect of Temperature and pH. *Biochemistry* **39,** 263–270 (2000).

27. Shannon, L. M., Kay, E., Jow, Y., Shannon, L. X. I. & Lew, J. Y. Peroxidase Isozymes from Horseradish Roots : I . ISOLATION AND PHYSICAL PROPERTIES from Horseradish Roots. *J. Biol. Chem.* **241,** 2166–2172 (1966).

28. Miranda, M. V, Ctor, H. I. & Lahore, M. F. Horseradish Peroxidase Extraction and Purification by Aqueous Two-Phase Partition. *Appl. Biochem. Biotechnol.* **53,** 147–154 (1995).

29. Haifeng, L., Yuwen, L., Xiaomin, C., Zhiyong, W. & Cunxin, W. Effects of sodium phosphate buffer on horseradish peroxidase thermal stability. *J. Therm. Anal. Calorim.* **93,** 569–574 (2008).

30. Cooper, V. A. & Nicell, J. A. REMOVAL OF PHENOLS FROM A FOUNDRY WASTEWATER USING HORSERADISH PEROXIDASE. **30,** 954–964 (1996).

31. Ryan, B. J., Carolan, N. & O'Fágáin, C. Horseradish and soybean peroxidases: comparable tools for alternative niches? *Trends Biotechnol.* **24,** 355–63 (2006).

32. Gillikin, J. W. & Graham, J. S. Purification and Developmental Analysis of the Major Anionic Peroxidase from the Seed Coat of Glycine max. *Plant Physiol.* **96,** 214–20 (1991).

33.   Ghaemmaghami, F., Alemzadeh, I. & Motamed, S. Seed Coat Soybean Peroxidase : Extraction and Biocatalytic Properties Determination. **7,** (2010).

34.   Sakharov, I. Y., Berlina, A. N., Zherdev, A. V & Dzantiev, B. B. Advantages of soybean peroxidase over horseradish peroxidase as the enzyme label in chemiluminescent enzyme-linked immunosorbent assay of sulfamethoxypyridazine. *J. Agric. Food Chem.* **58,** 3284–9 (2010).

35.   Nissum, M., Schiødt, C. B. & Welinder, K. G. Reactions of soybean peroxidase and hydrogen peroxide pH 2.4-12.0, and veratryl alcohol at pH 2.4. *Biochim. Biophys. Acta* **1545,** 339–48 (2001).

36.   Mceldoon, J. P. & Dordick, J. S. Unusual Thermal Stability of Soybean Peroxidase. *Biotechnol. Prog.* **7938,** 555–558 (1996).

37.   VanHaandel, M. J. *et al.* Characterization of different commercial soybean peroxidase preparations and use of the enzyme for N-demethylation of methyl N-methylanthranilate To produce the food flavor methylanthranilate. *J. Agric. Food Chem.* **48,** 1949–54 (2000).

38.   Kamal, J. K. A. & Behere, D. V. Thermal and Conformational Stability of Seed Coat Soybean Peroxidase. *Biochemistry* **41,** 9034–9042 (2002).

39.   Henriksen, A. *et al.* Structure of soybean seed coat peroxidase : A plant peroxidase with unusual stability and haem-apoprotein interactions. 108–115 (2001). doi:10.1110/ps.37301.108

40.   Top Four Things to Know About Glycan Analysis by Mass Spectrometry. *Thermo Fish. Sci. Inc.* Http://www.functionalglycomics.org/ (2013).

41.   Comstock, L. E. Importance of glycans to the host-bacteroides mutualism in the mammalian intestine. *Cell Host Microbe* **5,** 522–6 (2009).

42.   Dube, D. H. & Bertozzi, C. R. Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov.* **4,** 477–88 (2005).

43.   Nakano, M., Kondo, A., Kakehi, K. & Taniguchi, N. Glycomics – a new target for pharmaceuticals. *Drug Discov. Today Technol.* **3,** 39–47 (2006).

44.   Ma, B., Simala-Grant, J. L. & Taylor, D. E. Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16,** 158R–184R (2006).

45.   Rakha, E. a *et al.* Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. *Mod. Pathol.* **18,** 1295–304 (2005).

46.   Park, S. *et al.* marker for colon cancer : a combinatorial approach for development of glycan biomarkers. **2376,** 2366–2376 (2012).

47.   Gomord, V. *et al.* Production and glycosylation of plant-made pharmaceuticals: the antibodies as a challenge. *Plant Biotechnol. J.* **2,** 83–100 (2004).

48. Thobhani, S., Yuen, C.-T., Bailey, M. J. a & Jones, C. Identification and quantification of N-linked oligosaccharides released from glycoproteins: an inter-laboratory study. *Glycobiology* **19,** 201–11 (2009).

49. Varki, A. Sialic Acids as Ligands in Recognition Phenomena. *FASEB J.* **11,** 248–255 (1997).

50. Mandal, C. Sialic acid binding lectins. *Experientia* **46,** 433–441 (1990).

51. Schauer, R. Achievements and challenges of sialic acid research. *Glycoconj. J.* **17,** 485–99 (2001).

52. Misaizu, T. *et al.* Role of antennary structure of N-linked sugar chains in renal handling of recombinant human erythropoietin. *Blood* **86,** 4097–104 (1995).

53. Hilbrig, F. & Freitag, R. Protein purification by affinity precipitation. *J. Chromatogr. B* **790,** 79–90 (2003).

54. Ho, S. V., McLaughlin, J. M., Cue, B. W. & Dunn, P. J. Environmental considerations in biologics manufacturing. *Green Chem.* **12,** 755 (2010).

55. Hilbrig, F., Stocker, G., Schläppi, J.-M., Kocher, H. & Freitag, R. Utilization of Group Specific Ligands in the Downstream Processing of Proteins by Affinity Precipitation. *Food Bioprod. Process.* **84,** 28–36 (2006).

56. Schneider, M., Guillot, C. & Lamy, B. THE AFFINITY PRECIPITATION TECHNIQUE. *Ann. N. Y. Acad. Sci.* **369,** 257–264 (1981).

57.   Linné-Larsson, E., Galaev, I., Lindahl, L. & Mattiasson, B. Affinity precipitation of Concanavalin A with p-aminophenyl- α-D-glucopyranoside modified Eudragit S-100. *Bioseparation* **6,** 273–282 (1996).

58.   Taipa, M. a, Kaul, R. H., Mattiasson, B. & Cabral, J. M. Recovery of a monoclonal antibody from hybridoma culture supernatant by affinity precipitation with Eudragit S-100. *Bioseparation* **9,** 291–8 (2000).

59.   Urry, D. W. *et al.* HYDROPHOBICITY SCALE FOR PROTEINS BASED ON INVERSE TEMPERATURE TRANSITIONS. *Biopolymers* **32,** 1243–1250 (1992).

60.   Gray, W., Sandberg, L. & Foster, J. Molecular Model for Elastin Structure and Function. *Nature* **246,** 461–466 (1973).

61.   Luan, C. H., Parker, T. M., Gowda, D. C. & Urry, D. W. Hydrophobicity of amino acid residues: differential scanning calorimetry and synthesis of the aromatic analogues of the polypentapeptide of elastin. *Biopolymers* **32,** 1251–61 (1992).

62.   Meyer, D. E. & Chilkoti, A. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* **3,** 357–67 (2002).

63. Meyer, D. E. & Chilkoti, A. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules* **5,** 846–851 (2004).

64. Krukau, A., Brovchenko, I. & Geiger, A. Temperature-Induced Conformational Transition of a Model Elastin-like Peptide GVG(VPGVG)3 in Water. *Biomacromolecules* **8,** 2196–2202 (2007).

65. MacKay, J. A., Callahan, D. J., FitzGerald, K. N. & Chilkoti, A. Quantitative Model of the Phase Behavior of Recombinant pH-Responsive Elastin-Like Polypeptides. *Biomacromolecules* **11,** 2873–2879 (2010).

66. Simnick, A. J., Lim, D. W., Chow, D. & Chilkoti, A. Biomedical and biotechnological applications of elastin-like polypeptides. *Polym. Rev.* **47,** 121–154 (2007).

67. Urry, D. W., Trapane, T. L. & Prasad, K. U. Phase-Structure Transitions of the Elastin Polypentapeptide- Water System Within the Framework of. *Glass* **24,** 2345–2356 (1985).

68. Debelle, L. & Tamburro, A. M. Elastin: molecular description and function. *Int. J. Biochem. Cell Biol.* **31,** 261–272 (1999).

69. Debelle, L. & Alix, A. J. P. The structures of elastins and their function. *Biochimie* **81,** 981–994 (1999).

70. Reiersen, H., Clarke, a R. & Rees, a R. Short elastin-like peptides exhibit the same temperature-induced structural transitions as elastin polymers: implications for protein engineering. *J. Mol. Biol.* **283,** 255–64 (1998).

71. Meyer, D. E. & Chilkoti, a. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nat. Biotechnol.* **17,** 1112–5 (1999).

72. Wood, D. W., Wu, W., Belfort, G., Derbyshire, V. & Belfort, M. A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* **17,** 889–92 (1999).

73. Fong, B. A. & Wood, D. W. Expression and purification of ELP-intein-tagged target proteins in high cell density E-coli fermentation. *Microb. Cell Fact.* **9,** (2010).

74. Shamji, M. F. *et al.* Development and characterization of a fusion protein between thermally responsive elastin-like polypeptide and interleukin-1 receptor antagonist: sustained release of a local antiinflammatory therapeutic. *Arthritis Rheum.* **56,** 3650–61 (2007).

75. Christensen, T. *et al.* Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. *Protein Sci.* **18,** 1377–87 (2009).

76. Chow, D. C., Dreher, M. R., Trabbic-Carlson, K. & Chilkoti, A. Ultra-high expression of a thermally responsive recombinant fusion protein in E. coli. *Biotechnol. Prog.* **22,** 638–46 (2006).

77. Lin, M., Rose-John, S., Grötzinger, J., Conrad, U. & Scheller, J. Functional expression of a biologically active fragment of soluble gp130 as an ELP-fusion protein in transgenic plants: purification via inverse transition cycling. *Biochem. J.* **398,** 577–83 (2006).

78. Conley, A. J., Joensuu, J. J., Jevnikar, A. M., Menassa, R. & Brandle, J. E. Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol. Bioeng.* **103,** 562–73 (2009).

79. Tian, L. & Sun, S. S. M. A cost-effective ELP-intein coupling system for recombinant protein purification from plant production platform. *PLoS One* **6,** e24183 (2011).

80. Derbyshire, V. *et al.* Genetic definition of a protein-splicing domain: functional mini-inteins support structure predictions and a model for intein evolution. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11466–71 (1997).

81. Xu, M. *et al.* Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231,** 1–13 (1999).

82. Li, Y. F. Self-cleaving fusion tags for recombinant protein production. *Biotechnol. Lett.* **33,** 869–881 (2011).

83. Shi, C., Meng, Q. & Wood, D. W. A dual ELP-tagged split intein system for non-chromatographic recombinant protein purification. *Appl. Microbiol. Biotechnol.* **97,** 829–35 (2013).

84. Shen, Y. *et al.* Expression and purification of moricin CM4 and human β-defensins 4 in Escherichia coli using a new technology. *Microbiol. Res.* **165,** 713–8 (2010).

85. Wood, D. W. *et al.* Optimized single-step affinity purification with a self-cleaving intein applied to human acidic fibroblast growth factor. *Biotechnol. Prog.* **16,** 1055–63 (2002).

86. Banki, M. R., Feng, L. A. & Wood, D. W. Simple bioseparations using self-cleaving elastin-like polypeptide tags. *Nat. Methods* **2,** 659–661 (2005).

87. Sheth, R. D., Madan, B., Chen, W. & Cramer, S. M. High-throughput screening for the development of a monoclonal antibody affinity precipitation step using ELP-z stimuli responsive biopolymers. *Biotechnol. Bioeng.* **110,** 2664–76 (2013).

88. Arnold, L. & Chen, R. Novel thermo-responsive fucose binding ligands for glycoprotein purification by affinity precipitation. *Biotechnol. Bioeng.* **111,** 413–417 (2014).

89. Madan, B., Chaudhary, G., Cramer, S. M. & Chen, W. ELP-z and ELP-zz capturing scaffolds for the purification of immunoglobulins by affinity precipitation. *J. Biotechnol.* **163,** 10–6 (2013).

90. Gemeiner, P., Koz, T. & Tkac, J. Glycan and Lectin Microarrays for Glycomics and Medicinal Applications. *Contract* **30,** 394–418 (2010).

91. Van Damme, E. J. M. Lectins as Tools to Select for Glycosylated Proteins. *Methods Mol. Biol.* **753,** 289–297 (2011).

92. Singh, R. S., Tiwary, a. K. & Kennedy, J. F. Lectins: Sources, Activities, and Applications. *Crit. Rev. Biotechnol.* **19,** 145–178 (1999).

93. Rini, J. M. Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.* **24,** 551–77 (1995).

94. Imberty, A., Mitchell, E. P. & Wimmerova, M. Structural basis of high-affinity glycan recognition by bacterial and fungal lectins. *Curr. Opin. Struct. Biol.* **15,** 525–534 (2005).

95. Wu, A. M., Lisowska, E., Duk, M. & Yang, Z. Lectins as tools in glycoconjugate research. *Glycoconj. J.* **26,** 899–913 (2009).

96. Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. & Solís, D. From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* **36,** 298–313 (2011).

97. Sánchez-Pomales, G., Morris, T. a, Falabella, J. B., Tarlov, M. J. & Zangmeister, R. a. A lectin-based gold nanoparticle assay for probing glycosylation of glycoproteins. *Biotechnol. Bioeng.* **109,** 2240–9 (2012).

98.   Madera, M., Mann, B., Mechref, Y. & Novotny, M. V. Efficacy of glycoprotein
        enrichment by microscale lectin affinity chromatography. *J. Sep. Sci.* **31,** 2722–32
        (2008).

99.   Lee, L. Y. *et al.* An optimized approach for enrichment of glycoproteins from cell
        culture lysates using native multi-lectin affinity chromatography. *J. Sep. Sci.* **35,**
        2445–52 (2012).

100.   Yang, Z. & Hancock, W. S. Approach to the comprehensive analysis of
        glycoproteins isolated from human serum using a multi-lectin affinity column. *J.
        Chromatogr. A* **1053,** 79–88 (2004).

101.   Plavina, T., Wakshull, E., Hancock, W. S. & Hincapie, M. Combination of
        Abundant Protein Depletion and Multi-Lectin Affinity Chromatography (M-LAC
) for Plasma Protein Biomarker Discovery research articles. 662–671 (2007).

102.   Yang, Z. & Hancock, W. S. Monitoring glycosylation pattern changes of
        glycoproteins using multi-lectin affinity chromatography. *J. Chromatogr. A* **1070,**
        57–64 (2005).

103.   Varki, A., Cummings & Esko. *Essentials of Glycobiology*. (Cold Spring Harbor
        Laboratory Press, 2009).

104.   Lam, S. K. & Ng, T. B. Lectins: production and practical applications. *Appl.
        Microbiol. Biotechnol.* **89,** 45–55 (2011).

105. Kostlanova, N. *et al.* The fucose-binding lectin from Ralstonia solanacearum - A new type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. *J. Biol. Chem.* **280,** 27839–27849 (2005).

106. Sudakevitz, D. *et al.* A new Ralstonia solanacearum high-affinity mannose-binding lectin RS-IIL structurally resembling the Pseudomonas aeruginosa fucose-specific lectin PA-IIL. *Mol. Microbiol.* **52,** 691–700 (2004).

107. Kostlánová, N. *et al.* The fucose-binding lectin from Ralstonia solanacearum. A new type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. *J. Biol. Chem.* **280,** 27839–49 (2005).

108. Moustafa, I. *et al.* Sialic acid recognition by Vibrio cholerae neuraminidase. *J. Biol. Chem.* **279,** 40819–26 (2004).

109. Lehmann, F., Tiralongo, E. & Tiralongo, J. Sialic acid-specific lectins: occurrence, specificity and function. *Cell. Mol. Life Sci.* **63,** 1331–1354 (2006).

110. Crennell, S., Garman, E., Laver, G., Vimr, E. & Taylor, G. Crystal structure of Vibrio cholerae neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. *Structure* **2,** 535–44 (1994).

111. Choi, J. H. & Lee, S. Y. Secretory and extracellular production of recombinant proteins using Escherichia coli. *Appl. Microbiol. Biotechnol.* **64,** 625–35 (2004).

112.   Chen, R. R. Permeability issues in whole-cell bioprocesses and cellular membrane engineering. *Appl. Microbiol. Biotechnol.* **74,** 730–8 (2007).

113.   Müller, M. Twin-arginine-specific protein export in Escherichia coli. *Res. Microbiol.* **156,** 131–6 (2005).

114.   Wu, L. F., Ize, B., Chanal, a, Quentin, Y. & Fichant, G. Bacterial twin-arginine signal peptide-dependent protein translocation pathway: evolution and mechanism. *J. Mol. Microbiol. Biotechnol.* **2,** 179–89 (2000).

115.   Shin, H.-D. & Chen, R. R. Extracellular recombinant protein production from an Escherichia coli lpp deletion mutant. *Biotechnol. Bioeng.* **101,** 1288–96 (2008).

116.   Mori, H. & Ito, K. The Sec protein-translocation pathway. *Trends Microbiol.* **9,** 494–500 (2001).

117.   Economou, a. Following the leader: bacterial protein export through the Sec pathway. *Trends Microbiol.* **7,** 315–20 (1999).

118.   Franken, K. L. *et al.* Purification of his-tagged proteins by immobilized chelate affinity chromatography: the benefits from the use of organic solvent. *Protein Expr. Purif.* **18,** 95–9 (2000).

119.   Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr. Purif.* **48,** 1–13 (2006).

120. Hengen, P. Purificatin of His-Tag fusion proteins from Escherichia coli. *Trends Biochem. Sci.* **20,** 285–285 (1995).

121. Meyer, D. E., Trabbic-Carlson, K. & Chilkoti, a. Protein purification by fusion with an environmentally responsive elastin-like polypeptide: effect of polypeptide length on the purification of thioredoxin. *Biotechnol. Prog.* **17,** 720–8 (2001).

122. Wriggers, W., Chakravarty, S. & Jennings, P. A. Control of protein functional dynamics by peptide linkers. *Biopolymers* **80,** 736–746 (2005).

123. Lu, P. & Feng, M. G. Bifunctional enhancement of a beta-glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Appl. Microbiol. Biotechnol.* **79,** 579–587 (2008).

124. Robinson, C. R. & Sauer, R. T. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5929–34 (1998).

125. McPherson, D. T., Xu, J. & Urry, D. W. Product purification by reversible phase transition following Escherichia coli expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP. *Protein Expr. Purif.* **7,** 51–57 (1996).

126. Shimazu, M., Mulchandani, A. & Chen, W. Thermally triggered purification and immobilization of elastin-OPH fusions. *Biotechnol. Bioeng.* **81,** 74–9 (2003).

127. Straley, K. S. & Heilshorn, S. C. Design and adsorption of modular engineered proteins to prepare customized, neuron-compatible coatings. *Front. Neuroeng.* **2,** 9 (2009).

128. Widdick, D. a *et al.* The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 17927–32 (2006).

129. Liu, N. *et al.* Fusion Proteins from Artificial and Natural Structural Modules. 107–121 (2001).

130. Floss, D. M., Schallau, K., Rose-John, S., Conrad, U. & Scheller, J. Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. *Trends Biotechnol.* **28,** 37–45 (2010).

131. Koria, P. *et al.* Self-assembling elastin-like peptides growth factor chimeric nanoparticles for the treatment of chronic wounds. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 1034–9 (2011).

132. Wuhrer, M., Koeleman, C. A. M., Hokke, C. H. & Deelder, M. Protein Glycosylation Analyzed by Normal-Phase of Glycopeptides. **77,** 886–894 (2005).

133. Gray, J. S., Yang, B. Y., Hull, S. R., Venzke, D. P. & Montgomery, R. The glycans of soybean peroxidase. *Glycobiology* **6,** 23–32 (1996).

134. Wood, D. W. Non-Chromatographic Recombinant Protein Purification by Self-Cleaving Purification Tags. *Sep. Sci. Technol.* **45,** 2345–2357 (2010).

135. Kostal, J., Mulchandani, A. & Chen, W. Tunable Biopolymers for Heavy Metal Removal. *Macromolecules* **34,** 2257–2261 (2001).

136. Ge, X. & Filipe, C. D. M. Simultaneous phase transition of ELP tagged molecules and free ELP: An efficient and reversible capture system. *Biomacromolecules* **7,** 2475–2478 (2006).

137. Spiro, R. G. Studies on fetuin, a glycoprotein of fetal serum. *J. Biol. Chem.* **235,** 2860–2869 (1960).

138. Spiro, R. G. & Bhoyroo, V. D. Structure of the O -Glycosidically Linked Carbohydrate Units of Fetuin Structure of the O-Glvcosidicallv of Fetuin * Carbohydrate. 5704–5717 (1974).

139. Husain, Q. Peroxidase mediated decolorization and remediation of wastewater containing industrial dyes: a review. *Rev. Environ. Sci. Bio/Technology* **9,** 117–140 (2009).

CHAPTER II

DESIGN, CONSTRUCTION, AND PRODUCTION

2.1 Abstract

Novel thermo-responsive affinity sugar binders were developed by fusing a bacterial fucose lectin with a thermo-responsive polypeptide. These designer affinity ligands fusions were produced using an *E. coli* system capable of extracellular secretion of recombinant proteins and were isolated with a high recovery yield (95%) directly from growth medium by Inverse Temperature Cycling (ITC). In this work, we present a recombinant thermal responsive affinity ligand by combining the carbohydrate binding properties of a lectin with the thermal responsive properties of elastin to form lectin-ELP fusion proteins.

2.2 Introduction

Protein purification is an essential method required in the biotechnological industry, with many different techniques currently used and being developed. With the popularity of recombinant proteins grew the capability of transcriptional fusion proteins. Either by plasmid or genomic incorporation, transcriptional fusions have become essential in identification, characterization, and purification of many recombinant proteins in both eukaryotic and prokaryotic systems.

Various types of tags are used in research today. Glutathione S-transferase, maltose-binding protein, chitin-binding protein, and biotin tags are affinity tag examples often fused to target protein[1.] The quintessential tag is merely a string of histidine

47

residues. Specifically, more than 60% of proteins produced for structural studies include a polyhistidine tag[2]. With this small fusion of amino acids, use of affinity purification results in yields of 90% or more. However, while this small modification is efficient for purification, certain scenarios such as clinical application require removal of this tag. To combat expensive or detrimental tag removal processes, new methods exploiting endoproteases are employed.

Typically, a fusion protein is designed at the transcriptional level. A cloned segment of DNA is ligated into a plasmid vector containing the desired codons for the affinity tag expression[3.] Whether this fusion is at the N-terminus or C-terminus of the target is dependent on various constraints, such as ease of cloning or folding requirements. Furthermore, a linker is often required, so the additional presence of the tag does not interfere with folding or activity of the target protein. As this fusion gene is downstream from a promoter, expression is induced and the fusion gene is produced. Expression optimization typically follows to avoid inclusion body formation.

Expression optimization is essential in structuring to lower the cost of production for the desired fusion proteins. Increased production in a biological host decreases the cost when compared to other synthetic and chemical methods. Changes in growth temperature, medium, and induction procedure are all crucial in optimization[4].

In this fashion, fusion proteins will be engineered. However, instead of using chromatographic methods, the fusion proteins will contain an environmentally responsive tag, so purification is as easy as temperature fluctuation. These fusion proteins will act as the binding ligand to target the native protein of interest, requiring no derivatization of the final product. They will be produced from an E. coli expression system capable of

extracellular secretion. We illustrate the bi-functionality of this affinity precipitation tool.

2.3 Results

2.3.1 Design, Cloning, and Expression of Constructs

The constructs were carefully designed for functionality, expression, secretion, and flexibility for future work. The synthetic gene was deliberately constructed to allow for oligomerization of the ELP segment so further optimization of the fusion could be performed. Additional functionalities, such as signal peptides, fusion order, and linkers were investigated as well.

2.3.1.1 RSL-ELP20

The first construct to be cloned and expressed was of a lectin with a C-terminal ELP fusion. This design then laid the platform for further development of the constructs. The following are the design considerations and cloning methods employed to produce the first construct.

DESIGN

A lectin from *Ralstonia solanacearum* (RSL) was selected for its size and functionality. A bacterial lectin of 9.9 kDa is small and amenable to expression in a bacterial host. Furthermore, as a bacterial lectin, minimal codon optimization was necessary compared to that of plant or mammalian lectins. In its natural form, this lectin forms a trimer, forming six binding sites with affinity to fucose[5,6]. Fucose is a prominent biomarker in many diseases, such as certain cancers, liver disease, rheumatoid arthritis, Alzheimers disease, cystic fibrosis, and asthma[7-10]. Fucose is also pertinent to the

growing realm of therapeutic proteins, as core-fucosylation is found on the heavy chain of many antibodies, which makes it a desirable monosaccharide to target[11].

The ELP segment required additional modifications. First of all, ELP's are comprised of a pentapeptide repeat, VPGXG, where X is a guest residue of any amino acid except proline. However, the guest residue dramatically affects the inverse solubility properties of this peptide[12,13]. The majority of previous ELP purification work uses either the single substitute of valine or a block copolymer design with varying ratios of alanine, valine, and glycine. The main specification to be optimized is the transition temperature. If the transition temperature is too high, protein denaturation and loss of activity are likely to occur. If the transition temperature is too low, the construct may precipitate inside the cell during culture cultivation and form unusable inclusion bodies. Moreover, if the construct is to be applied to industrial applications, heating a solution to a higher temperature would require an excessive amount of energy, thus increasing cost. A second specification to be met is the size of the ELP fusion. If the ELP chain has too many repeats, it will likely burden the cell, hinder expression, and possibly sterically hinder the binding of the lectin domain. In bacteria, decreasing ELP length from 90 to 20 pentapeptides increased expression yield by 4 times[14].

In an effort to keep the ELP fusion small, but with a low transition temperature, previous work done by Meyer's lab was consulted. As shown by the predictive model and accompanying experiments from a study by Meyer et. al., Figure 2.1 illustrates the lowest transition temperature attainable is achieved by the ELP containing only valine as the guest residue.
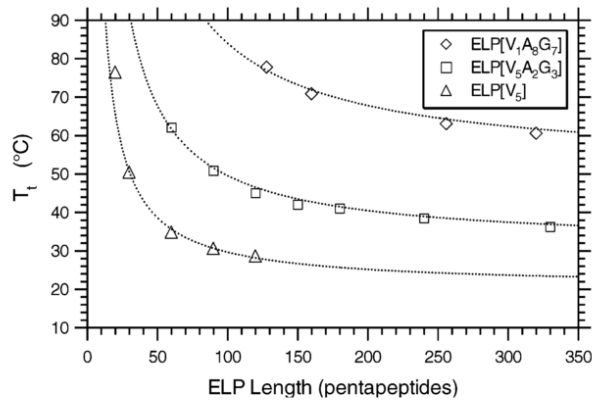
Figure 2.1 Transition temperature, T_t, as a function of ELP chain length for three different ELP libraries[15].

The Tt asymptote is at 24°C, which is too low, as cultivation is usually done at about 30°C. However, this Tt occurs at 120 repeats of the pentapeptide, whereas reducing the number of repeats exponentially increases the Tt. To obtain the most functional transition temperature, valine was selected as the guest residue. Valine was also predicted to be successful based on its hydrophobicity, as shown in previous studies[12]. A small library of varying pentapeptide repeats was experimented with to optimize ELP length to balance expression and Tt, as addressed in 2.3.1.2. But the initial gene was designed with only 20 pentapeptide repeats, and thus is designated RSL-ELP20.

The next section designed was a spacer or linker region between the two functional domains. In fusion protein architecture, it is often crucial to incur some distance between two domains, so misfolding or interference does not occur[16]. However, the linker itself can often restrict functionality, and varying lengths and flexibility must be investigated. A popular flexible peptide linker often used is a glycine rich linker[17]. Due to the lack of R group on the C_{alpha}, there is increased flexibility. Different sized

51

cassettes have been studied at length and seem to require a case by case analysis. However, some consistencies have emerged. An average linker length between averaged sized protein domains is typically greater than 10 and less than 60 residues[18]. With this in mind, a glycine rich linker of (GGGGS)$_2$ was chosen to span the distance between the two functional domains. This putatively allowed enough spacer between the two relatively small domains, while keeping the total size of the fusion protein to a minimum.

Fusion order was a nontrivial decision made in the design as well. As fusion order can control solubility, folding, and expression of recombinant proteins, it was investigated as well. In a study done by Christensen, ELP fusions are designed as N- and C- terminal fusions. In every case, the C-terminal fusion of the ELP was expressed to a higher level with higher probability of correct folding than that of the N-terminal ELP fusion. It is speculated that this is due to lower levels of mRNA in the ELP to protein direction. Intracellular degradation of the translated ELP portion is also a possible scenario resulting in decreased expression[19]. In any case, it was decided that the ELP would be fused to the C-terminus of the lectin for optimal expression.

To enable secretion, an N-terminal signal sequence was cloned before the lectin sequence. This is further detailed in **Section 2.3.2.1.**

Finally, the construct was adorned with an N-terminal His-tag marker after the signal sequence before the lectin. This was done as an additional method of purification to confirm expression before functionality is confirmed. It is also to enable near complete capture so expression levels could be quantified. The final gene design is depicted by Figure 2.2.

Tat	6xHis tag	RSL	Linker	ELP20	

Figure 2.2 Gene design for RSL-ELP20 fusion. Blank boxes represent restriction sites.

CLONING

The gene was directionally cloned into a pQE80L derivative using sites BamHI and HindIII in the multiple cloning site. The plasmid map is shown in Figure 2.3. The pQE80L vector is a high copy number plasmid with an ampicillin resistance marker and a multiple cloning site with an N-terminal 6xHis tag under the control of an inducible T5 promoter. The resulting plasmid was designated pQE20.
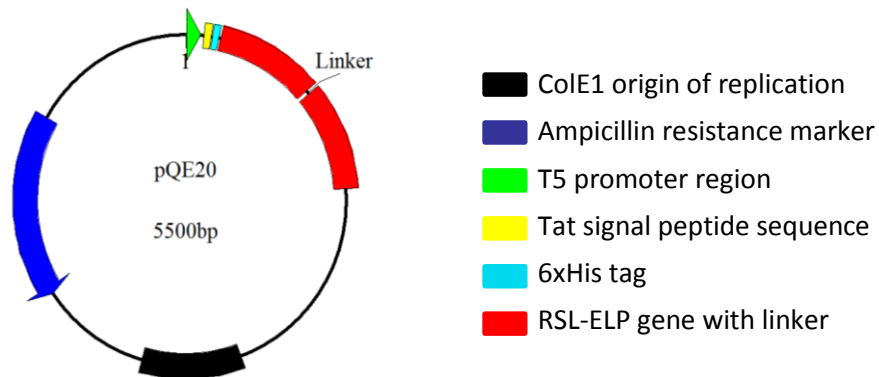
Figure 2.3 Plasmid map of designed plasmid pQE20.

EXPRESSION

After cloning of the synthetic gene into the plasmid was complete, the plasmid was transformed into a cloning host, *E. coli* K12 UT5600. This strain, designated UT5600/pQRE20, was cultured. To induce expression, the procedure described in

Materials and Methods was followed. After allowing the culture to reach exponential phase, expression was induced by addition of isopropyl-beta-D-thiogalactopyranoside (IPTG) to the culture medium. This was then allowed to cultivate for 24-48 hours. The presence of the protein was confirmed by SDS-PAGE of the cell lysate, with broad bands identifying the expressed recombinant proteins, as visualized in Figure 2.4. The apparent sizes of the ELP containing constructs were larger than the actual molecular weights, a result noted in other research[20,21]. This disparity is likely due to the extended chain length of the ELP absent of any charged residues, which results in slower gel migration during electrophoresis. It's also important to note that the RSL in its native form associates into a trimer formation. A unique band 2 or 3 times the size of the desired protein was also investigated, as trimer or dimer formation was likely. Trimer association was optimal, as the maximum number of fucose binding sites were formed with full homotrimer association.
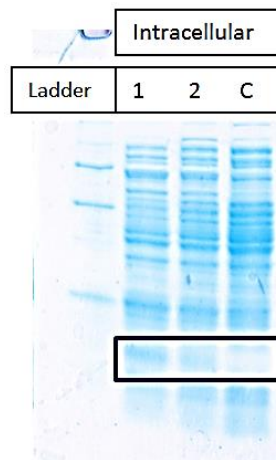
Figure 2.4 Two samples expressing RSL-ELP20 (Lane 1 and 2) alongside a control containing an empty plasmid (Lane C).

2.3.1.2 RSL-ELP40 and RSL-ELP60

After successful design, cloning, and expression of the first construct, fusion proteins of similar yet altered structure could be pursued in hopes of optimizing the expression and functionality. As discussed in Section 2.3.1.1, increased ELP length reduces the phase transition temperature, so fusion proteins with increased ELP length were cloned and investigated.

CLONING

In order to create this library of varying ELP length, a repetitive cloning technique was incorporated into the initial design of the gene, as is often done to create recursive ELP segments[22,23]. By using restriction sites with compatible sticky ends, such as XhoI and SalI, the ELP region was clipped out of the gene and isolated by gel extraction. PCR was not utilized, as the GC rich region required for the pentapeptide repeats in the ELP made primer specification and amplification difficult. Instead, large amounts of vector were extracted via Maxi-Prep kits, and multiple digestions performed to isolate the required amount of ELP segment (about 300 base pairs in length). By digesting the host vector at only site SalI, a non-directional ligation occurred at the sites of the matching sticky ends comprised of TCGA/AGCT. A schematic of this recursive cloning method is illustrated below in Figure 2.5.

1.

RSL	Linker	SalI	ELP20	XhoI

2.

TCGA	ELP20	AGCT

TCGA	ELP20	AGCT

TCGA	ELP20	AGCT

3.

RSL	Linker	SalI	ELP20	XhoI

4.

RSL	Linker	SalI	ELP20	-	ELP20	XhoI
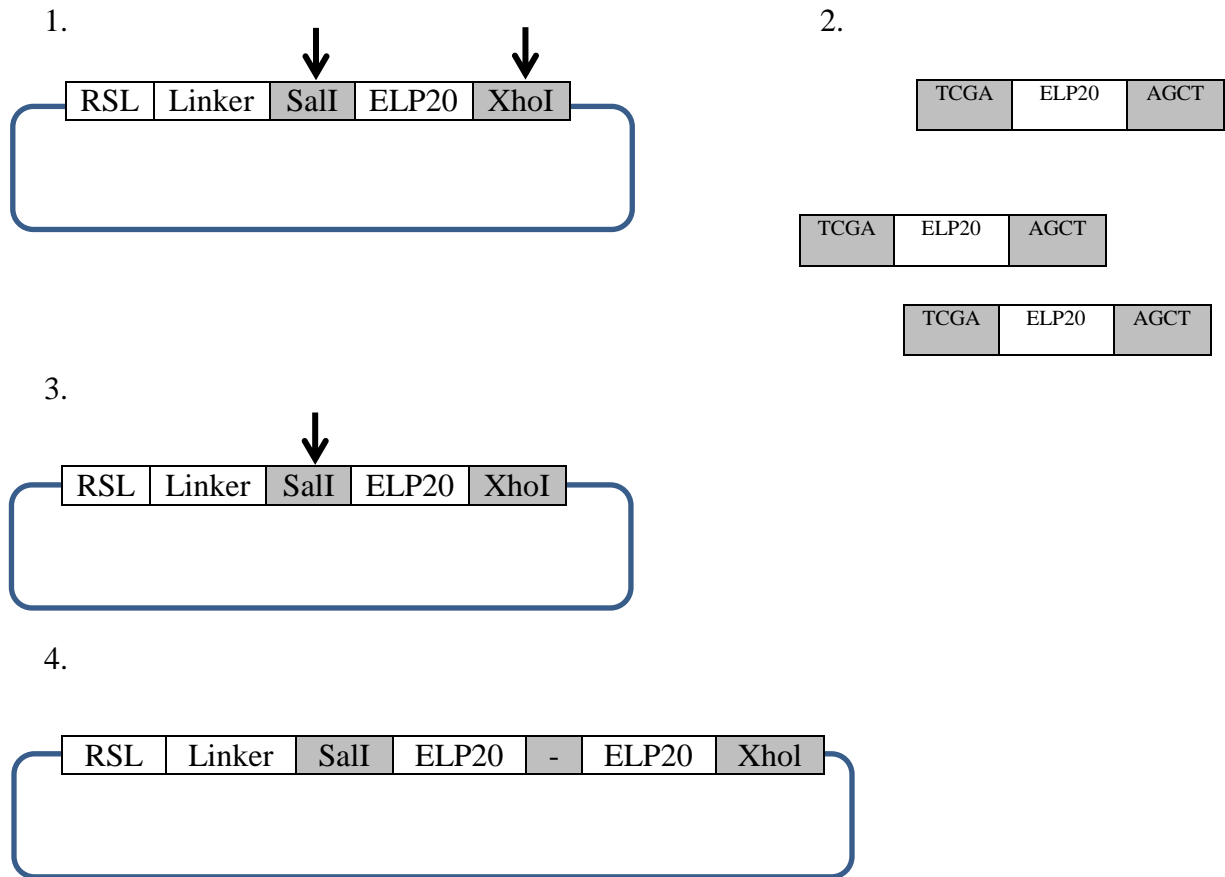
Figure 2.5 Schematic of oligomerization of ELP gene. Step 1: Plasmid is digested at indicated sites. Step 2: ELP20 containing fragments are collected, overhanging nucleotides are included in gray. Step 3: Original plasmid is digested only at SalI. Step 4. ELP20 fragments collected in Step 2 are ligated into the SalI site.

As the ligation does not have a directional preference, at least 50% of the successfully ligated clones should be in the correct direction. This must be verified by sequencing or protein activity, as the sizes will be identical. Ligation at the mismatch sites of SalI of the vector and XhoI of the insert forms a non-restriction site combination of the two, thus successive cloning of longer ELP regions is possible at the single complete SalI site. The initial synthetic gene codes for 20 repeats of the pentapeptide, while cloning of repetitive regions allows for a library that includes repeats in multiples

56

of 20. Only 20, 40, and 60 ELP pentapeptide repeats were cloned, as longer polymers did

not aid in significantly decreasing the transition temperature after 60 repeats.

EXPRESSION

To confirm the elongation of the ELP domain, the gene was expressed and the

protein identified for validation. As with the RSL-ELP20, the plasmid containing the

RSL-ELP40 and RSL-ELP60 inserts were transformed into a cloning host UT5600. The

transformation procedure is described in Materials and Methods. The colonies were

screened and the potential candidates cultured. The cultures were induced with IPTG and

cultured for 24-48 hours. The presence of the proteins was confirmed by SDS-PAGE of

the cell lysate, with broad bands identifying the expressed recombinant protein. The

RSL-ELP40 construct has large band a little below 50 kDa, as seen in the crude fraction

SDS-PAGE in Figure 2.6a. Figure 2.6b is also the crude intracellular protein, but with a

large band displaying expression of RSL-ELP60, at about 60 kDa.
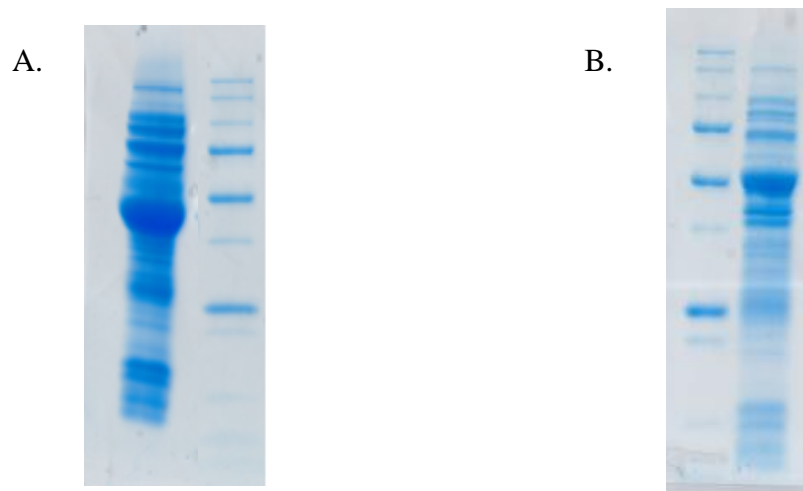
Figure 2.6 Crude cell lysate showing intracellular expression of

A) RSL-ELP40 and B) RSL-ELP60

Cloning and Expression of VCNA-ELP40

   To expand upon the initial design even further, a different lectin was employed for the fusion design. By using the lectin with affinity to a different carbohydrate moiety, a broader range of application was achievable. This also sets the stage for further development using a variety of lectins for the thermo responsive affinity design.

CLONING

   After successful cloning of the fucose binding ligand, an alternative ligand with affinity to sialic acid was investigated. To target sialic acid, a binding site from *Vibrio cholera* was isolated from a neuraminidase enzyme, NanH. The binding site is comprised of 191 amino acids, requiring 573 codons for translation. This lectin, named VCNA, was cloned for a VCNA-ELP40 fusion. The majority of work was done with the 40 repeat fusion of the ELP domain, given the suitable transition temperature and reasonable expression level, detailed in Chapter 3. Due to this success, only the ELP40 fusion was pursued.

   The NanH enzyme has three primary domains. The N-terminal domain is comprised of a sialic acid binding lectin. This gene encoding this section was optimized for *E. coli* and synthesized by GenScript.

   The primers were designed to transcriptionally fuse the desired linker to C-terminus of the lectin portion before cloning into the ELP40 containing vector. During

PCR amplification of the synthetic gene, the linker was added, so the final insert included both the lectin and linker, as seen in Figure 2.7.

| Tat | | 6xHis tag | | VCNA | Linker | | ELP40 | | |

Figure 2.7 Gene design for VCNA-ELP40 fusion. Blank boxes represent restriction sites.

EXPRESSION

As with the other constructs, the plasmid containing the fusion insert was transformed into cloning host UT5600. Colonies were selected by screening in the presence of ampicillin, and the plasmid size in these putative colonies was confirmed. These colonies were then cultivated and induced to express the protein of interest.

In the case of VCNA expression, it was noted in previous literature that a heat shock induction method was used in tandem with traditional IPTG induction. Although the protein band of the fusion protein was visible, attempts of purification using ITC to exploit the functionality of the ELP were unsuccessful. It was hypothesized that poor folding of the VCNA domain resulted in misfolding of the ELP region. In order to rectify the issue and aid in correct folding, the heat shock method, as described in Moustafa, was used at the time of induction[24]. After allowing the culture to progress to an OD_{600} of 0.4, the culture flask is removed from the 37˚C incubator and submerged in a 42˚C bath. The culture is subject to this higher temperature for 20 minutes. This short period of heat shock likely expresses additional chaperonines which aid in folding of the

59

recombinant protein. After the heat shock, the culture is returned to room temperature for 10 minutes before IPTG is added.

2.3.2 Secretion of RSL-ELP Constructs

In order to become a cost effective option, the construct must be able to be resourcefully and economically produced and isolated. By utilizing the fast growing and easily manipulated host, *E. coli*¸ the fusion protein can be expressed. By utilizing translocation pathways and increased outer membrane permeability, the fusion protein can be secreted, and thus isolated by exploiting the thermo responsive nature of the construct. This avoids inclusion bodies, sonication steps, and harmful protease activity. Two different secretion pathways were investigated to find the optimal production level.

2.3.2.1 Tat Pathway

After confirmation of expression, the construct containing plasmids were transformed into *E. coli* E609Y, a mutant strain of E609 containing a deletion of the *lpp* encoded lipoprotein. This mutant strain confers a leaky outer membrane phenotype. When combined with a translocation pathway, such as the Tat dependent pathway, extracellular secretion is achieved[25]. All constructs were designed to include the N-terminal Tat signal sequence.

E609Y cultures containing plasmids with the constructed genes were induced and incubated at 18°C. Cells and extracellular milieu were harvested over the course of the 48 hour cultivation. The intracellular, periplasmic, and extracellular fractions were separated and analyzed to illustrate the localization of the constructs during cultivation and the secretion process. 12 hours after induction, proteins secreted by the strain were

present in the extracellular fraction, albeit in low quantities. Each sample had secreted approximately 20 mg/L to 40 mg/L, as seen in Figure 2.8. After SDS-PAGE analysis, displayed in Figure 2.9a, it was noted no desired protein was found in the extracellular portion at this time point. However, by 24 hours, all constructs excluding the RSL-ELP60 construct were discernible by SDS-PAGE in the extracellular portion. This is observable in Figure 2.9a. These bands indicating secretion continue to broaden over the course of the 48 hour cultivation, detailing an increased amount of protein secretion over time, as seen in Figure 2.9b.
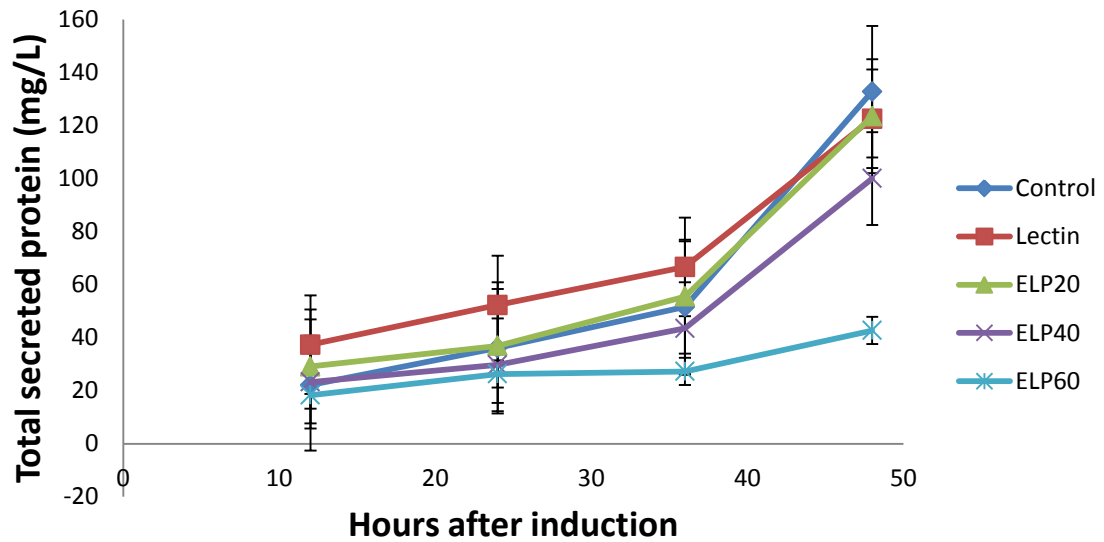
Figure 2.8 Total extracellular protein content as a function of time after induction.

61

A)   B)
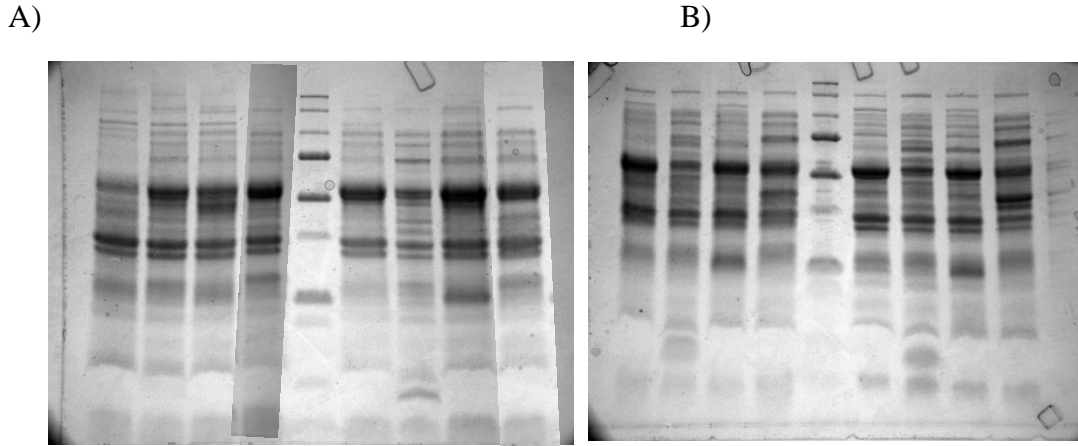
Figure 2.9 Extracellular protein profile at A) 12 and 24 hours B) 36 and 48 hours

As plotted in Figure 2.10, the amount of secreted recombinant protein reaches 35 mg/L of culture for the smallest construct, the lectin with no fusion. The RSL-ELP20 and RSL-ELP40 constructs are secreted at levels of 30 mg/L and 18 mg/L, respectively, over the 48 hour incubation period. No RSL-ELP60 was detectable in the extracellular milieu after 48 hours of cultivation.
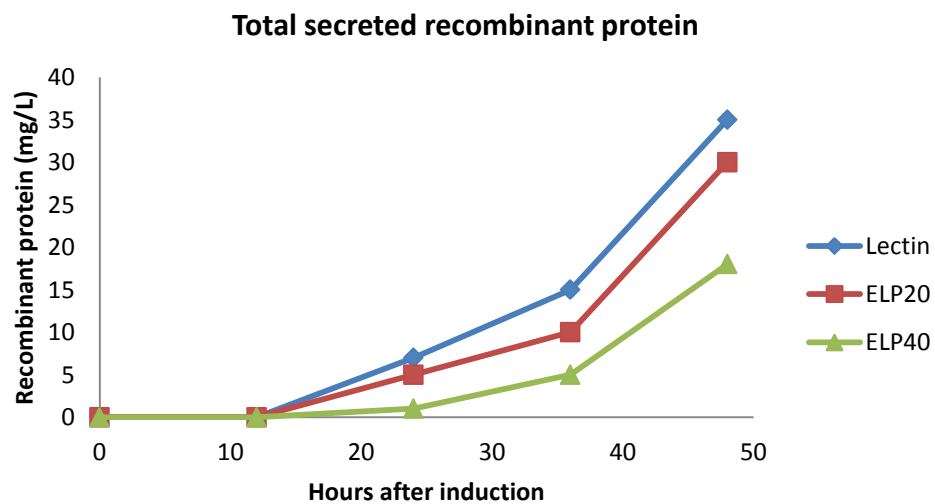
Figure 2.10 Secreted recombinant protein as a function of time after induction

<u>2.3.2.2 Sec Pathway</u>

In an effort to optimize extracellular secretion in a biological host, a different secretion pathway was also pursued. It was a possibility that the translocation of a preprotein fusion would be more efficient than that of a folded and active protein. With this in mind, that Sec pathway for secretion was investigated.

First, the fusion gene was digested from the pQTH vector at the BamHI and HindIII sites, comprising the entire gene sequence minus the Tat N-terminal signal sequence. This was ligated into the pET20b+ vector at the same restriction sites to fuse to the Sec signal sequence. This new vector was transformed into *E. coli* BL21star(DE3), compatible with the T7 promoter. The screened clones were induced, cultured, and fractionated to localize expression.

These cultures were fractionated along with a control strain containing an empty plasmid. Each localized region of three lanes in the SDS-PAGE below (Figure 2.11) contains first the control (pET20b+), followed by two separate cultures of the clone (pETRSL). The expression seems to be localized periplasmically; however, there doesn't seem to be a difference between the control and the two clones. Little to no expression is visible. Although it's possible the dark band seen above the 25 kDa marker is the target fusion protein, further characterization did not confirm. No distinct characterizing band was seen in the other fractions.
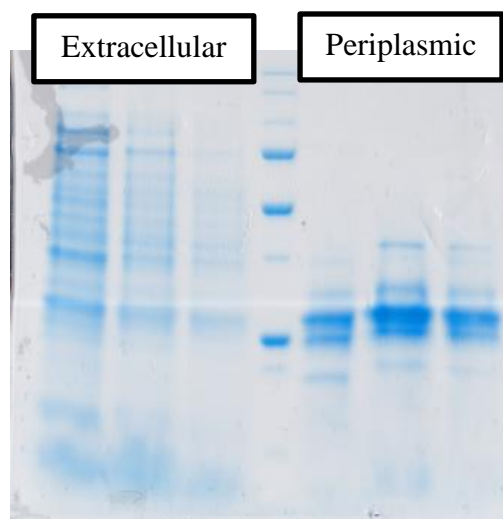
Figure 2.11 Secreted protein profile via Sec pathway. Samples are control plasmid containing strain followed by 2 pETRSL containing strains.

In an effort to possibly mediate or increase expression, I also transformed the cloned pETRSL vector into BL21(DE3)/pLysS. I also compared the expression of my protein in BL21star(DE3) to expression in BL21(DE3)/pLysS. The addition of the pLysS vector allows tighter control over gene induction and expression. Figure 2.12 is of the periplasmic space of each of these two strains plus a control.

Figure 2.12 Periplasmic protein profile from Sec pathway combined with pLysS vector.
Lane 1: MW Ladder Lane 2: BL21(DE3)/pLysS/pETRSL
Lane 3. BL21star(DE3)/pETRSL

Again, although there appeared to be a prominent band in the periplasm, the target was *only* localized to the periplasm, yielding very low expression when compared to the extracellular secretion of the Tat secretion pathway. The pLysS plasmid attempt seemed successful, but the success of this pathway is overshadowed by the efficiency of the other secretion pathway.

2.3.2.3 Cell Localization

In using this relatively new method of protein secretion, it was important to monitor the location of the recombinant proteins in the process. At this 12 hour time point, all desired proteins were localized to the periplasm, as noted in Figure 2.13. As much as 4 mg/L of desired protein was localized to the periplasm in the case of RSL-ELP20, with slightly lower levels of the other constructs, and a negligible amount of the largest construct, RSL-ELP60. Intracellular localization of each construct was negligible

at 12 hours as well. While levels of RSL, RSL-ELP20, and RSL-ELP40 in the periplasm either decreased or remained relatively constant, SDS-PAGE analysis displays periplasmic build-up of RSL-ELP60, as seen in Figure 2.14. Intracellular levels of all constructs were negligible except in the case of RSL-ELP40. After 48 hours, it appeared as if a detectable (< 1 mg/L) of RSL-ELP40 was still localized to the cytoplasm. The majority of this construct was detected as a band at about 150 kDa, which identifies with the trimer configuration of the RSL-ELP40 construct.
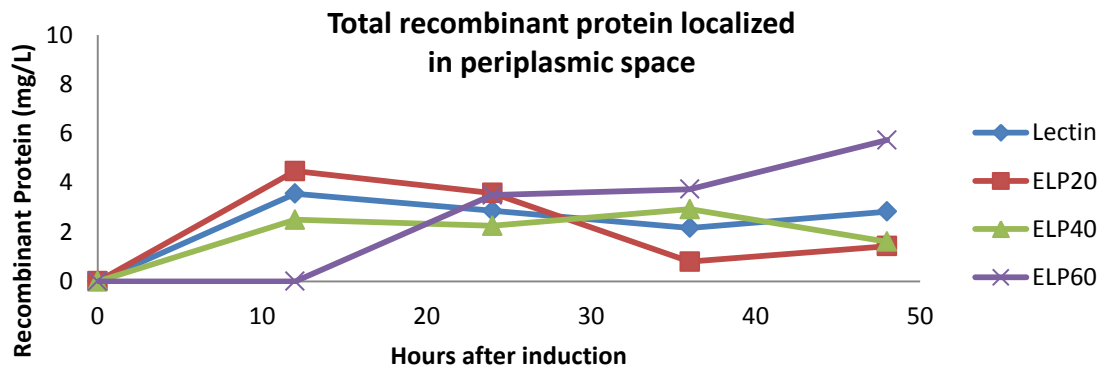
Figure 2.13 Recombinant protein localized to periplasm as a function of time.
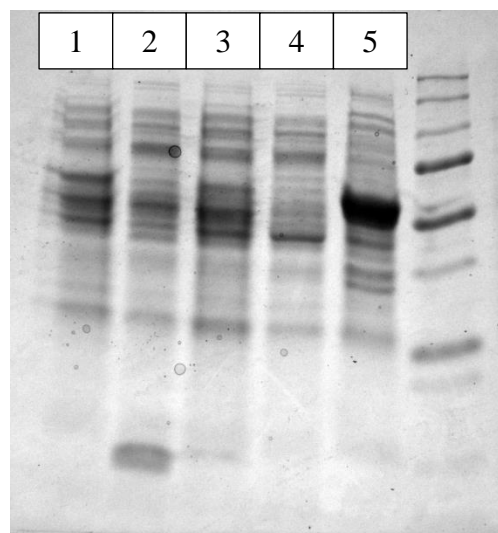
2.3.3 Secretion of VCNA-ELP Construct

Based on the success of the Tat pathway in the E609Y strain, the production of the VCNA-ELP fusion construct was done solely using this method of extracellular expression. Initially, a cultivation temperature of 18˚C was used, as this was the primary temperature used for the RSL-ELP expression. There was concern that too high a cultivation temperature may cause the ELP *in vivo* to fold and precipitate, facilitating early inclusion bodies that are not translocated. Two different IPTG concentrations were used to induce, and the total protein and recombinant protein levels were monitored at 24 and 48 hours.

Table 2.1 Protein levels secreted by E609Y at 18˚C cultivation.

A. Total Protein (mg/L)		
	0.1 mM IPTG	1.0 mM IPTG
24 hours	374	217
48 hours	510	331

B. VCNA-ELP40 (mg/L)		
	0.1 mM IPTG	1.0 mM IPTG
24 hours	75	43
48 hours	102	66

With promising results of about 100 mg/L being produced, a similar cultivation was attempted at a higher temperature, hoping to promote cell growth, though possibly at the expense of protein production. The same parameters were observed, though at a higher temperature of 30˚C after induction. The results are summarized in Table 2.2 a and b.

Table 2.2 Protein secreted by E609Y at 30˚C cultivation

A. Total Protein (mg/L)		
	0.1 mM IPTG	1.0 mM IPTG
24 hours	312	144
48 hours	340	281

B. VCNA-ELP40 (mg/L)		
	0.1 mM IPTG	1.0 mM IPTG
24 hours	62	29
48 hours	68	56

As visible from the quantification of both total protein and recombinant fusion protein, the increased temperature did nothing to improve the production levels. It was concluded that a 48 hour, 18˚C cultivation after induction with 0.1 mM IPTG produced the best production levels of the VCNA-ELP40 fusion construct in E609Y.

2.3.4 Endotoxin Presence

By secreting the construct, it was hypothesized that avoidance of cell lysis would result in decreased endotoxin presence in the protein fraction. Each strain (Y type and wild type) was cultivated according to the determined optimal conditions from the

previous experiment. The working solution containing the recombinant RSL was measure for endotoxin units per volume (EU/mL) using a ToxinSensor Chromogenic LAL Endotoxin Assay Kit. For the E609, this was the cell lysate, for the E609Y, it was the extracellular portion. Although the E609Y measured 13000 EU/mL less than the cell lysate solution, the total units of endotoxin were still in the same order of magnitude for each sample, as seen in Table 2.3.

Each of these was then purified via His-tag affinity gel. The purification yielded a large reduction of endotoxin in solution. However, again, while the Y-type strain resulted in a lower EU/mL, the result is not drastically different. The benefits with respect to reduced endotoxin levels were non-conclusive.

Table 2.3 Endotoxin levels in protein fraction and after His-tag purification

	Total		Histag Purified	
Strain	**E609**	**E609Y**	**E609**	**E609Y**
EU/mL	277000 ± 2200	264000 ± 2100	1640 ± 40	1220 ± 20

2.4 Discussion

In this study, lectin ELP fusions of varying ELP length were cloned, expressed, and secreted from the *E. coli* host. These novel recombinant fusions were characterized and the dual functionality of the fusion was manipulated to illustrate a glycoprotein purification process using affinity precipitation.

While affinity precipitation is not a new concept, the ability to produce an active thermo-responsive affinity ligand entirely by a bacterial host is a valuable concept

introduced in the study. No chemical coupling methods are required as in the case of Eudragit S-100 or other synthetic polymers[26-28]. Avoiding the use of solvents and eliminating the loss of ligand due to incomplete coupling are beneficial for production on the industrial scale, and production in E. coli is a reputable method of production in industry today. Extracellular secretion from this versatile host is also beneficial on the industrial scale for ease of isolation. No large scale homogenization or cell lysis step is required to attain the ligand. Secretion also aids in avoiding periplasmic located proteases, which minimizes proteolytic degradation and eliminates the requirement for protease inhibitors. Production of this purification mechanism bacterially also greatly diminishes the cost of current affinity purification systems, such as affinity chromatography. Avoiding expensive resins that require regeneration and that are easily fowled is a large cost benefit in downstream purification processes.

However, while this secretion mechanism may be ideal for smaller constructs, it appears not to favor highly hydrophobic sequences. The largest construct, RSL-ELP60, was not secreted to the extracellular milieu. It is not likely a size limitation of the secretion mechanism, as the band pattern of the extracellular portion in SDS-PAGE analysis shows bands up to about 120 kDa in weight. More likely, the halted secretion and localization to the periplasm of this construct is indicative to its higher hydrophobicity of the (VPGVG)60 sequence. Active translocation of the fully formed and folded protein via the Tat dependent pathway is functional, as portrayed by the large build-up of RSL-ELP60 in the periplasm. But the diffusion portion of the secretion pathway exploiting the leaky outer membrane is less amenable for highly hydrophobic proteins.

70

While the Tat dependent translocation does not seem to discriminate against more hydrophobic proteins, as seen by successful translocation of RSL-ELP60, it does seem to have limitations. While RSL-ELP40 is successfully translocated and secreted in desirable levels, there is still a detectable amount found within the cytoplasm of the cell. This is particularly noticeable as a lone band of about 120 kDa is detectable, illustrating the presence of the RSL-ELP40 trimer construct. At first glance, the residual RSL-ELP40 in the cytoplasm in trimers appears to be left behind due to size limitations of the translocation pathway. However, the largest translocated heterologous protein in Streptomycetes via Tat is 146 kDa, while E. coli is generally accepted to translocate 10-100 kDa using this pathway[29,30]. The trimer formation is 84 kDa in actuality, which is within the range of known translocated proteins. However, upon forming the trimer structure, it is possible the N-terminal signal sequence required for translocation may be shielded, leaving these trimers to remain in the cytoplasm.

Finally, a variety of conditions were tested in order to improve the secretion and production levels of both of the fusion constructs. Using this method, production of about 100 mg/L has been reached, which is a little low for industrial production cases. However, with some improvement, this platform still presents as a cost effective way to produce ligand for affinity-based purification.

2.5 Materials and Methods

Bacterial Strains and Plasmids

Escherichia coli strain K12 UT5600 (*F- ara-14 leuB6 secA6 lacY1 proC14 tsx-67 Δ(ompT-fepC)266 entA403 trpE38 rfbD1 rpsL109 xyl-5 mtl-1 thi-1*) was used as a host for all cloning steps. Plasmid pQTAT containing an N-terminal TAT signal sequence was used as the host vector. The fusion gene was synthesized by GenScript, and consisted of the lectin fused to 20 VPGVG repeats. A 10 amino acid glycine rich linker was used to provide flexibility between the lectin and elastin fusion.

Construction of Recombinant Plasmids and ELP extension

The synthetic gene was digested from the carrier plasmid with BamHI and HindIII restriction enzymes (NEB). It was inserted in the multiple cloning site of pQTH at these same restriction sites, and ligated to form pQRE20. A large amount of this plasmid was then cultivated and digested at sites SalI and XhoI to isolate the fragment of gene encoding 20 repeats of VPGVG. The host vector, pQRE20, was then digested at SalI, revealing sticky ends compatible with both XhoI and SalI restriction sites of the insert. The product was plasmid pQRE40, a fusion of a lectin to 40 repeats of VPGVG. This was repeated to produce plasmid pQRE60.

Protein Expression

E. coli strain E609Y[25] was transformed by heat shock at 42˚C for 1 minute followed by a regeneration incubation in LB media at 37˚C for 1 hour. The resulting culture was plated on LB agar containing 100 µg/mL of ampicillin as a selection marker.

A single colony with the plasmid carrying the synthetic gene was selected and cultured in 3 mL LB broth and ampicillin overnight at 37 C and 250 RPM. This seed culture was used to inoculated 100 mL LB broth culture containing ampicillin to an OD_{600} of 0.1. This larger culture was grown at 37C and 250 RPM to a cell density of OD_{600} of about 0.4, at which point isopropyl-β-D-thio-galactoside (IPTG) was added to the culture to a final concentration of 1mM to induce the production of the fusion protein. The cultivation temperature was then reduced to 18˚C and the cells were grown for an additional 48 hours.

Heat Shock Induction

Cultures were grown first at 37˚C after seeding. Upon reaching an optical density of 0.5-0.6 at OD_{600}, the flasks were subjected to a heat shock of 42˚C for 20 minutes. The cultures were then removed and allowed to cool to room temperature for 10 minutes before the addition of IPTG to initiate protein expression. The cultivation temperature was then reduced to 30˚C and the cells were grown for an additional 48 hours[24].

Fractionation

The extracellular fraction was collected as the supernatant after pelleting the cells of 10 mL of culture media. This cell removal was done with centrifugation at 6,000g at 4˚C for 25 min. The extracellular milieu was concentrated for further analysis using centricones and centrifugation at 4,000g for 10 minutes. This was repeated until a workable concentration was reached. Periplasmic and cytoplasmic fractions were prepared according to the method described in the pET system manual (EMD Chemicals,

San Diego, CA). A ten milliliter culture was harvested by centrifugation and re-suspended in 5 ml of shock buffer solution comprised of 30 mM Tris–HCl, pH 8.0, 20% sucrose, and 1mM EDTA. The cell suspension was incubated at room temperature for 10 min and pelleted by centrifugation at 6,000g at 4˚C for 10 min. The cells were re-suspended in 0.5 ml ice-cold 5 mM $MgSO_4$ and incubated on ice for 10 min. The cells were pelleted by centrifugation as before, and the supernatant was collected as a periplasmic fraction. The pellet was re-suspended in 1 ml phosphate buffered saline (PBS) solution, sonicated, and centrifuged at 6,000g at 4˚C for 20 min. The supernatant was saved as a cytoplasmic fraction. Each fractionated sample was analyzed using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) or enzymatic activity assays.

Isolation with Affinity Chromatography

To ensure lectin binding activity, the fusion protein was purified using affinity chromatography with a Fucose Separopore® (Agarose) matrix, obtained from bio-WORLD. The crude protein solution was added to this matrix and equilibrated with PBS. This mixture was kept on ice on an orbital shaker at 60 RPM overnight to allow for binding. The unbound fraction was washed from the column with PBS, and the bound fraction was eluted from the column with an elution buffer of 0.1M fucose in PBS. Protein purity was assessed by SDS-PAGE.

Isolation with Inverse Temperature Cycling (ITC)

The solution was heated to the transition temperature of the ELP in an Eppendorf Thermomixer for 20 minutes at 300 RPM. The solution was then centrifuged at 16,000 x g for 1 minute. The supernatant was removed, and the pellet resuspended in PBS or PBS with 0.5M NaCl. The resuspended solution was chilled at 4˚C overnight to maximize resolubilization of the ELP constructs. This chilled solution was centrifuged again for 5 min at 16,000 x g for 5 minutes to remove unspecific precipitates, and the supernatant was collected as the purified protein fraction.

Protein Quantification

Protein concentration was measured by the method described by Bradford using a protein reagent dye (Bio-Rad), with bovine serine albumin as the standard.

SDS-PAGE Analysis

Each sample was combined in a 1:1 ratio with SDS sample buffer (10% SDS, 10% b-mercaptoethanol, 0.3 M Tris–HCl (pH 6.8), 0.05% bromophenol blue, 50% glycerol), boiled for 5 min, and resolved by 12.5% (w/v) SDS–PAGE. Each gel was detected by Coomassie blue staining (Bio-Rad).

2.6 References

1.   Franken, K. L. et al. Purification of his-tagged proteins by immobilized chelate affinity chromatography: the benefits from the use of organic solvent. Protein Expr. Purif. 18, 95–9 (2000).

2.   Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr. Purif. 48, 1–13 (2006).

3.   Hengen, P. Purificatin of His-Tag fusion proteins from Escherichia coli. Trends Biochem. Sci. 20, 285–285 (1995).

4.   Chow, D. C., Dreher, M. R., Trabbic-Carlson, K. & Chilkoti, A. Ultra-high expression of a thermally responsive recombinant fusion protein in E. coli. Biotechnol. Prog. 22, 638–46 (2006).

5.   Sudakevitz, D. et al. A new Ralstonia solanacearum high-affinity mannose-binding lectin RS-IIL structurally resembling the Pseudomonas aeruginosa fucose-specific lectin PA-IIL. Mol. Microbiol. 52, 691–700 (2004).

6.   Kostlanova, N. et al. The fucose-binding lectin from Ralstonia solanacearum - A new type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. J. Biol. Chem. 280, 27839–27849 (2005).

7.   Nakano, M., Kondo, A., Kakehi, K. & Taniguchi, N. Glycomics – a new target for pharmaceuticals. Drug Discov. Today Technol. 3, 39–47 (2006).

8.   Becker, D. J. & Lowe, J. B. Fucose: biosynthesis and biological function in mammals. Glycobiology 13, 41R–53R (2003).

9. Rakha, E. a et al. Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. Mod. Pathol. 18, 1295–304 (2005).

10. Park, S. et al. marker for colon cancer : a combinatorial approach for development of glycan biomarkers. 2376, 2366–2376 (2012).

11. Gomord, V. et al. Production and glycosylation of plant-made pharmaceuticals: the antibodies as a challenge. Plant Biotechnol. J. 2, 83–100 (2004).

12. Urry, D. W. et al. HYDROPHOBICITY SCALE FOR PROTEINS BASED ON INVERSE TEMPERATURE TRANSITIONS. Biopolymers 32, 1243–1250 (1992).

13. Luan, C. H., Parker, T. M., Gowda, D. C. & Urry, D. W. Hydrophobicity of amino acid residues: differential scanning calorimetry and synthesis of the aromatic analogues of the polypentapeptide of elastin. Biopolymers 32, 1251–61 (1992).

14. Meyer, D. E., Trabbic-Carlson, K. & Chilkoti, a. Protein purification by fusion with an environmentally responsive elastin-like polypeptide: effect of polypeptide length on the purification of thioredoxin. Biotechnol. Prog. 17, 720–8 (2001).

15. Meyer, D. E. & Chilkoti, A. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. Biomacromolecules 5, 846–851 (2004).

16. Wriggers, W., Chakravarty, S. & Jennings, P. A. Control of protein functional dynamics by peptide linkers. Biopolymers 80, 736–746 (2005).

17. Lu, P. & Feng, M. G. Bifunctional enhancement of a beta-glucanase-xylanase fusion enzyme by optimization of peptide linkers. Appl. Microbiol. Biotechnol. 79, 579–587 (2008).

18. Robinson, C. R. & Sauer, R. T. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. Proc. Natl. Acad. Sci. U. S. A. 95, 5929–34 (1998).

19. Christensen, T. et al. Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. Protein Sci. 18, 1377–87 (2009).

20. McPherson, D. T., Xu, J. & Urry, D. W. Product purification by reversible phase transition following Escherichia coli expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP. Protein Expr. Purif. 7, 51–57 (1996).

21. Shimazu, M., Mulchandani, A. & Chen, W. Thermally triggered purification and immobilization of elastin-OPH fusions. Biotechnol. Bioeng. 81, 74–9 (2003).

22. Meyer, D. E. & Chilkoti, A. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. Biomacromolecules 3, 357–67 (2002).

23. Straley, K. S. & Heilshorn, S. C. Design and adsorption of modular engineered proteins to prepare customized, neuron-compatible coatings. Front. Neuroeng. 2, 9 (2009).

24. Moustafa, I. et al. Sialic acid recognition by Vibrio cholerae neuraminidase. J. Biol. Chem. 279, 40819–26 (2004).

25.   Shin, H.-D. & Chen, R. R. Extracellular recombinant protein production from an Escherichia coli lpp deletion mutant. Biotechnol. Bioeng. 101, 1288–96 (2008).

26.   Hilbrig, F. & Freitag, R. Protein purification by affinity precipitation. J. Chromatogr. B 790, 79–90 (2003).

27.   Hilbrig, F., Stocker, G., Schläppi, J.-M., Kocher, H. & Freitag, R. Utilization of Group Specific Ligands in the Downstream Processing of Proteins by Affinity Precipitation. Food Bioprod. Process. 84, 28–36 (2006).

28.   Taipa, M. a, Kaul, R. H., Mattiasson, B. & Cabral, J. M. Recovery of a monoclonal antibody from hybridoma culture supernatant by affinity precipitation with Eudragit S-100. Bioseparation 9, 291–8 (2000).

29.   Widdick, D. a et al. The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. Proc. Natl. Acad. Sci. U. S. A. 103, 17927–32 (2006).

30.   Müller, M. Twin-arginine-specific protein export in Escherichia coli. Res. Microbiol. 156, 131–6 (2005).

CHAPTER III

AFFINITY LIGAND CHARACTERIZATION

3.1 Abstract

These novel recombinant fusions were characterized and the dual functionality of the fusion was manipulated to illustrate a glycoprotein purification process using affinity precipitation. The preserved functionality of the fusion protein was tested, along with varying parameters to ensure robustness. The RSL-ELP fusions exhibited phase transitions at 80°C, 45°C, and 36°C respectively to 20, 40, and 60 repeats of ELP pentapeptide. The VCNA-ELP40 fusion exhibited phase transition at 51°C. Addition of salt resulted in depressed transitions temperatures. In the case of the RSL fusion, trimer formation needed to be investigated. As formation of the trimer complex doubles the number of binding sites, it was crucial to ensure heating, salt, and other external factors used in the purification process did not disrupt trimerization. The sensitivity of trimerization was tested and deemed to not be easily susceptible to dissociation. The bi-functionality of this affinity precipitation tool is illustrated by purification of a model glycoproteins, horseradish peroxidase (HRP) and fetuin, using a simple method of temperature cycling and centrifugation. Furthermore, the fusion protein is tested against high contaminant solutions and competing glycoproteins to test its robustness. Using this purification scheme, $98 \pm 2.8\%$ recovery of the glycoprotein is achieved while maintaining activity of the purified model glycoprotein. $90 \pm 8.6\%$ recovery is achieved in the case of high contaminant solutions.

3.2 Introduction

Fusion proteins have long been produced via genetic engineering or chemical conjugation techniques. Often, the fusing of proteins is mean to exploit bifunctionality of the final product, possibly having a synergistic effect or additional benefits. Sometimes fusion is done simply as a method to determine location, study conformational aspects, or elucidate evolutionary traits[1,2]. In any of the possible cases, it is crucial that the designed protein not lose function due to the fusion.

More specifically, ELP's have been chemically conjugated to various different targets, resulting in ELP-drug fusions[3-5]. In most cases, these fusions have been shown to both retain thermally induced aggregation properties as well as biological activity[6]. However, different variations, such as fusion order or linker length, can often detrimentally affect the activity or binding properties of the fused domain[7,8]. Because of this, extensive characterization is required to confirm retained function of the multi-domain construct.

Once it's been confirmed that each domain retains the expected function when in the fusion construct, the fusion protein as a whole is tested for operational function. In short, the thermo-responsive ligand is tested as a method to purify a glycoprotein. Various factors must be investigated in ensure a straight-forward, feasible isolation method can be utilized. The transition temperature, salt concentration, and stability of the binding region must be within reasonable constraints to maintain the integrity of the target glycoprotein.

In order to do this, a fucosylated model glycoprotein was chosen. Horseradish peroxidase (HRP), is a commonly used peroxidase in biomedical research. It's

performance and structure is well characterized, as well as methods for assaying activity. In its native form, it has 9 potential sites for N-glycosylation, 8 of which are typically (60%) core (α1-3) fucosylated[9,10]. The dominant glycan is shown below in Figure 3.1. However, HRP has notably poor stability in regards to both pH and temperature.
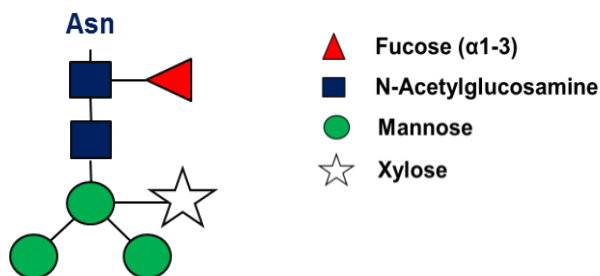
Figure 3.1 Dominant glycan present on HRP.

Because this model glycoprotein has peroxidase activity, it's possible to monitor activity found in each step of the purification process. By doing this, it is ensured that a biologically active product is available at the final step of the purification. Furthermore, it helps localize where the most protein is lost, should it not all be targeted in the process. The short comings of the purification process as a whole can be more directly observed by monitoring activity.

In addition to activity and maintained integrity of the target model protein, it is also important that this fusion protein have high specificity. Particularly, if this construct is going to be applied to industrial or pharmaceutical applications, high purity of the final product is a must. In some cases of novel purification methods, there is an aspect on non-specific binding that occurs. It is desired that this is kept to a minimum to avoid the possible carry-though of endotoxins and other by-products of production and

fermentation. Furthermore, in the case of lectins, there is promiscuous binding observed. In the case of the RSL lectin, while there is a high binding affinity for fucose, there is also a small, but measurable affinity for mannose. Mannose is a very popular monosaccharide: a main building block comprising the main core of N-linked glycans. If enough mannose is present in a sample, it's possible it could out-compete the lectin's preference for fucose. Because of this, the final purity of the desired substance must be monitored, as well as extensive experiments performed pertaining to competitive contaminants in solution.

3.3 Results

3.3.1 ELP Functionality

The transition properties of ELPs have been well characterized[11,12], but the fusion of this peptide to various proteins has had differing effects on usually predictable characteristics. While literature values for the transition temperatures of peptides ELP20, ELP40, and ELP60 were 80, 50, and 36˚C, respectively, inverse temperature cycling was performed at various temperatures to confirm the actual transition temperatures when these peptides were fused to the lectin monomer[13,14]. The theory behind the selection of ELP lengths is based on the achievable transition temperature of the fusion construct. Ideally, a lower transition temperature would be advantageous, in an effort to not denature the target glycoprotein as well as to minimize the energy requirement for the purification process.

Initially, higher than predicted transition temperature were used to ensure the purification of the protein. This was done to ensure initial functionality of the ELP domain. To confirm the isolation of the fusion protein using ITC, the ITC product was visualized by SDS-PAGE analysis. This was done for RSL-ELP20, RSL-ELP40, and RSL-ELP60, as shown respectively in Figure 3.2. The second, higher molecular weight band is addressed in Section 3.3.2.

A)   B)   C)
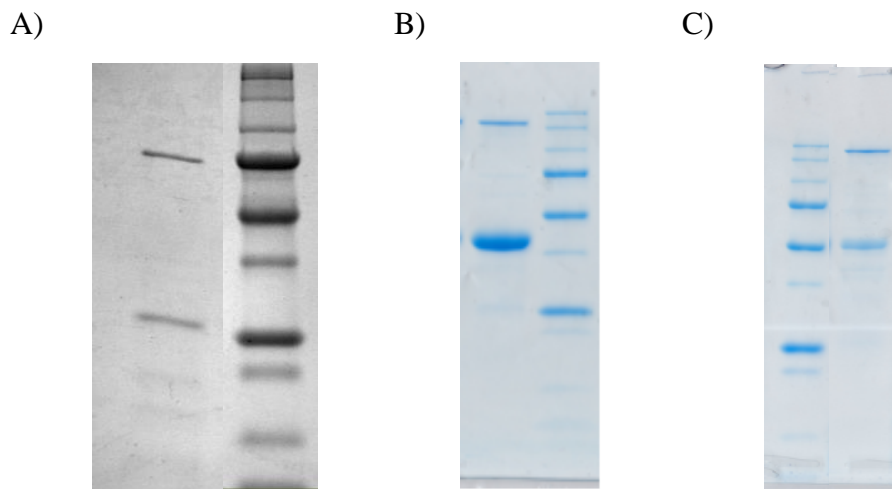
Figure 3.2 SDS-PAGE of isolated A) RSL-ELP20 B) RSL-ELP40 C) RSL-ELP60

The turbidity of the constructs in solution was monitored to determine the exact temperature range where aggregation occurs. The transition temperature is typically defined as the point at which 50% turbidity has occurred[15]. As shown in Figure 3.3, the transition temperature for the RSL-ELP40 construct was determined to be 45˚C, which could be depressed by addition of 0.5M NaCl to a transition temperature of 36˚C. Addition of 1.0M NaCl promoted precipitation of the RSL-ELP40 at room temperature. The RSL-ELP60 in PBS with no additional NaCl displayed aggregation at 38˚C. The

aggregation process was fully reversible and behaved similarly over the course of multiple cycles of temperature ramping and cooling.

Furthermore, it was noted the addition of fucose had no effect on the transition temperature. It was possible that, in the case of the trimerized construct, an additive property of the ELP's could effectively lower the transition temperature. It has been noted that increased concentration of ELP in solution lowers the Tt, so it's possible to conclude that cooperative ELP's in the same complex may reduce the over transition temperature of the trimer construct. However, after complete trimerization of the constructs was confirmed with the addition of 0.1M fucose and SDS-PAGE, the turbidity study was repeated, with no varying results.
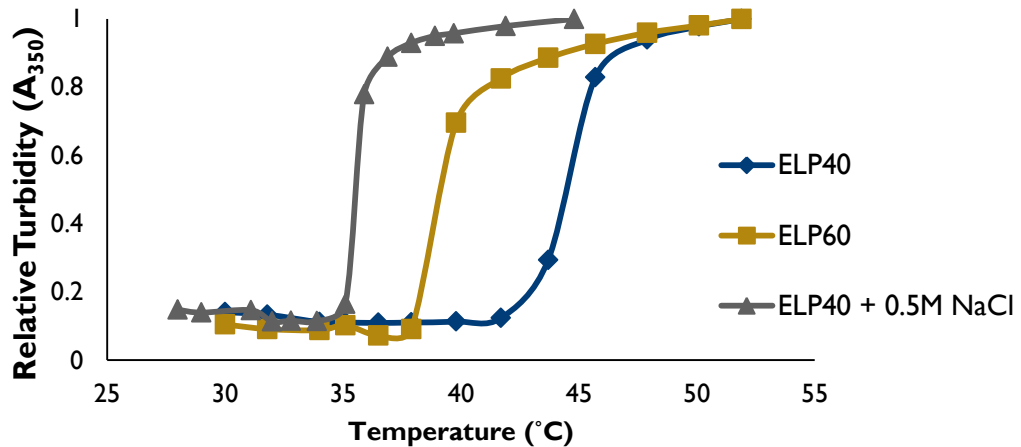
Figure 3.3 Turbidity profile for RSL-ELP40 and RSL-ELP60, as well as influence of salt.

This turbidity study was identically done to ensure the functionality of the ELP domain in the VCNA-ELP40 fusion construct. The turbidity profile is shown in Figure 3.4. In this case, there were no multimers or additional binding issues, so the experiment

85

was carried out simply with PBS and varying salt concentrations. A summary of the two most successful phase transition observations is illustrated below. It was noted that the transition temperature of the fusion construct as a whole was higher than the RSL case. This is likely a result of using a larger lectin. Where the RSL lectin is about 10 kD, the VCNA lectin is about 21kD, comprising of a larger portion of the fusion protein, while the ELP remains the same. However, addition of salt still works as an effective additive, and lowers the transition temperature to about 38°C.
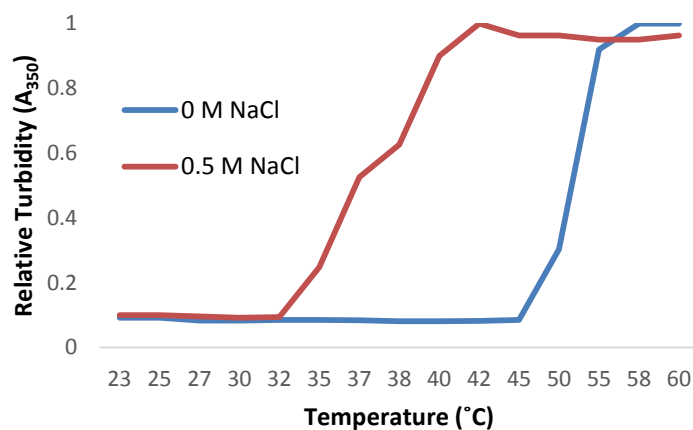
Figure 3.4 Turbidity profile of VCNA-ELP40, with and without addition of salt.

In observing the precipitation characteristics of the ELP, it was also important to monitor the total concentration of ELP in solution. Length of ELP has been linked to concentration of recombinant proteins in plants, where the maximum concentration peaked at 5-30 pentapeptide repeats[16]. Also, higher concentrations of ELP protein lead to more efficient recovery during the ITC purification process. This has been previously noted in fusion protein studies[13,17]. In some cases, addition of free ELP has been added to

the solution as an enrichment strategy. In this study, it was noted the concentration of ELP had to be above 0.5 mg/mL to see a distinct precipitation profile.

3.3.2 RSL Functionality

To ensure functionalization of the RSL in the fusion construct, the construct was purified using affinity chromatography with a fucose bound agarose column. This matrix consists of cross-linked agarose beads with a fucose ligand separated by a five atom spacer. By exploiting the fucose binding characteristics of the RSL, the constructs should be separable by this chromatographic method. By allowing for binding and elution with a fucose buffer, the constructs could be isolated using this method. Indeed, as depicted below by the SDS-PAGE gel in Figure 3.5, an (almost) single band was attainable using this method. There is a small amount of smearing, which is attributed to the trimer nature of the lectin. As the RSL is natively found in trimer form in the presence of fucose, the elution from the fucose affinity chromatography column was washed repeatedly in an attempt to remove fucose and elucidate a single band at the monomer molecular weight. While there is a prominent band seen in Figure 3.5, there is still some smear, suggesting the dialysis did not remove all the fucose.
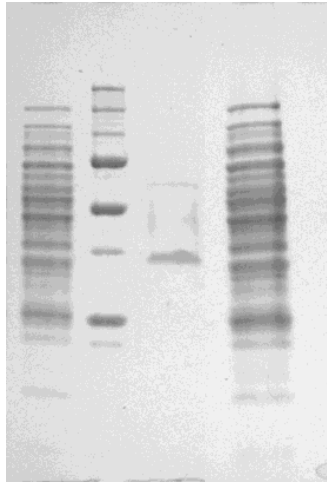
Figure 3.5 SDS-PAGE detailing fucose affinity chromatography results.
Lane 1: Effluent unbound to the column. Lane 2: MW marker. Lane 3: Isolated RSL-ELP40.
Lane 4: Crude protein lysate.

It appears that the RSL still managed to bind, and the smear appears to be a mix of a trimer and a dimer present in solution. In order to investigate the stability and conditions under which this trimer was observable, the lectin was expressed without the ELP fusion.

Once the RSL was expressed in singularity, without fusion, it was purified with the same affinity chromatography method used above. The eluate was washed with DI water 3 times, and the apparent lectin band was shown to drop down to the molecular weight of the monomer, 10 kDa. It was also noted that fusion constructs associated in trimer formations after the addition of fucose. These trimers were observed with such stability that they remained in this configuration even after SDS-PAGE preparation and analysis, enduring addition of reducing agents, boiling, and electrophoresis. This is trimer configuration is required for maximal binding efficiency of the lectin, as 3 binding

sites are intra-monomeric and 3 binding sites are inter-monomeric. To achieve 6 complete sites of fucose association, the trimer configuration must be attained. Although the lectin samples had been washed, they had been in the refrigerator for at least 24 hours, and it appears trimerization occurs without the presence of fucose. This suggests that the natural state and stability of the lectin is optimal in trimer form, and this association will occur in solution with no perturbation.

Experiments were performed to ensure trimer association of the different sized constructs was favorable at a range of conditions. Complete trimerization of all monomers occurred with all constructs in PBS at room temperature with the addition of fucose. It had been hypothesized that incomplete association was likely in the case of the larger ELP constructs due to steric hindrance of the larger ELP fusions. This association is facilitated with the addition of as little as 2 μM fucose, as suggested by the low detection limit in Kostlanova[18]. This is also illustrated by Figure 3.6.
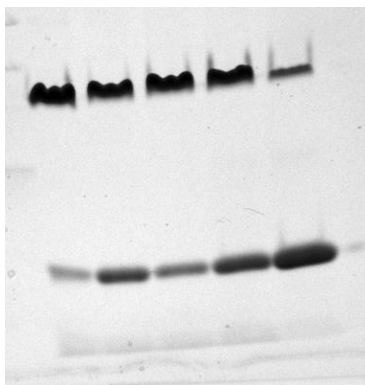
Figure 3.6 Increased trimer presence of RSL with increased fucose concentration. Lane 1: 10 μM fucose. Lane 2: 5 μM fucose. Lane 3: 3 μM fucose. Lane 4: 2 μM fucose. Lane 5: 1 μM fucose.

It's important that the ITC purification does not disrupt the formation of trimers, which are crucial for the targeting of the fucose containing. To confirm this, I observed the trimerization induced by the addition of fucose at varying temperatures. After isolating RSL-ELP20 constructs with ITC, I placed the samples at various temperatures, as shown below, for 15 minutes. After taking a sample from each, I added, 0.5 M fucose to each while keeping each sample at its respective temperature. I allowed 15 minutes for trimerization to occur, then took another sample. It was also hypothesized that a certain destabilizing temperature existed, above which trimers of the RSL would not form and maximum binding efficiency would not be reached. To test this, fucose was added to a solution of fusion constructs over a range of temperatures increasing up to the transition temperature of the ELP. The lectin complex was monitored at each of these temperatures to observe if higher temperatures were detrimental to association. In the most extreme temperature cases for the RSL-ELP20, where the transition temperature is 80˚C, trimerization was observed in all cases save the final stage where precipitation of the construct had already occurred. This is seen in the last lane of Figure 3.7. This suggests that the lectin trimer formation is amenable for all stages of the purification process.

Lanes:
1. 0 C
2. 21 C
3. 37 C
4. 55 C
5. 80 C

6. 0 C with fucose
7. 21 C with fucose
8. 37 C with fucose
9. 55 C with fucose
10. 80 C with fucose
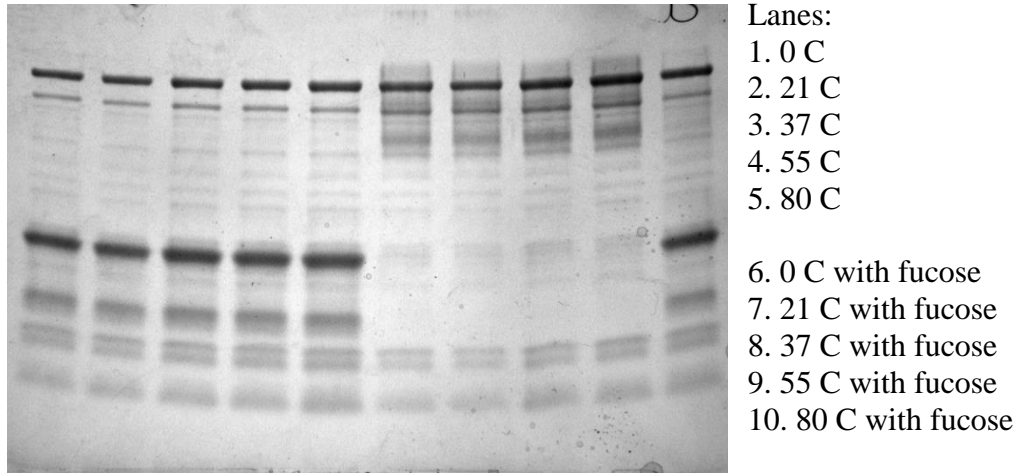
Figure 3.7 SDS-PAGE depicting presence of trimers at different temperatures with and without fucose.

As sodium chloride is often used as an additive to depress the transition temperature of ELPs in the purification process, it is important that increased NaCl not interfere with the association of the monomers of the RSL. As the effect of up to 3.0 M NaCl in solution is investigated in Wood, et al., a similar study was performed. High salt concentrations did not inhibit the trimer formation of the RSL, as pictured in Figure 3.8, so optimal binding sites can still be created.

Molarity of NaCl plus 0.1M fucose					
0.5	1.0	1.5	2.0	2.5	3.0
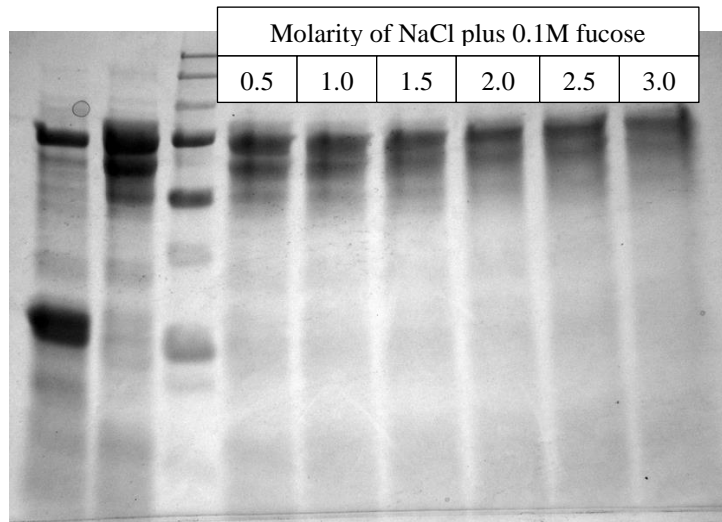
Figure 3.8 Trimerization as a function of salt concentration.

3.3.3 VCNA Functionality

The VCNA-ELP40 fusion construct was tested for retained lectin functionality. After ensuring ELP functionality, covered in section 3.3.1, the lectin portion was tested by targeting a model glycoprotein and a negative control. The model glycoprotein used here was fetuin, a sialic acid containing glycoprotein. VCNA-ELP40 was added to a fetuin solution in a 2:1 molar ratio of lectin to glycoprotein. The predetermined ideal ITC method of 1M salt and transition temperature of 37 was used to isolate the fusion protein and the bound target, and the protein was eluted from the lectin using 2mM sialic acid solution. If the lectin is functional, an isolate of the fetuin should be attainable, while only a single band of VCNA-ELP40 should be seen after attempting to isolate the asialo fetuin. As seen below in the SDS-PAGE analysis of the purification in Figure 3.9, fetuin was isolated. Of course, not all the fetuin was removed from solution; the ratio of 2:1 was not sufficient in targeting all the desired glycoprotein. However, this capture efficiency is further explored in Section 3.3.4.4. This ensures the VCNA lectin domain remains functional after fusion to the ELP domain.
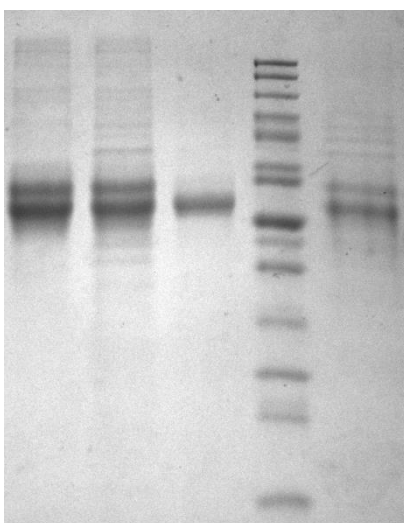
1. Fetuin stock solution
2. Uncaptured fetuin after VCNA and ITC
3. Isolated VCNA after fetuin purification
4. MW ladder
5. Isolated fetuin

Figure 3.9 Purification of fetuin with VCNA.

3.3.4 Glycoprotein Purification with Model Glycoprotein

To illustrate the functionality of this fusion protein, the constructs were used to isolate a fucose containing glycoprotein, horseradish peroxidase. Horseradish peroxidase (HRP) was chosen as model protein, as it is core-fucosylated at eight out of nine glycosylation sites[10]. At 44 kDa, HRP is easily detectable using SDS-PAGE analysis, and confirmation of retained peroxidase activity after purification is easily confirmed with a commercial assay.

3.3.4.1 Purification with ITC

This purification study was performed with constructs RSL-ELP20 and RSL-ELP40. After initial isolation from the cell culture, a purified solution containing the constructs was prepared with ITC. HRP was added in a 1:1 molar ratio to the concentration of fusion protein. This solution was kept on ice overnight to promote complete binding. ITC protocol was used to purify based on the thermo-responsiveness of the ELP, and the recovery of the HRP and construct was confirmed by monitoring protein concentration. This was done using a standard Bradford Assay and SDS-PAGE analysis, as seen in Figure 3.10. Using the RSL-ELP20 construct, a recovery of 56% was achieved of the HRP.
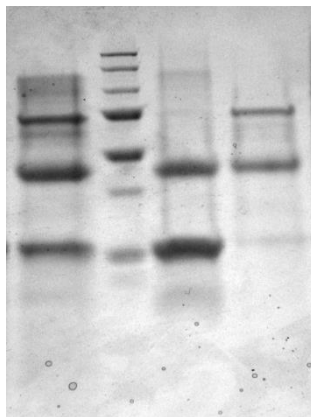
Figure 3.10 SDS-PAGE of HRP isolation using RSL-ELP20.

Retained activity after the purification process implementation was also monitored. By comparing the activity of the HRP before and after the process, as well as HRP undergoing the same conditions, a relative activity was observed. A sample from a mixture of ELP construct and HRP was diluted 100 times and measure for HRP activity of peroxidase over a 5 minutes period. The mixture was submitted to a cycle of ITC or addition of salt to isolate the HRP. After resuspending the pellet containing the ELP construct and bound HRP, a sample was taken and diluted, and the activity of the resulting remaining protein was compared to that of the initial mixture. It was observed that heating of the HRP resulted in a dramatic decrease in activity, as seen in Figure 3.11(a-b). Therefore, the purification that required no heating, just addition of salt, retained the most activity (Figure 3.11c). In fact, with this method, the HRP retained about 98% activity, which suggests that the quantity purified with this method is high as well.
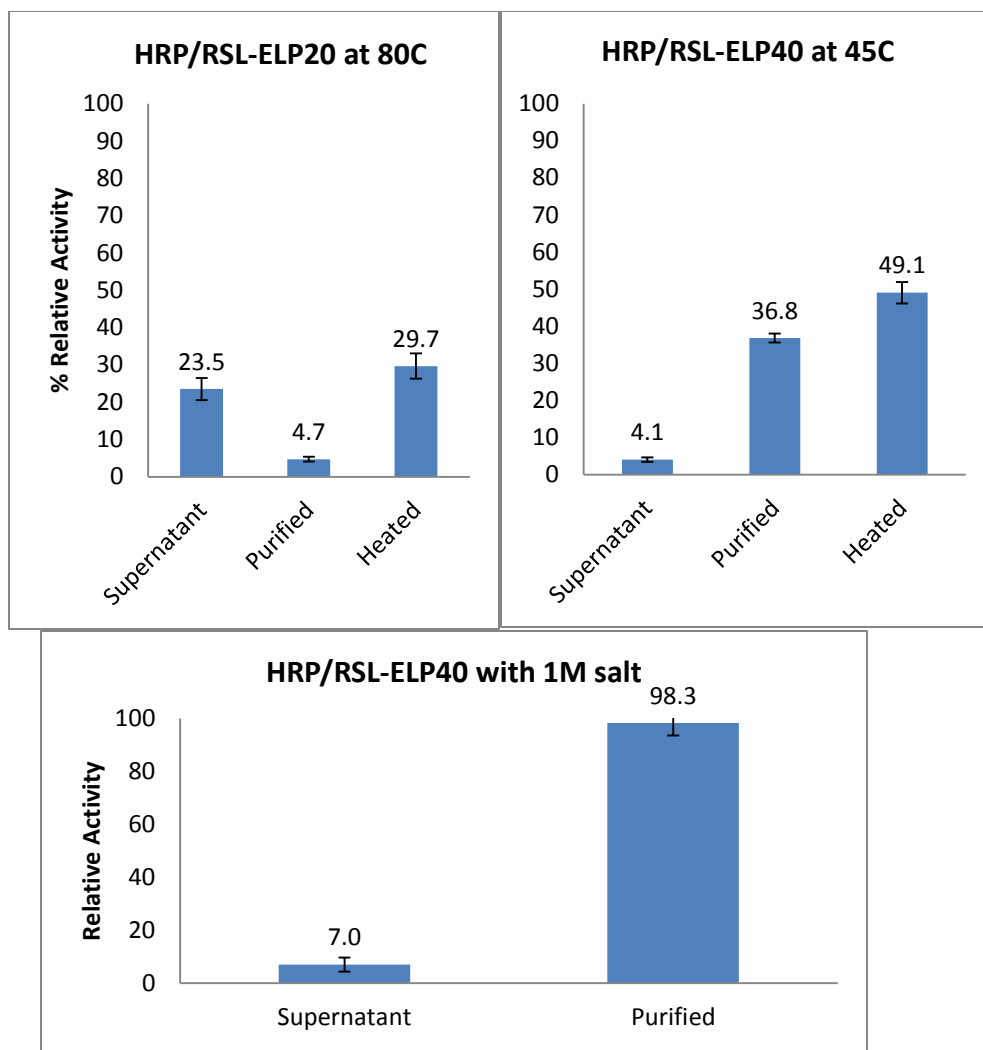
Figure 3.11 Relative activity of HRP at different stages of purification.

3.3.4.2 Purification from High Contaminant Solution

To observe the specificity in a high contaminant solution, HRP was purified from a cell lysate mixture. With no other glyoproteins present in the solution, the RSL-ELP40 construct was tested to ensure non-specific binding was not occurring. *E. coli* whole cell lysate with protein concentration 10 mg/ml was mixed with target protein 0.25 mg/ml (5.7 µM) HRP. RSL-ELP40 fusions were added to a final concentration of 0.5 mg/ml

(18 μM). After one cycle of ITC, followed by addition of fucose to precipitate ligand,

HRP was obtained in highly pure form, shown as a single band (Fig. 3.12), suggesting

that the quality of purified protein maintained even in the presence of large amount of

protein contaminants and other cellular debris.   Additionally, SDS-PAGE analysis was

run on the purified mixtures, and is depicted in Figure 3.12. While the captured activity

may read as high, it's still possible that non-specific binding is occurring, and the final

product is not pure. To check for this, the SDS-PAGE analysis below was observed. A

very clean and specific purification of the HRP was obtained from the lysate mixture,

showing both high yield and purity. Admittedly, some activity retained in the supernatant

after precipitation suggests some HRP is lost in the purification process.   Further

experimentation was required to determine whether it was never initially targeted or

simply lost in the process.

Figure 3.12. HRP isolated from a crude cell lysate solution.

3.3.4.3 Competitive Glycoproteins

        However, the previous section only deals with proteins from a bacterial cell

lysate. As bacteria typically do not perform post translational modifications on their

proteins, it is highly unlikely there were any competing glycoproteins in solution. To further illustrate that the purification is specific for fucose-containing glycoproteins, an experiment was conducted with fetuin used as a glycoprotein target competitor. Fetuin is a non-fucose containing glycoprotein of size 48 kDa, isolated from fetal calf serum[19]. The experiment was carried out as described above except the target protein is fetuin. As shown in Figure 3.13, fetuin appeared only in the soluble fraction after ITC (lane 2), and there was no fetuin present in re-solublized pellet (lane 3), indicating its inability to form complex with RSL-ELP40. Another experiment was conducted with the glycoprotein mixture containing an equimolar ratio (0.1 M) of HRP and fetuin, HRP was recovered with no reduction of yield when compared to the case without fetuin addition (not shown). These results indicate that the presence of a non-fucose-containing glycoprotein did not interfere with the purification and the affinity precipitation is sufficiently specific to fucose containing glycoprotein.
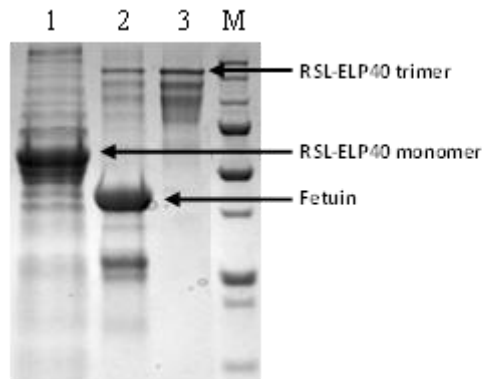
Figure 3.13 SDS-PAGE of fetuin purification, negative control.

In a similar fashion, the VCNA-ELP40 was tested. However, the VCNA lectin is characterized to bind to the sialic acid found on fetuin; instead, asialo fetuin was used as

a negative control. VCNA was added to an asialo fetuin solution in a molar ratio of 2:1 lectin to glycoprotein, and ITC was performed by both addition of 1M salt and a temperature increase to 37˚C. A sample of the supernatant and resuspended pellet was analyzed with SDS-PAGE, as shown below in Figure 3.14. The single band of VCNA seen in the third column shows no presence of asialo fetuin, suggesting the lack of sialic acid present on the glycoprotein ceases the binding and purification interactions. Furthermore, there was no decrease of asialofetuin present when comparing the original solution to the unbound, supernatant fraction, further supporting that no competitive glycoproteins are targeted unless they contain the desired monosaccharide.

1. Asialofetuin stock solution

2. Uncaptured asialofetuin after VCNA and ITC

3. Resuspended VCNA pellet (no asialofetuin seen)

4. MW marker
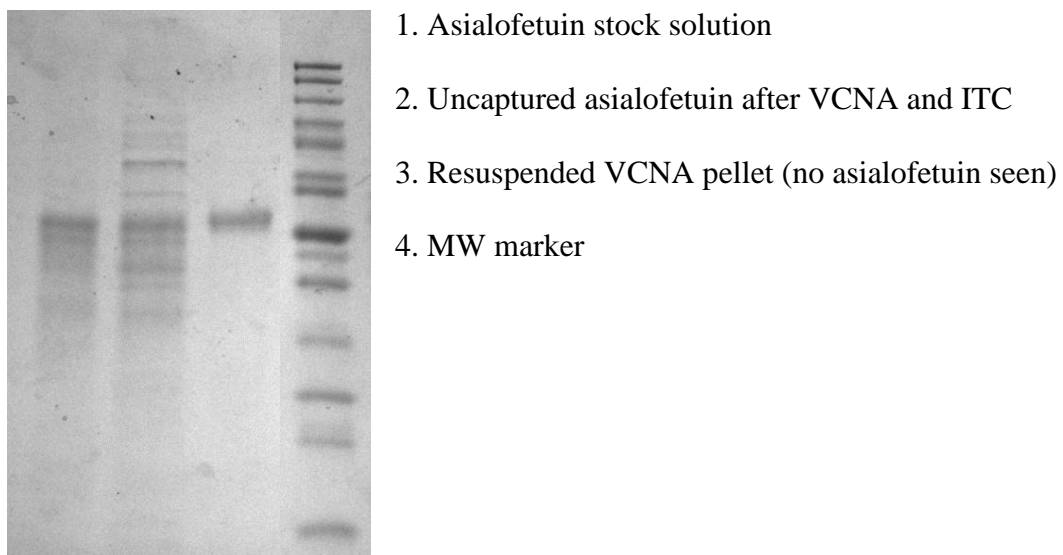
Figure 3.14 SDS-PAGE of asialo-fetuin purification, negative control.

3.3.4.4 Complete Capture

To investigate a relationship between recovery yield and the amount of ligand used in the purification process, the molar ratio of ligand to HRP was varied from 1:2 to

1:20. Results showed that recovery yield increased with the increase of ligand relative to the amount of target protein present in the mixture and full recovery was approached with 15-20 molar excess of ligand. As shown in Figure 3.15, it was observed that the relative activity obtained in a cell lysate mixture was lower than that of a purification done without additional proteins present in the starting mixture. By adjusting the ratio of the ELP trimer complex to the amount of HRP in solution, a higher relative activity was obtained.

The recovery yield in this case was 49±0.9%, slightly lower than the case without contaminating debris, suggesting no major impact on recovery yield by contaminants, illustrating the advantage of the affinity purification process. Even at very low target protein concentration, the presence of cell lysate reduced the recovery of HRP only slightly, from 98 ± 2.8% to 90± 8.6%.
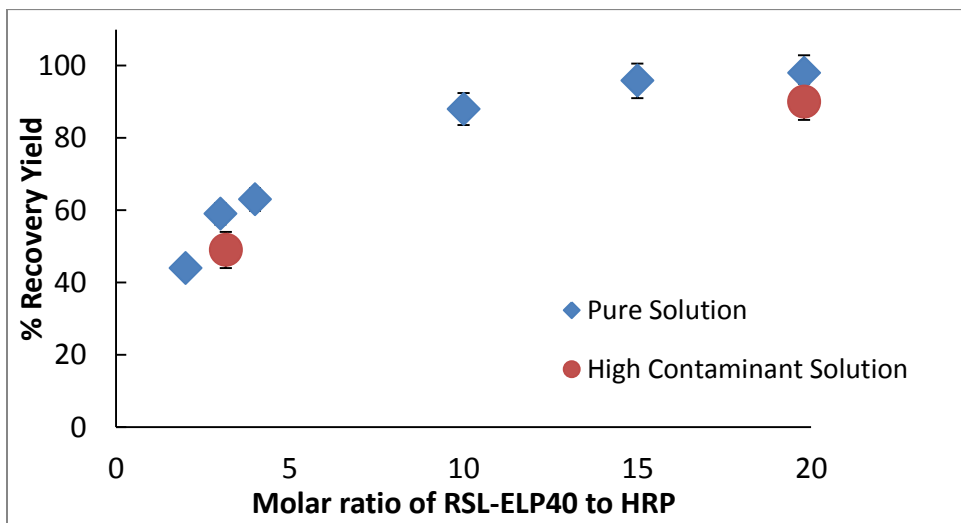
Figure 3.15 Recovery yield as a function of RSL-ELP40 ligand to target ratio.

Additionally, this was done with the VCNA-ELP40 fusion protein in relation to the sialylated model protein, fetuin. The VCNA binding portion, unlike the RSL lectin, does not form multimers, so with only 1 binding site per lectin, the ratio required was expected to be higher than previous RSL experiments. Also, fetuin has, on average, about 13.6 sialic acid residues per molecule, which is also more than the previous fucose content of HRP experiments[20]. Because it's been hypothesized that the ratio is related to complete titration of lectin to target sites, the estimated the molar ratio for complete capture would be around 13:1 VCNA to fetuin. To get a full profile, several ratios leading up to 13 were tested and observed. As seen in the graph below, Figure 3.16, complete capture occurred around 10:1 VCNA to fetuin, which is better than expected, yet still fairly close to the predicted value.
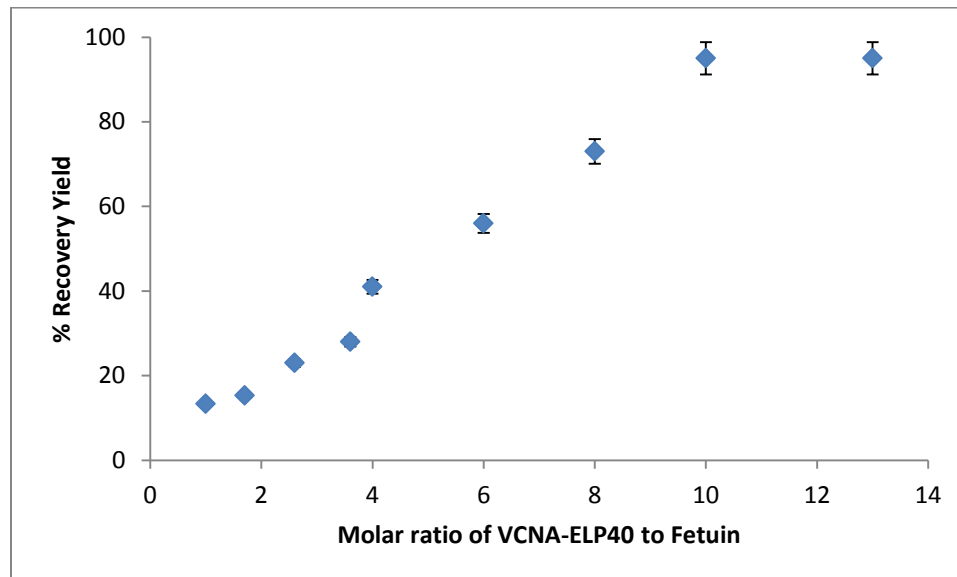
Figure 3.16 Recovery yield as a function of VCNA-ELP40 ligand to target ratio.

3.4 Discussion

After extensive characterization, it appears as if the fusing of the two functional domains did not disrupt the folding of either, as each retained the intended functionality. Furthermore, it was shown that an environmental change manipulating the functionality of one end did not disrupt the functionality of the other. Specifically, the addition of salt or fluctuation of temperature to precipitate the ELP did not disrupt the binding domain of the lectin portion; the addition of fucose for RSL binding did not disrupt the transition characteristics of the ELP. Having confirmed the effectiveness of the construct for its intended use, further experimentation could be followed for purification methodization. These constructs are then applied to a purification scheme to purify a model glycoprotein, horseradish peroxidase. Overall recovery of about 56% was reached in the case of RSL-ELP20. This recovery could be increased using a number of variables. First of all, it is possible that addition of the glycoprotein in an equimolar ratio is over loading the binding capacity of the lectin. The trimerized form of the lectin has 6 binding sites in close proximity, and the HRP has a minimum of 8 fucose saccharides per protein. It is likely that the high avidity of the lectin, due to the polyvalency of the trimer form of the lectin, is binding multiple times to the same molecule of HRP. It is also possible that some form of aggregation is occurring, where one HRP is bound to several lectin, after which the sterics of the ELP prevent further association of these lectin to other HRP proteins. Optimization of this ratio must be done to ensure that 56% is indicative of the actual amount of HRP purified, and not just the amount able to bind to the present RSL-ELP

101

constructs. Once this ratio was investigated, a capture efficiency of 98% was attainable, confirming the ratio hypothesis.

However, these experiments were carried out using a pure solution of isolated HRP in a buffer. As this holds little to no consequence for application, a contaminated solution, comprised of a cell lysate protein solution or competing glycoproteins, was spiked with HRP. This would illustrate a more practical purification scheme. The recovery of the HRP using the lectin-ELP construct was poor at first, which led to further investigation into the stability of HRP. It was found through monitoring the activity of the HRP throughout the process that increasing the temperature was damaging the structural integrity of the enzyme. To remedy this, salt was used to induce ELP aggregation instead of temperature. It is beneficial to have this alternative precipitation method, so other fragile glycoproteins can still be targeted using this purification method. This expands the applicability of this construct to a wider range of less stable proteins.

Once this salt instigation method was ascertained, the experiments pertaining to purification of HRP from a contaminant solution were done. A similar relative recovery of about 90% was seen from the cell lysate solution, displaying minimal disruption of the lysate on the construct and purification method. Similarly, with the competing glycoprotein, fetuin, no fetuin was targeted using this method, portraying a specific affinity to fucosylated glycoproteins. No non-specific binding was seen in either instance.

This thermo-responsible affinity ligand development is a platform for further variation. But capitalizing on the range of lectins available, similar constructs can be created to isolate sample based on the oligosaccharide composition. This is particularly

amenable in the field of glycomics. By utilizing this thermoresponsive ligand, a focused separation will be achieved, avoiding the lengthy and multi-step process of electrophoresis, blotting, extraction, and clean-up. Fractionation of biological samples with respect to a biomarker, such as fucose, will aid in glycomic profiling of diseases and disease progression. This is also a scalable process easily applicable to industrial scale purification. The inexpensive production of the lectin-ELP fusion in a bacterial host avoids the costs of expensive affinity resins. Affinity precipitation is a promising economic alternative to chromatography[21].

3.5 Materials and Methods

Isolation with Affinity Chromatography

Fucose binding functionality within the fusions were tested with a Fucose Separopore® (Agarose) matrix, obtained from bio-WORLD (Dublin, OH, USA). The crude protein solution was added to this matrix and equilibrated with PBS. This mixture was kept on ice on an orbital shaker at 60 RPM overnight to allow for binding. The unbound fraction was washed from the column with PBS, and the bound fraction was eluted from the column with an elution buffer of 0.1M fucose in PBS. Samples from each stage of the process were collected and analyzed with SDS-PAGE.

Inverse Phase Transition

The phase transition was facilitated by one of two methods: 1. The solution was heated to the transition temperature of the ELP in an Eppendorf Thermomixer for 20 minutes at 300 RPM. 2. The appropriate concentration of NaCl was added to solution, either promoting the precipitation in itself, or aiding in precipitation by proceeding the aforementioned heating step. The solution was then centrifuged at 16,000 x g for 1 minute. The supernatant was removed, and the pellet resuspended in PBS or PBS with 0.5M NaCl. The resuspended solution was chilled at 4C overnight to maximize resolubilization of the ELP constructs. This chilled solution was centrifuged again for 5 min at 16,000 x g for 5 minutes to remove unspecific precipitates, and the supernatant was collected as the purified protein fraction.

Turbidity Measurements

The transition temperature of the constructs was determined spectrophotometrically in a 96-well microplate reader (Spectramax M5, Molecular Devices). Each well contained 200 microliters of PBS with 0, 0.5, or 1.0 M NaCl with construct at a concentration of 5 mg/mL. Temperature was increased (in 2°C increments) from 30 to 50°C and the absorbance at 310 nm was measured.

SDS-PAGE Analysis

Each sample was combined in a 1:1 ratio with SDS sample buffer (10% SDS, 10% b-mercaptoethanol, 0.3 M Tris–HCl (pH 6.8), 0.05% bromophenol blue, 50% glycerol), boiled for 5 min, and resolved by 12.5% (w/v) SDS–PAGE. Each gel was detected by Coomassie blue staining (Bio-Rad).

Peroxidase Assay

For peroxidase assay, a solution of 0.0017 M hydrogen peroxide was prepared by diluting 1 mL of 30% hydrogen peroxide to 100 mL with DI water. 1 mL of this solution was further diluted into 50 mL of 0.2 M potassium phosphate buffer pH 7.0. A 0.0025 M 4-aminoantipyrine solution was prepared in 0.17 M phenol solution. The hydrogen peroxide solution and aminoatipyrine solution were mixed 1:1 ratio, and diluted horseradish peroxidase was added. Absorbance of the resulting mixture was measured at 510 nm over a 5 minute time period at 25°C in a microplate reader, Spectramax M5 (Molecular Devices).

Contaminant Solution

The purification of HRP from a "contaminant" solution was done with a cell lysate mixture. A 50 mL culture of *E. coli* was grown to an OD of 1 then centrifuged to collect a cell pellet. This pellet was separated from the supernatant and resuspended in 5 mL of PBS. After the addition of 0.1mM phenylmethanesulfonyl fluoride (PMSF) to inhibit protease activity, the pellet was subject to sonication. The lysed cell debris was collected by centrifugation, and the remaining supernatant was measure for soluble protein concentration. This supernatant was spiked with HRP for the experimental amount and used as the high contaminant solution.

3.6 References

1. Liu, N. et al. Fusion Proteins from Artificial and Natural Structural Modules. 107–121 (2001).

2. Roque, a C. a, Lowe, C. R. & Taipa, M. A. Antibodies and genetically engineered related molecules: production and purification. Biotechnol. Prog. 20, 639–54 (2004).

3. Floss, D. M., Schallau, K., Rose-John, S., Conrad, U. & Scheller, J. Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. Trends Biotechnol. 28, 37–45 (2010).

4.   Koria, P. et al. Self-assembling elastin-like peptides growth factor chimeric nanoparticles for the treatment of chronic wounds. Proc. Natl. Acad. Sci. U. S. A. 108, 1034–9 (2011).

5.   Lin, M., Rose-John, S., Grötzinger, J., Conrad, U. & Scheller, J. Functional expression of a biologically active fragment of soluble gp130 as an ELP-fusion protein in transgenic plants: purification via inverse transition cycling. Biochem. J. 398, 577–83 (2006).

6.   Shamji, M. F. et al. Development and characterization of a fusion protein between thermally responsive elastin-like polypeptide and interleukin-1 receptor antagonist: sustained release of a local antiinflammatory therapeutic. Arthritis Rheum. 56, 3650–61 (2007).

7.   Christensen, T. et al. Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. Protein Sci. 18, 1377–87 (2009).

8.   Shimazu, M., Mulchandani, A. & Chen, W. Thermally triggered purification and immobilization of elastin-OPH fusions. Biotechnol. Bioeng. 81, 74–9 (2003).

9.   Wuhrer, M., Koeleman, C. A. M., Hokke, C. H. & Deelder, M. Protein Glycosylation Analyzed by Normal-Phase of Glycopeptides. 77, 886–894 (2005).

10.   Gray, J. S., Yang, B. Y., Hull, S. R., Venzke, D. P. & Montgomery, R. The glycans of soybean peroxidase. Glycobiology 6, 23–32 (1996).

11. Luan, C. H., Parker, T. M., Gowda, D. C. & Urry, D. W. Hydrophobicity of amino acid residues: differential scanning calorimetry and synthesis of the aromatic analogues of the polypentapeptide of elastin. Biopolymers 32, 1251–61 (1992).

12. Wood, D. W. Non-Chromatographic Recombinant Protein Purification by Self-Cleaving Purification Tags. Sep. Sci. Technol. 45, 2345–2357 (2010).

13. Meyer, D. E. & Chilkoti, A. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. Biomacromolecules 5, 846–851 (2004).

14. Meyer, D. E. & Chilkoti, a. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. Nat. Biotechnol. 17, 1112–5 (1999).

15. Kostal, J., Mulchandani, A. & Chen, W. Tunable Biopolymers for Heavy Metal Removal. Macromolecules 34, 2257–2261 (2001).

16. Conley, A. J., Joensuu, J. J., Jevnikar, A. M., Menassa, R. & Brandle, J. E. Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. Biotechnol. Bioeng. 103, 562–73 (2009).

17. Ge, X. & Filipe, C. D. M. Simultaneous phase transition of ELP tagged molecules and free ELP: An efficient and reversible capture system. Biomacromolecules 7, 2475–2478 (2006).

18.   Kostlánová, N. et al. The fucose-binding lectin from Ralstonia solanacearum. A new type of beta-propeller architecture formed by oligomerization and interacting with fucoside, fucosyllactose, and plant xyloglucan. J. Biol. Chem. 280, 27839–49 (2005).

19.   Spiro, R. G. Studies on fetuin, a glycoprotein of fetal serum. J. Biol. Chem. 235, 2860–2869 (1960).

20.   Spiro, R. G. & Bhoyroo, V. D. Structure of the O -Glycosidically Linked Carbohydrate Units of Fetuin Structure of the O-Glvcosidicallv of Fetuin * Carbohydrate. 5704–5717 (1974).

21.   Arnold, L. & Chen, R. Novel thermo-responsive fucose binding ligands for glycoprotein purification by affinity precipitation. Biotechnol. Bioeng. 111, 413–417 (2014).

CHAPTER IV

METHOD OPTIMIZATION WITH SOYBEAN PEROXIDASE

4.1 Abstract

Effective methods for isolation and purification of glycoproteins and other glycoconjugates are important to biopharmaceutical industry and diagnostic industry. They are also critical to an emerging field of glycoproteomics aimed to understand structure and function relationships of glycans. Previously, we reported development of thermal-responsive affinity ligands by fusing elastic like polymer (ELP) to a bacterial lectin. In this work, we apply the newly developed affinity ligand in an affinity precipitation process to isolate fucosylated soybean peroxidase (SBP) from a dilute crude plant extract. Under optimal conditions, one step binding and precipitation resulted in >95% recovery yield directly from crude extract and a 22.7 fold purification. The SBP isolated using this affinity precipitation process meets or exceeds the quality specifications of comparable reagent grade products by Sigma. We showed that the recovery yield had a strong dependence on the molar ratio of ligand to target fucosylated protein, with a ratio of three being optimal, which could be predicted based on the total fucose content per protein molecule and the number of binding site per ligand molecule. The optimal condition for elution of target protein was determined and a relatively high concentration of cognate sugar for the lectin (1mM fucose) was needed for effective elution (≥80%). To regenerate the ligand, fucose bound at the binding sites was washed off with a pH buffer containing no fucose. This simple wash process allowed the ligand

to regain much of the binding ability. However, a gradual decline of binding ability was observed after each reuse. In spite of the decline, we demonstrated that the ligand could be used for 10 times, giving an averaged 80% isolation yield based on initial input of soybean peroxidase.

4.2 Introduction

4.2.1 Soybean Peroxidase

Plant peroxidases are used widely as research reagents. With their activity easily and inexpensively detected, they are included in numerous commercial diagnostic kits. They are also used widely in food processing and in waste water treatment[1,2]. Horseradish peroxidase (HRP), for example, has been used as a conjugate for CL-ELISA for detection of antibody-mediated binding events, N-de-methylation to produce food flavor, and in removal of dyes from polluted water[2-4]. These many uses of HRP could well be replaced with the functionally equivalent soybean peroxidase (SBP). Moreover, since SBP is more thermal and environmentally stable than HRP, it can be used in applications requiring more harsh conditions than HRP can tolerate[5-7]. SBP could potentially be derived from soybean hull, a byproduct of the soybean industry[8,9]. However, as SBP exists in a complex plant crude extract with a low concentration of 3% of the total hull protein content[9], it requires an efficient isolation and purification process to a suitable purity dictated by the target application. A previous economic analysis concluded that, using liquid enzyme extraction as the method of isolation, the main cost barrier is the final spray drying step. Affinity precipitation, could potentially address the

cost barrier by concentrating the target protein early in the process and reducing processing volume by virtue of specific interaction of ligand and protein.

4.2.2 Affinity Precipitation Process Optimization

Previously, we reported development of novel thermal-responsive sugar-binding ligands by fusing a small bacterial fucose lectin with an ELP[10]. In this work, we apply the fucose-binding ligand to isolate SBP, a fucosylated protein from complex plant crude extract. We demonstrate a one-step affinity precipitation process that achieves >95% recovery and a product quality surpassing commercial reagent grade. In chapter 3, it was briefly addressed that modifying some basic parameters, such as ratio of ligand to target, result in dramatic increases in purification yield. However, this is only one parameter in one step of a much larger process. In order to effectively optimize the novel purification process that works in tandem with the fusion protein, it must be properly constrained by a variety of parameters at all steps. Ratios, salt concentration, pH, elution buffers, and temperatures are often the focus of ELP purification systems[11,12].

An effective method of affinity precipitation was developed, which could be used to enrich a low abundant target glycoprotein from a complex mixture with a high recovery yield. Although the method is only shown here for a fucosylated protein, it is conceivable that this method of purification can be extended to other glycoproteins by developing thermal responsive ligands with lectins specific to sugars other than fucose, such as the VCNA construct cover in other chapters.

4.3 Results

4.3.2 Purification Method

The method used in tandem with the fusion construct is a simple, straightforward method derived by understanding the dual functionality of the fusion construct. By outlining the process, the areas that require optimization can be pinpointed. Furthermore, the specifications of the crude and final product can be investigated to give a clear illustration of the capability of the purification process.

<u>4.3.1.1 Process</u>

A schematic representation of the affinity precipitation process is shown in Figure 4.1. The purification is initiated by addition of the thermal-responsive ligand, RSL-ELP fusion protein, to the crude plant extract. The mixture is incubated at an ambient temperature (~25°C) for 30 minutes to allow the lectin to bind to the fucose on the surface of SBP, resulting in formation of RSL-ELP-SBP complex. This is designated as "binding" in the figure below. Subsequently, the mixture is brought to 50°C, the complex precipitates out of the solution within 3 minutes. The fraction of SBP activity captured by the ligand, calculated by subtraction of residual SBP activity left in supernatant from the total initial SBP activity, is taken as recovery yield. The above process was developed using 50°C precipitation, which takes advantage of SBP's thermal stability. For proteins that are less thermal stable, the transition temperature could be lowered by including salt in the precipitation step. The RSL-ELP construct was determined to have a Tt of 45°C by turbidity studies. Addition of 0.5M NaCl resulted in a reduction of this temperature to about 35°C. Thus the method could be easily adapted to a wide range of proteins with thermal sensitivity.

After the soluble contaminants are centrifuged out, the SBP-ligand bound complex is resolubilized. However, to consolidate steps, this resolubilization is done in an elution buffer. This both allows for the target and lectin-ELP to become soluble, as well as allows for dissociation of the complex. This elution buffer should now contain only the fusion protein and target glycoprotein. At this point, another round of ITC is performed. This precipitates the ELP portion, which is no longer bound to the target. The pellet that forms from the aggregation is the lectin-ELP only, and the SBP remains in the soluble fraction. The pellet is collected via centrifugation, and either washed or simply added back to a crude solution for recycle usability. The soluble fraction is measured for SBP activity, and this is denoted as "isolation yield."
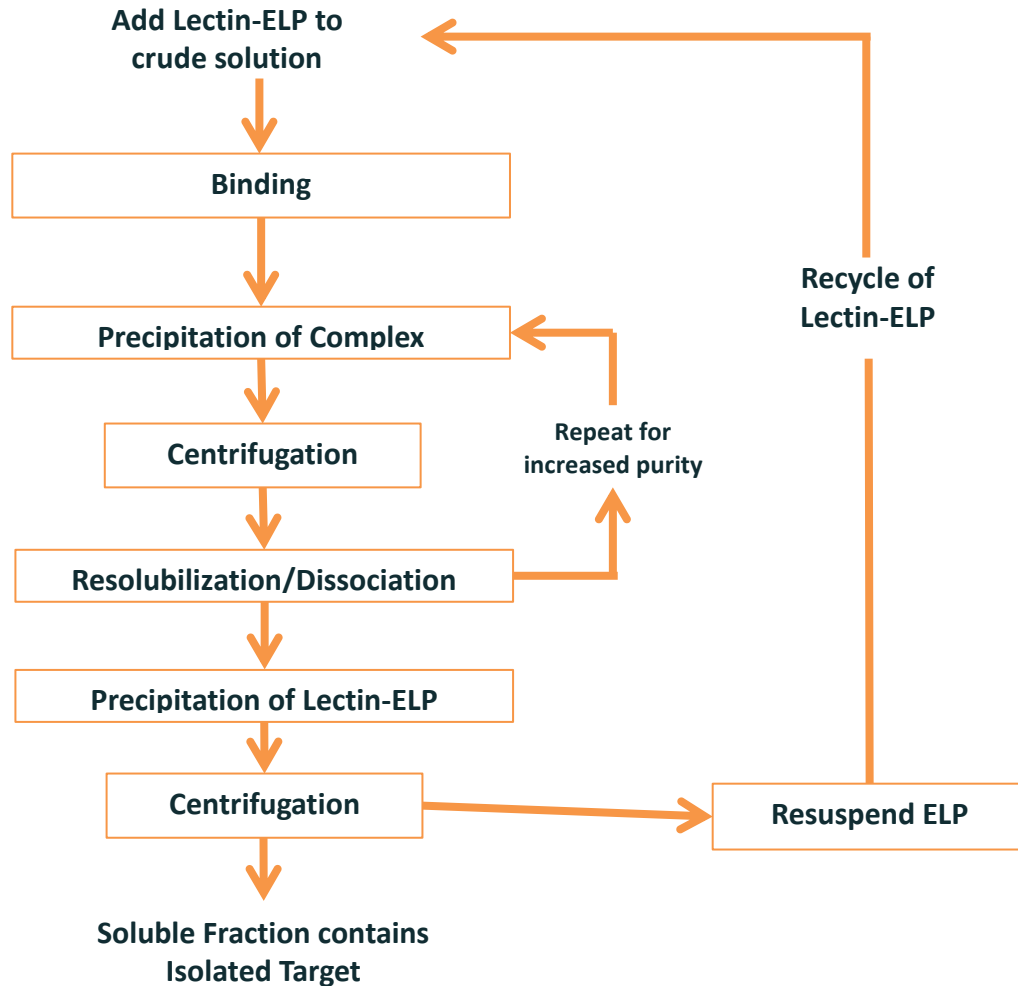
Figure 4.1 Purification method schematic.

4.3.1.2 Natural Extract Characterization

The initial crude soybean extract had a protein concentration of 3.2 mg/mL, and the initial peroxidase activity was 18.5 U/mg. Ligand was added to the protein solution. Using molar ratio of ligand and target SBP of 3, precipitation at 50˚C the recovery yield of SBP was 95%. Using 1mM fucose in elution buffer resulted in 80% release of SBP, or isolation yield. In a typical experiment, after the elution step, 420 U/mg soybean

peroxidase activity in the supernatant was measured, representing a 8.4 fold of purification. A summary of purification was shown in Table 4.1. The quality of product was further indicated by the RZ (Reinheiszahl) value, which is a ratio absorbance of A403/A275, a measure of hemin content. The RZ value was increased from 0.19 for soybean extract to 0.75 after affinity precipitation. Compared to specifications of Sigma product, the SBP from the affinity precipitation has a specific activity up to 7 fold higher and a RZ value higher than the specified minimum (Table 4.1).

Table 4.1 Summary of average initial and purified quality parameters alongside the Sigma-Aldrich equivalent

	Crude	Purified	Sigma
Total protein (mg)	3.8	0.16	--
RZ value	0.19	0.75	≥ 0.5
U/mg protein	50	420	50-150

The product stream was further analyzed by SDS-Page, which showed significantly reduced number of bands present. While a single homogenous solution was not obtained, the target protein was significantly enriched in the product stream (Fig. 4.2), corroborating with the 8.4 fold increase in specific activity. It is likely the widened band representing the SBP is a result of the different glycoforms found in a natural solution.
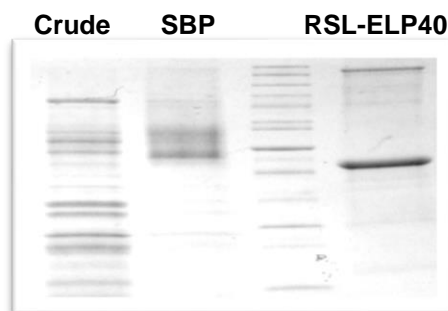
Figure 4.2 SDS-PAGE of crude soybean extract, purified SBP, and RSL-ELP40.

4.3.1.3 Glycan Profile

Additionally, total carbohydrate and fucose content of each purification step were analyzed. This was done to show fractionation of the purified sample in preference of fucose. Each sample was completely hydrolyzed, so glycans were broken down into monosaccharides. Using HPAEC, the monosaccharide profile and compositions were measured.

After purification, there was a significant reduction of both mannose and xylose content in the product stream which accompanied an increase in fucose content, as depicted in Figure 4.3. The concentration of the fucose increased from 4% of the total sugar content to 12%, indicating fractionation in favor of fucose. The other sugars were dramatically eliminated, dropping from 56% of the content to a little less than 5%. This clearance of non-identified sugars was very promising, as it indicates clearance of high contaminants from the natural extract.
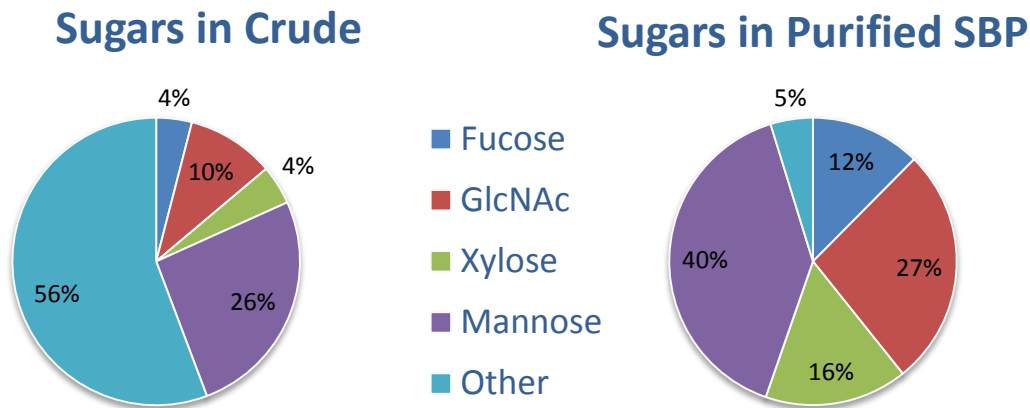
**Figure 4.3 Fractional compositions of the monosaccharides
in the crude and purified samples.**

Notably the fucose content in purified sample was close to the level reported for purified SBP[13], suggesting that a majority of fucose present in the sample are associated with SBP (Table 4.2). Furthermore, the ratio of remaining sugars in the purified fraction was consistent with the determined ratio of present monosaccharides in the native SBP glycans. These ratios are calculated by normalization to the 2 N-acetylglucosamine residues present on the core of every glycan of SBP. From this, the respective ratios of the other monosaccharides are compared, as labeled in Table 4.2. This strongly suggests no prominent other glycoproteins were present in the isolate solution. Taken together, while a homogenous SBP was not evident, most impurity present in plant crude extract was removed and the product stream is significantly enriched with SBP, which should meet most common applications of SBP such as research reagent.

Table 4.2 Molar ratios of detected monosaccharides compared to literature value of SBP glycans

	Molar Ratios			
	Fucose	GlcNAc	Xylose	Mannose
Literature SBP[13]	0.9	2	0.7	3.3
SBP pure	1	2	0.9	3.4

4.3.2 Ratio Optimization

In Chapter 2, a ratio study was done with RSL-ELP40 and HRP. Clear improvement of capture was demonstrated by monitoring the activity of the HRP in different fractions of the purification process. In order to optimize the SBP purification process, this experiment was repeated with the new target glycoprotein. As shown in Figure 4.4, the recovery yield is dependent on the molar ratio of ligand to SBP. As there are two binding sites per lectin[14], a theoretic low limit of ratio is 0.5:1, corresponding to a scenario that a single lectin molecule binds to two separate SBP protein molecules. As depicted in Figure 4.4, this ratio only resulted in a 20% recovery. This could be explained by steric hindrance that prevents a single lectin molecule from binding two fucose sugars on two separate SBP molecules. Increasing the ratio to 1:1, however, improved the recovery only to 50%, indicating the binding of lectin to target protein was not as simple as one ligand: one protein as the ratio would suggest. About 95% recovery yield could be achieved using a ratio of three, representing two molar excess for the ligand. Since on average there are 5.6 fucose moieties per SBP[13], this ratio corresponds to the case that all fucose present on the SBP are bound to lectin. Further increase in ratio did not increase the recovery, as at this point all fucose on SBP were bound to lectin and there was no

more free fucose to bind to excess lectin. This study suggests that fucose on the same protein molecule binds preferentially than to fucose from another protein molecule. This is a likely scenario, as often lectin exhibit avidity, or multivalency that preferentially targets nearby monosaccharides.[15,16] The highest yield is achieved when all fucose are bound to lectins. Thus the ratio that gives a near complete recovery could be predicted by the number of cognate sugar on a protein. The highest recovery of 95% with molar ratio of three is a significant improvement over our last report which requires 20 molar excess for full recovery of a model protein, HRP[10].
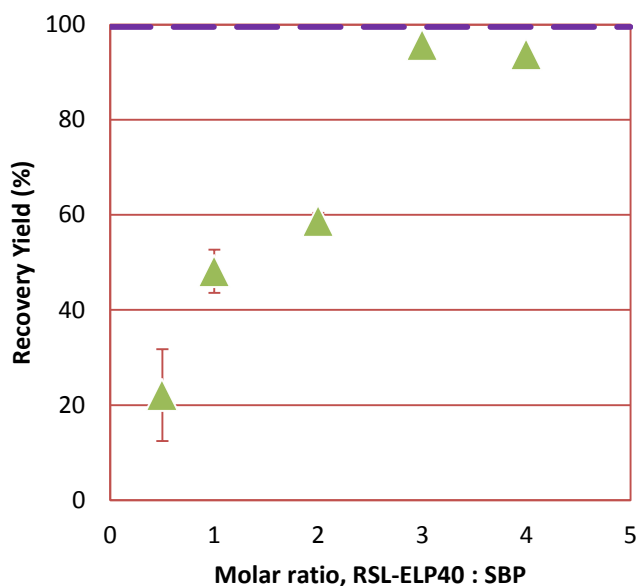
Figure 4.4 Recovery yield as a function of ligand to target ratio.

To ensure that no loss of activity was occurring due to denaturation or inactivity, the SBP activity was also measured in the remaining crude solution after a round of purification. The residual SBP activity left in the crude is labeled as "unbound," and is

120

observed in Figure 4.5 below. The absolute activity, in U/mg, is reported for both the purified and unbound fractions.
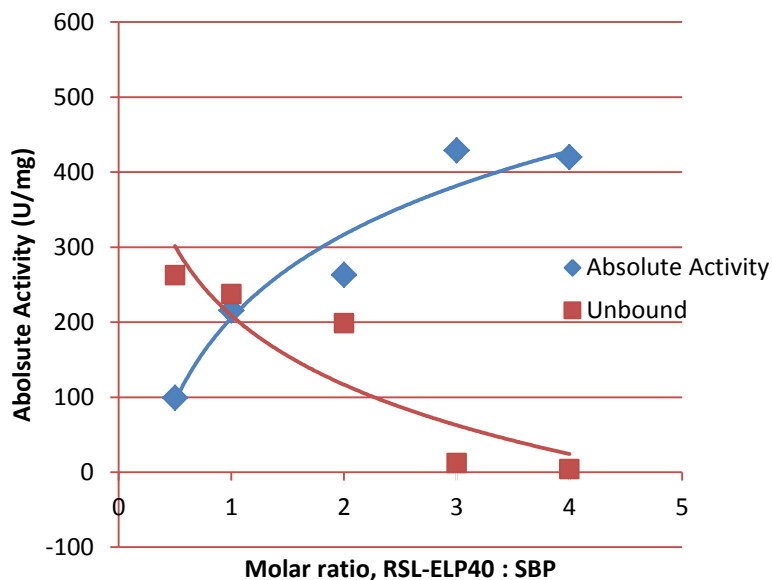
Figure 4.5 Purified and unbound fraction: reported as absolute activity measurements.

4.3.3 Elution Step

As shown previously in Figure 4.1, dissociation of protein complex to elute target protein and to release ligand is an essential step in affinity precipitation. Various elution methods were tested for optimal release of ligand. The elution step must effectively release the target from the ligand, or else there's no way to isolate the glycoprotein. Furthermore, optimization is preferred so as to minimize the use of hazardous or expensive chemicals, while allowing for near complete release of the target glycoprotein. Different methods such as pH, solvents, and competitive binding buffers are investigated.

4.3.3.1 Fucose Concentration Optimization

The most effective method for ligand release for lectins is a buffer solution with cognate sugar. As fucose is relatively expensive compared to other monosaccharides, minimizing its use is important for the SBP purification. We investigated the effect of fucose concentration on the ligand release. Shown in Figure 4.6 is SBP elution efficiency as a function of fucose concentration. Elution efficiency is calculated based on the initial input of SBP at the beginning of an affinity precipitation experiment. The resulting activity is compared to the initial peroxidase activity of the crude sample, thus the relative activity is reported as a fraction of the possible activity. The remaining SBP still bound to the ligand is measured as well, and is denoted as "activity bound." This is activity found in the resuspended solution after a second round of ITC meant to isolate the RSL-ELP from the target. Measuring this activity also ensures that no loss of activity occurs due to inactivation or inhibition from the ITC or elution process. Of the range of fucose concentrations investigated, 1 mM of fucose released about 80% SBP from protein complex, based on activity measured in the supernatant.
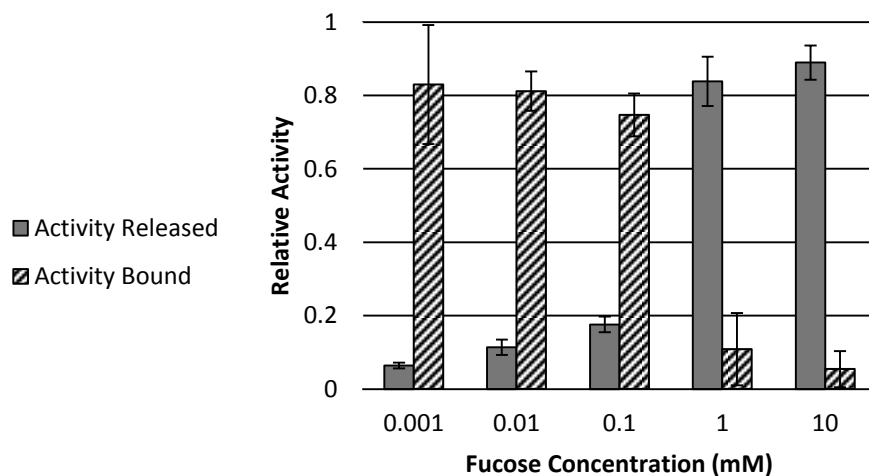
Figure 4.6 Activity in each fraction using a fucose containing elution buffer.

122

4.3.3.2 Avidity of Partially Hydrolyzed Mannose

In many lectin chromatography cases, often a cognate monosaccharide it used to elute the bound fraction. In the case of the RSL lectin, where fucose is the primary target, it would be quite beneficial to substitute a more cost effective monosaccharide in the elution buffer, as fucose is typically about $25/g from common chemical companies. Mannose ($1.10/g from Sigma) is a cheaper and more viable option. However, more of it is required to release the target glycoprotein from the affinity of the lectin, as the binding to mannose is 100 times less than that of fucose. Experimentation and optimization is required to determine if this substitution is a viable economic or process improvement. So, using the procedure similar to that of the pH alteration, the elution step was performed with varying concentrations of mannose in the buffer. Anticipating a high concentration requirement due to the low affinity of the RSL to mannose, the concentrations were 0.1, 0.5, 1, 1.5, and 2M mannose in PBS. Again, the activity of the crude, first supernatant from purification, second supernatant after dissociation, and the remaining pellet were measured. Using the ELP40 construct, it was noted that about 85% of activity was purified from the initial purification. At this point, the pelleted construct and target protein were resuspended in the respective mannose solutions. Following another round of ITC to isolate the ELP from the target protein, the relative activity of the purified SBP is reported.
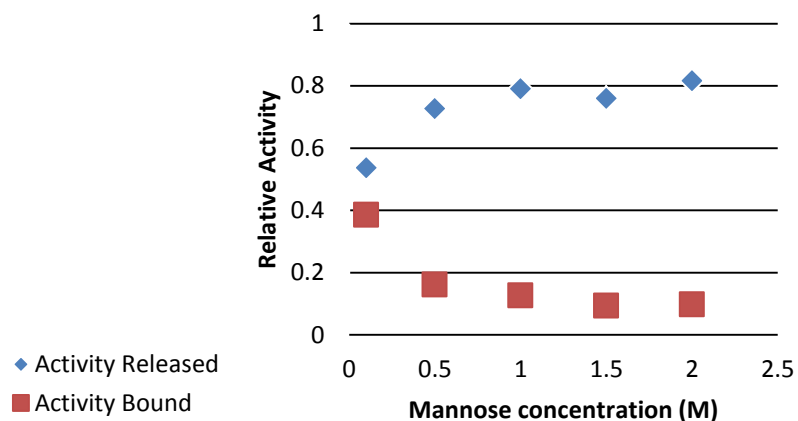
Figure 4.7 Activity in each fraction using a mannose containing elution buffer.

A phenomenon that is often discussed in tandem with lectin affinity is the avidity of lectin containing regions. Avidity refers to the "association constant of a polyvalent interaction"[17]. When there is an increased ligand density, the binding of one lectin to an open site can synergistically aid in the binding of a nearby ligand to another open site. Typically, this is described in terms of surface glycans, where the nearby lectins all associate together to glycans of close proximity. It was hypothesized that the multivalency of the trimer complex would increase the avidity to a surface that provided multiple binding sites.

Table 4.3 Isolation yield using mannose, mannan, or hydrolyzed mannan containing elution buffers at 2 different concentrations.

		Recovered		
% w/v	Molarity	Mannose	Mannan	Hydrolyzed Mannan
0.1	5.6 mM	26.1 ± 2.4%	5.6 ± 0.5%	34.5 ± 1.2%
1	56 mM	30.5 ± 1.5%	6.6 ± 0.7%	**68.6 ± 5.2%**

The partially hydrolyzed mannan averages about 2.5 times better than that of mannose, and at much lower concentrations than required of mannose (56 mM vs. 1.0M) Also, the hydrolyzed mannan performs about 1 order of magnitude better than that of mannan alone, which does not provide many binding sites for the lectin, as the degree of polymerization is high.

4.3.3.3 pH and Solvent Release

As the RSL binds to fucose primarily via hydrogen bonding, it was hypothesized that release of SBP or other fucosylated substrate from the construct could be accomplished by addition of a solvent. DMSO, isopropanol, and ethanol were tested in SBP purification from a crude protein solution using ELP40 and ITC. DMSO and isopropanol did not yield sufficient results, but the ethanol release consistently maintained an activity of about 65% compared to the crude mixture. While this is relatively efficient, it is not performing as well as fucose. A cost analysis will be done comparing this lower yield achieved with a relatively inexpensive solvent to see if there is any merit in using ethanol as a method of target protein release. Low pH buffers (pH 2.2 and 3) were most efficient in SBP release but resulted in the loss of SBP activity. Solvents such as ethanol, isopropanol, and dimethyl sulfoxide, were also tested but resulted in either poor release or enzyme inactivation.
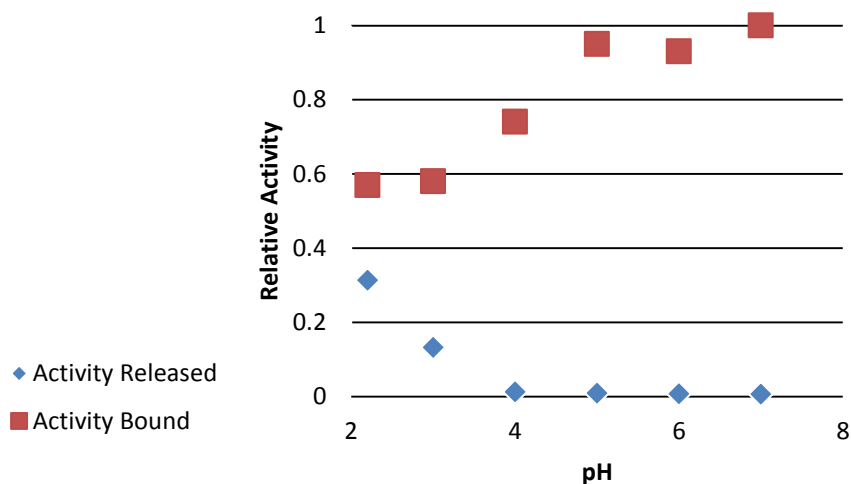
Figure 4.8 Activity of each fraction using elution buffers of varying pH.

4.3.4 Recycling the Ligand

As affinity ligand is possibly a major cost contributor in an affinity precipitation process, we investigated the possibility of its reuse. To this end, after each purification, the ligand was isolated from SBP by temperature-induced precipitation and resuspended in a fresh batch of crude soybean extract to test its reusability. In preliminary experiments, decrease in recovery yield was observed beginning at the third re-use and the recovery yield fell to 25%. The reason for this may be due to the residual fucose that still bound with ligand after each use and ligand release. To regenerate ligand, a wash step was added to the recycle procedure. After the final precipitation step to isolate the target, the pelleted ligand was resuspended in 100 uL PBS solution and resolubilized, rinsing residual fucose from the ligand. The wash solution was subject to ITC again to pellet the ligand for the next addition of crude, and the resulting wash fluid was analyzed for trace activity. This additional step restored the ligand function for additional rounds

of recycle. Figure 4.9 details the purification results of the SBP each time the RSL-ELP was recycled, and Figure 4.10 illustrates localization of all measured activity throughout the purification process. The isolation yield averages at 81.4% over the course of 10 recycles, but a steady decline is visible. It is likely that this drop is due to lost RSL-ELP in the process or possible fouling of the affinity ligand over time. Increasing SBP activity remains in the unbound fraction, suggesting it no longer is binding to the RSL-ELP40. Repetitive temperature fluctuations may aid in the unfolding of the lectin after multiple recycles, so binding is lost. These results showed that the affinity ligand can be reused for 8-10 times without losing binding functionality as long as a wash step to remove residual fucose is added between two purifications. The ability to reuse ligand is a unique advantage in the design of our affinity precipitation.
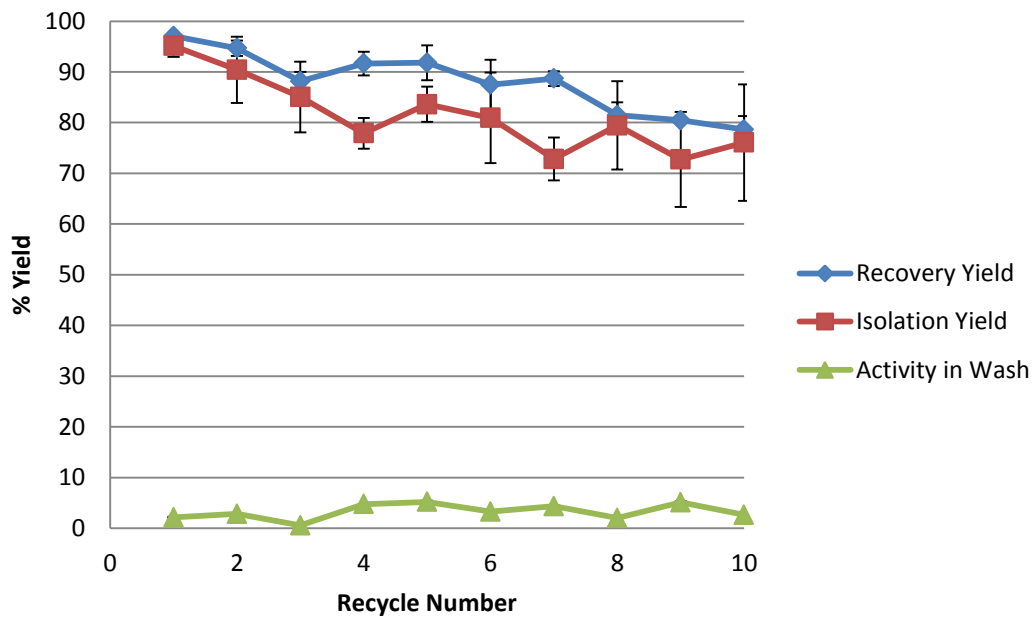
Figure 4.9 Yield activity of each purification step.
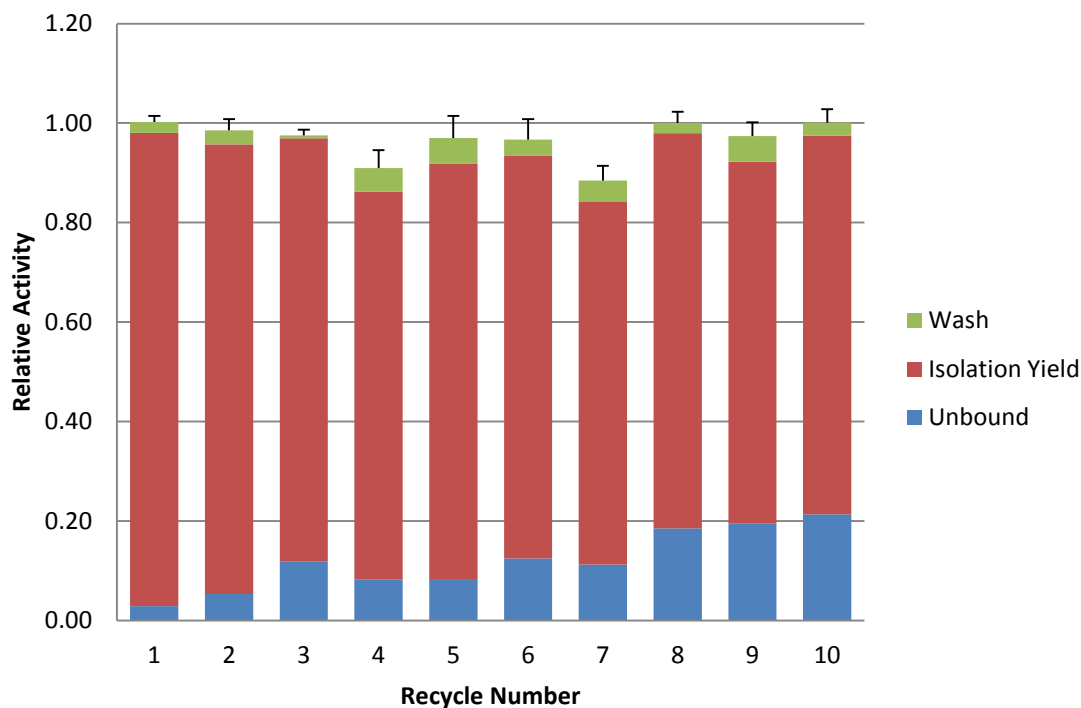
Figure 4.10 Localized activity in each fraction over the course of 10 recycles.

Furthermore, the RSL-ELP can be recycled in previously extracted crude solutions to target any remaining SBP. After the initial precipitation step, the majority of the target glycoprotein is found bound to the RSL-ELP, but often some activity remains in the supernatant. These supernatants, before the recycle ITC, had activities remaining in them from the previous purification ranging from 17-22% of the original activity of the crude. After purification was complete, the isolated fusion protein was recycled into this supernatant to bind to the remaining SBP. After second round of ITC, none of the supernatants had activity above 0.17%, suggesting that the recycle, even with the previously used RSL-ELP construct, was very efficient in capturing the remaining SBP

from the crude. This suggests that SBP makes up 4.2% of the total crude, which is very close to the literature value of 3% SBP content in a crude protein extract.

4.4 Discussion

4.4.1 Purification Capability

We demonstrated in this study a successful use of our newly developed thermal-responsive fucose binding ligand in affinity precipitation for purification of soybean peroxidase from a complex plant crude extract, where the target protein is a minor component (~3%). Under optimal conditions, one step binding and precipitation resulted in >95% recovery yield directly from crude extract and a 22.7 fold purification. The product isolated using this affinity precipitation meets or exceeds the quality specifications of comparable products by Sigma. We showed that the recovery yield had a strong dependence on the molar ratio of ligand to target fucosylated protein, with a ratio of three being optimal, which could be predicted based on the total fucose content per protein molecule and the number of binding site per ligand molecule. The optimal condition for elution of target protein was determined and a relatively high concentration of cognate sugar for the lectin (1mM fucose) was needed for effective elution ($\geq$80%). To regenerate the ligand, fucose bound at the bind sites was washed off with a pH buffer containing no fucose. This simple wash process allowed the ligand to regain much of the binding ability. However, a gradual decline of binding ability was observed after each reuse. In spite of the decline, we demonstrated that the ligand could be used for 10 times, giving an averaged 80% isolation yield based on initial input of soybean peroxidase.

Taken together, an effective method of affinity precipitation was developed, which could be used to enrich a low abundant target glycoprotein from a complex mixture with a high recovery yield.

4.4.2 Economic Analysis

Affinity precipitation is an attractive purification method largely due to low cost of the process. Avoiding the use of expensive resins and chromatographic machines minimizes major expenses. An assessment of the laboratory costs is included in this work to provide a demonstration of the cost effectiveness of this method. All materials are purchased by Sigma. This assessment is outlined in Table 4.4. First, a cultivation of 2 liters of *E. coli* was calculated for material costs. From previous studies, assume a production of 18 mg/L of RSL-ELP40 construct extracellularly after 48 hours of cultivation. Collection of RSL-ELP from the culture is simply done by temperature cycling the extracellular milieu. For the SBP affinity purification process, assumptions are made concerning the crude protein extract. SBP is 3% of total crude hull protein, and total crude hull protein is 0.5% of the total soybean hull. The soybean hull is 5% of the total bean weight. If 2.3 kg of soybeans are processed for the crude hull protein extract, a calculated 17.4 mg of SBP is putatively found in solution. Assume addition of RSL-ELP in a 3:1 molar ratio with SBP.

In calculating the value of this process, the market cost of purified SBP from Sigma is used. Sigma sells SBP for $104.50 for 5 kU. Assuming 75% purification of the SBP in solution, and the average units per milligram achieved, the worth of this process totals more than the expenses at laboratory scale. There is revenue of $33.40. However,

this is calculated after one purification round. If the remaining 15% of SBP were to be extracted from the crude, an additional $38 would be gained, bringing gain to $71.59. Furthermore, this is without capitalizing on the recyclability of the construct. Even assuming 75% purification efficiency with each cycle, and recycling a modest 8 times, the gain comes to $491.93. Even at bench scale, the inexpensive nature of the affinity purification method yields great benefits, and has much potential for scale up and applicability to other target glycoproteins.

Table 4.4 An economic analysis of total gain using this purification process.

A.) Total Production Costs

E. coli cultivation	
LB broth, Ampicillin, IPTG	32.10

Soybean protein extract	
Soybeans, Acetone, Buffer	47.90

Process buffers	
PBS, Elution buffer	1.14

Total Cost **$81.15**

B.) Total Product Value

Value of SBP	
Units SBP purified	5,481
Value per kU	$20.90

Total Value **$114.55**

C.) Total Revenue

With Recycle x 8	
Soybean extract	383.21
Buffers	9.18
Income	916.42

Total Revenue	$491.93

4.5 Materials and Methods

Chemicals

All chemicals were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) unless otherwise specified and were used without further purification. Soybeans were purchased from a local farmers market.

Fusion Protein Expression and Fractionation

The RSL-ELP40 fusion protein was expressed in E. coli and obtained as described in Arnold et al. 2014. A single colony of E. coli UT5600 transformant was selected and cultured in 3mL LB broth and ampicillin overnight at 37˚C and 250 rpm. This seed culture was used to inoculate LB broth culture containing ampicillin to an OD600 of 0.1. This larger culture was grown at 37˚C and 250 rpm to a cell density of OD600 of about 0.4, at which point isopropyl-b-D- thio-galactoside (IPTG) was added to the culture to a final concentration of 1mM to induce the production of the fusion protein. The cultivation temperature was then reduced to 18˚C and the cells were grown for an additional 48 hours.

Periplasmic and cytoplasmic fractions were prepared according to the method described in the pETsystem manual (EMD Chemicals, San Diego, CA). A culture was harvested by centrifugation. The pellet was re- suspended in phosphate buffered saline (PBS) solution, sonicated, and centrifuged at 6,000g at 4˚C for 20 min. The supernatant was saved as a cytoplasmic fraction.

Soybean Protein Extraction

Soybeans were soaked in 10% (w/v) in 0.1 M potassium phosphate buffer (pH 7.0) at room temperature overnight. The hydrated hull was removed from the bean, and the isolated hulls were again soaked in the same buffer for overnight, then homogenized by vortexing for 2 minutes. The mixture was centrifuged for 20 minutes at 4°C at 13,000 x g. The supernatant was collected and concentrated about 20 times by acetone precipitation.

Monosaccharide Characterization

Monosaccharide concentrations were measured by Dionex analysis. To analyze the crude carbohydrate composition, the glycans are first hydrolyzed. The hydrolysis is achieved by addition of 6 N HCl to the protein mixture and heating to 121°C for 45 minutes. After neutralizing with NaOH to a pH of 7, the sample was centrifuged for 30 seconds at 15,000 g to removed precipitates. The sample was diluted for analysis.

The release monosaccharide samples were analyzed by High Performance Anion-Exchange Chromatography (HPAEC) using a DIONEX system equipped with an ED50 electro-chemical detector (Sunnyvale, CA). Monosaccharides were separated on a CarboPac PA-20 column (Dionex). Detection was through pulsed amperometry. (waveform: t = 0.41 s, p = −2.00 V; t = 0.42 s,p = −2.00 V; t = 0.43 s, p = 0.60 V; t = 0.44 s, p = −0.10 V; t = 0.50 s, p = −0.10 V) The mobile phase consisted of degassed solution A containing 100 mM sodium hydroxide and solution B containing deionized water. The mobile phase was pressurized with inert gas (He) to prevent interference of airborne carbon dioxide. An isocratic method containing 2% buffer A and 98% buffer B was pumped at a flow rate of 0.5 mL/min.

DNS Assay

To measure the concentration of sugar in solution, a dinitrosalicylic acid reagent solution (DNS) assay of the solution was used (Rutter 2013)[18]. A 1% reagent solution was used. DNS solution was prepared as follows: 0.75% 3,5-dinitrosalycylic acid, 1.4 % sodium hydroxide, 21.6 % potassium sodium tartrate, 0.55 % phenol, and 0.55 % sodium metabisulfate, dissolved in water. In order to determine soluble reducing sugar concentrations, 100 μL of sample was added to 900μL of DNS solution. These mixtures were then boiled for 5 min, centrifuged at 15,000g for 5 min, and optical density of the supernatant at 550 nm was measured. Reducing sugar concentrations were calculated using glucose as standards.

Partial Hydrolysis of Mannan

Partial mannan hydrolysis was prepared mirroring the method of preparation of agar oligosaccharides (Joo 2003)[19]. 2 w/v% mannan in water was mixed 50/50 with 5% HCl. This mixture was subject to 121 ˚C for 60 minutes. The resulting partially hydrolyzed mixture was neutralized with NaOH.

Acetone Precipitation

Chilled acetone was added in a ratio of 5 to 1 to the crude protein solution. The mixture was vortexed, then refrigerated at -20˚C for an hour. The precipitates were collected by centrifugation at 6,000 x g for 15 minutes. The acetone supernatant was removed, and the pellet was resuspended in PBS. An SBP activity assay as performed to ensure no loss of activity.

Inverse Temperature Cycling

The transition temperature of the constructs was determined spectrophotometrically in a 96-well microplate reader (Spectramax M5, Molecular Devices, Sunnyvale, California). Each well contained 200 mL of PBS with 0, 0.5, or 1.0MNaCl with construct at a concentration of 5 mg/mL. Temperature was increased (in 2˚C increments) from 30 to 50˚C and the absorbance at 310 nm was measured.

SBP Activity Assay

For peroxidase assay, a solution of 0.0017M hydrogen peroxide was prepared by diluting 1mL of 30% hydrogen peroxide to 100mL with DI water. One milliliter of this solution was further diluted into 50mL of 0.2M potassium phosphate buffer pH 7.0. A 0.0025M 4-aminoantipyrine solution was prepared in 0.17M phenol solution. The hydrogen peroxide solution and aminoatipyrine solution were mixed 1:1 ratio, and diluted horseradish peroxidase was added. Absorbance of the resulting mixture was measured at 510nm over a 5 min. time period at 25˚C in a microplate reader, Spectramax M5 (Molecular Devices).

Specific activity is calculated as follows:

$$\text{Units/mg} = \frac{\Delta Abs_{510}/min}{6.58 \times mg\ enzyme/mL\ reaction\ mixture}$$

Samples were run in triplicate to calculate standard deviation of error.

SDS- PAGE Analysis

Each sample was combined in a 1:1 ratio with SDS sample buffer (10% SDS, 10% b-mercaptoethanol, 0.3M Tris–HCl (pH 6.8), 0.05% bromophenol blue, 50%

glycerol), boiled for 5 minutes, and resolved by 12.5% (w/v) SDS–PAGE. Each gel was detected by Coomassie blue staining (Bio-Rad).

4.6 References

1. Chattopadhyay, K. & Mazumdar, S. Structural and Conformational Stability of Horseradish Peroxidase: Effect of Temperature and pH. Biochemistry 39, 263–270 (2000).

2. VanHaandel, M. J. et al. Characterization of different commercial soybean peroxidase preparations and use of the enzyme for N-demethylation of methyl N-methylanthranilate To produce the food flavor methylanthranilate. J. Agric. Food Chem. 48, 1949–54 (2000).

3. Husain, Q. Peroxidase mediated decolorization and remediation of wastewater containing industrial dyes: a review. Rev. Environ. Sci. Bio/Technology 9, 117–140 (2009).

4. Sakharov, I. Y., Berlina, A. N., Zherdev, A. V & Dzantiev, B. B. Advantages of soybean peroxidase over horseradish peroxidase as the enzyme label in chemiluminescent enzyme-linked immunosorbent assay of sulfamethoxypyridazine. J. Agric. Food Chem. 58, 3284–9 (2010).

5. Henriksen, A. et al. Structure of soybean seed coat peroxidase : A plant peroxidase with unusual stability and haem-apoprotein interactions. 108–115 (2001). doi:10.1110/ps.37301.108

6.   Kamal, J. K. A. & Behere, D. V. Thermal and Conformational Stability of Seed Coat Soybean Peroxidase. Biochemistry 41, 9034–9042 (2002).

7.   Mceldoon, J. P. & Dordick, J. S. Unusual Thermal Stability of Soybean Peroxidase. Biotechnol. Prog. 7938, 555–558 (1996).

8.   Gillikin, J. W. & Graham, J. S. Purification and Developmental Analysis of the Major Anionic Peroxidase from the Seed Coat of Glycine max. Plant Physiol. 96, 214–20 (1991).

9.   Hailu, G., Weersink, A. & Cahlík, F. Examining the Prospects for Commercialization of Soybean Peroxidase. AgBioForum 13, 263–273 (2010).

10.   Arnold, L. & Chen, R. Novel thermo-responsive fucose binding ligands for glycoprotein purification by affinity precipitation. Biotechnol. Bioeng. 111, 413–417 (2014).

11.   Sheth, R. D., Madan, B., Chen, W. & Cramer, S. M. High-throughput screening for the development of a monoclonal antibody affinity precipitation step using ELP-z stimuli responsive biopolymers. Biotechnol. Bioeng. 110, 2664–76 (2013).

12.   Madan, B., Chaudhary, G., Cramer, S. M. & Chen, W. ELP-z and ELP-zz capturing scaffolds for the purification of immunoglobulins by affinity precipitation. J. Biotechnol. 163, 10–6 (2013).

13.   Gray, J. S., Yang, B. Y., Hull, S. R., Venzke, D. P. & Montgomery, R. The glycans of soybean peroxidase. Glycobiology 6, 23–32 (1996).

14.   Kostlánová, N. et al. The fucose-binding lectin from Ralstonia solanacearum. A new type of beta-propeller architecture formed by oligomerization and interacting

with fucoside, fucosyllactose, and plant xyloglucan. J. Biol. Chem. 280, 27839–49 (2005).

15.   Imberty, A., Mitchell, E. P. & Wimmerova, M. Structural basis of high-affinity glycan recognition by bacterial and fungal lectins. Curr. Opin. Struct. Biol. 15, 525–534 (2005).

16.   Horan, N. et al. Nonstatistical binding of a Protein to Clustered Carbohydrates. 96, 11782–11786 (2012).

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

18.   Rutter, C., Mao, Z. & Chen, R. Periplasmic expression of a Saccharophagus cellodextrinase enables E. coli to ferment cellodextrin. Appl. Microbiol. Biotechnol. 97, 8129–38 (2013).

19.   Joo, D.-S., Kim, O.-S., Cho, S.-Y. & Lee, C.-H. Preparation Conditions of Agar Oligosaccharides with Organic Acids. J. Korean Soc. Fish. Technol. 36, 6–10 (2003).

CHAPTER V

APPLICATION TO THERAPEUTIC PROTEIN

5.1 Abstract

The goal of this work was to illustrate the efficacy of the expanded platform of the designed purification method using relevant therapeutic glycoproteins of clinical interest. By utilizing two lectin fusions sequentially in a novel multi-lectin affinity precipitation (MLAP) method, a pertinent glycoprotein, erythropoietin, could be targeted. By constraining the process conditions required for the VCNA-ELP40 construct, optimal purification conditions were determined as ITC at 37°C and 0.5 M NaCl with elution conditions of 1 mM sialic acid containing PBS buffer. Replicated culture conditions mimicking industrial production were derived using old CHO cell growth media spiked with target glycoprotein. The isolation ability of the target from this imitation media was promising, with purification yields at $64 \pm 3\%$ using the VCNA construct and $87 \pm 4\%$ using the RSL construct. However, upon obtaining the final formulation of Epogen for experimentation, high presence of human serum albumin (HSA), required further investigation into over abundant glycoprotein depletion. In replicated high albumin solutions, depletion of high abundant proteins consistently ranged from 95-99% while enriching the product concentration in the isolate. EPO in the replicated used CHO medium was targeted at levels of $60 \pm 2.9\%$ and $42 \pm 5\%$ by the RSL and VCNA

constructs, respectively, and resulted in albumin depletion of about 85% and 99%, respectively.

5.2 Introduction

The area of downstream pharmaceutical purification stands to receive the greatest financial benefit with the development of an affinity chromatography alternative. Over 80% of the cost associated with therapeutic protein production is found in the downstream processing. Despite a projected 48% general expansion of mammalian and microbial product production, the improvements in downstream processing, specifically the capacity of chromatography, are incremental.[1] Similarly, for post-translational understanding, pre-analytical purification and concentration of the sample could also benefit from improvement.

In order to apply this affinity precipitation process to an industrially relevant situation, the fucosylated and sialylated glycoprotein erythropoietin was investigated. Erythropoietin (EPO) is a naturally occurring glycoprotein hormone produced by the liver or kidney. EPO indirectly regulates the oxygen in the blood by modulating erythrocytes, and is used in clinical treatments of anemia since the 1980's.[2–4] Typically, it is given to treat people with anemia due to critical illness, such as cancer or kidney failure. Due to its effectiveness, it became the single greatest drug expenditure paid for by Medicare in 2012.[5]

The production and subsequent purification of specific glycoforms of this drug are extremely significant for efficacy. Specifically, the N-linked glycans dictate the activity and half-life of EPO, affecting the biosynthesis, pharmacodynamics, and overall biological activity[2,6,7]. On the three N-linked sites of glycosylation, the chief player in the altering the efficacy is more explicitly attributed to the sialic acid residues present[8,9]. This is an issue in the production phase, as CHO and BHK cells have specific, preferential carbohydrate motifs. However, extensive research has been done on CHO cell production of EPO, as this is the current method of large scale production in use[10,11]. Various perturbations in media affect the number of multi-antennary glycan structures, as well as sialylation. The target glycoprotein is secreted by the CHO cell into the extracellular matrix, from which it then must be purified.

Additionally, in the case of CHO cell cultivation, over-abundant glycoproteins also present with a unique issue in purification. The high concentrations of these contaminants can mask the presence or targeting ability of certain purification methods, and a depletion method must be employed. One such of these over-abundant glycoproteins is albumin, found in high concentrations in blood and serum. Methods that diminish the presence of albumin while enriching in favor of the target glycoprotein are valuable.[12]

The deliverable therapeutic protein must have consistent efficacy. The glycans at all three N-linked sites are typically identical, although one or two can sometimes present with increased branching. The most abundant glycan is shown in Figure 5.1. Additionally, the O-linked glycosylation site may carry additional sialic acid residues, but this is typically on a GlcNAc disaccharide with no other sugar residues present[13,14].

Current methods of isolating this therapeutic are limited to combining multiple types of chromatography.[15]

Figure 5.1 Glycan of erythropoietin, adapted from Wang et. al 2013

To apply the optimized lectin-ELP affinity purification method to purification of the EPO from CHO cell media, a multi-lectin method was employed. Within research utilizing lectin chromatography emerged a cooperative method of combining lectins capable of recognizing different glycan structures, termed multi-lectin affinity chromatography (MLAC). MLAC has been useful in glycoprotein enrichment strategies, and offers a new platform by which to target and analyze a glycoprofile.[16,17] By simulating this method of utilizing multiple lectins, a more targeted process is constructed. By using a novel multi-lectin affinity precipitation (MLAP) approach, a target glycoprotein can be enriched from a crude, high protein content solution, such as the extracellular media from CHO cells. As seen in Figure 5.1, the glycans present on EPO contain both terminal sialic acid residues, as well as a core fucosylation site. By

employing both the RSL and the VCNA lectin fusion constructs, the therapeutic glycoprotein can be targeted and extracted via sequential precipitation steps.

5.3 Results

5.3.1 VCNA-ELP40 Process Optimization

In order to proceed with the most efficient method for EPO isolation, the method using VCNA-ELP40 needed to be optimized as the RSL-ELP40 ligand had previously been (Chapter 4). While optimal ITC conditions had previously been determined for VCNA-ELP40 (Chapter 3), elution conditions had not been investigated.

5.3.1.1 Optimal SA Concentration for Elution

As was performed with the RSL-ELP40 in relation to the SBP purification in Chapter 4, an elution experiment was done with the VCNA with respect to its model target, fetuin. First, a cognate sugar was used at varying concentrations in the elution buffer, and the isolation yield was monitored as a function of sialic acid concentration. At first, an arbitrary 10 mM was used for release in these studies, which was an educated guess based on the Kd of the VCNA lectin. It was likely in excess, which is wasteful of resources and possibly harmful to certain protein due to the positive charge of the sialic acid. In order to reduce the amount of excess SA used in the elution buffer, an optimization of SA concentration in the release buffer was tested. Table 5.1 summarizes the isolation yield as a function of SA concentration in the PBS elution buffer.

Table 5.1 Isolation Yield as a function of SA concentration in Elution Buffer

Sialic Acid Concentration	10 mM	1 mM	0.1 mM	0.01 mM
Isolation Yield	92 ± 4.6%	96 ± 4.8%	61 ± 6.1%	44 ± 4.4%

At 0.01 mM, a fair amount of release was observed: about half. However, it seems sufficient to use 1 mM, as this was generating the highest release, and therefore, highest overall yield. Again, it's possible that the higher concentrations, such as 10 mM, would be detrimental to the target glycoprotein, as SA carries a positive charge.

5.3.1.2 Adjusted pH for Elution

Again, in an effort to create a feasible and cost effective alternative, varying pH was investigated as a release mechanism. Particularly due to the charged nature of this sugar and its relation to the binding properties of the lectin, it was very likely that altered pH would be effective in terminating the affinity between the VCNA and fetuin. The elution was performed by resuspending the VCNA- fetuin pellet in McIlvaine buffers of pH 2.2, 3, 4, 5, and 6. Isolation yield is as tabulated below in Table 5.2, but isolate for pH 3 and 4 were unreadable.

Table 5.2 Isolation Yield as a function of pH of elution buffer

pH	2.2	3	4	5	6
Isolation Yield	30 ± 6.0%	--	--	16 ± 3.2%	10 ± 2.0%

In this case, it appears the low pH is responsible for impairing the fusion protein. At low pH, the VCNA-ELP40 doesn't seem to precipitate correctly. This experiment

was repeated several times with similar results. At higher pH, the change is not enough to facilitate release, so low recovery is seen. Because of the poor results shown in Table 5.2, using a pH change to facilitate elution does not seem to be effective.

5.3.2 Model Glycoprotein Purifications

In order to confirm the presence of competing glycoproteins, the purification technique was first tested on the crude CHO media by spiking with the desire amount of target model glycoprotein. The purpose of this was three-fold. First, a balanced ratio of ligand to target based on previous experiments was used to ensure no loss of capture efficiency in the CHO cell media solution. Secondly, the purified product was monitored to confirm no high abundance of competing glycoprotein contaminants was targeted. Finally, other large bands in the crude were scrutinized to observe depletion of other significant contaminants.

5.3.2.1 CHO media

The used CHO cell media was run on SDS-PAGE, Figure 5.2, to identify the proteomic profile contained in the extracellular matrix. Overall, while there were many visible bands over a broad range of weights and concentrations, the most prominent band was the presence of albumin. The albumin was from the fetal bovine serum added to the media to promote the growth of the CHO cells, not a product from the cells themselves. The CHO cell media is spiked with 10% fetal bovine serum, which, according to the product specifications, contains 17-34 mg/mL albumin.
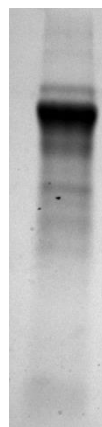
Figure 5.2 Used CHO cell media protein profile.

5.3.2.2 VCNA-ELP40 and fetuin

First, the VCNA-ELP40 construct was used to target a fetuin spiked media sample. 100 µL of extracellular CHO growth media was spiked with 18 µg of fetuin. VCNA-ELP40 was added in a ratio of 10 :1 to the fetuin, and performed ITC using the VCNA-ELP40 favored conditions (37C + 0.5 M NaCl). The elution of the fetuin from the ligand was done with an elution buffer containing 2 mM SA in PBS, and the ligand was isolated with another round of ITC. The SDS-PAGE of the fractions is shown below in Figure 5.3.

As depicted in the SDS-PAGE, the VCNA-ELP40 was successfully isolated using the ITC and elution methods. No remaining fetuin is visible in the ELP ligand lane. The unbound fraction is still predominantly albumin, with some variation of other contaminants seen at other molecular weights as well. The purified fraction, which ideally would contain only fetuin, still shows the presence of some albumin. Based off protein quantification, about $60 \pm 3\%$ of the fetuin was isolated. The estimated purity in the SDS-PAGE is about 70%. The amount of remaining albumin in the product stream

was measured as well. The reduction of this abundant albumin ranges from 94-97%

reduction from the original media. This is a promising way to reduce highly abundant

proteins.
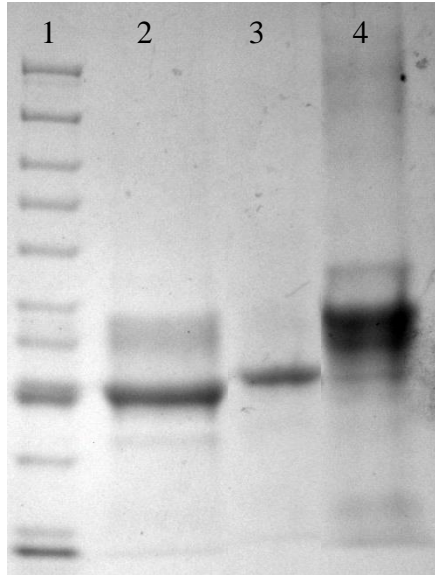
Figure 5.3 Fetuin purified from CHO cell extracellular media by VCNA-ELP40.
Lane 1: Molecular weight marker. Lane 2: Purified fraction. Lane 3: Isolated VCNA-ELP40.
Lane 4: Unbound fraction.

5.3.2.3 RSL-ELP40 and HRP

The ability of RSL-ELP40 to purify the model glycoprotein, HRP, from the used CHO

media was also tested. 20 µg of HRP was added to the media, and RSL-ELP40 was

added in excess. I performed a purification using the favored ITC conditions (50°C) and

used an elution buffer of 1mM fucose. The melting point of EPO is noted as 53°C, so

purification at 50°C should not be detrimental to the structure[18]. The ligand was

separated from the HRP product, and ran this product, the ligand, and the unbound

fraction on an SDS-PAGE for analysis, as shown in Figure 5.4. Similar results were seen

with the VCNA, where the target glycoprotein was successfully targeted, but there is a remaining band that is likely attributed to the large abundance of albumin. It is noticeable that in the purified fraction, there is an RSL-ELP40 trimer band present. The concentration of ELP in solution was presumably too low, and while some ELP was removed during the elution process, the low concentration of ELP resulted in incomplete precipitation, so some ligand is still present in the product stream. The lower band represented in Lane 1 is indeed HRP, based on activity measurements. The purified fraction is still in the elution buffer, so all ligand present will only be in trimer form. This method resulted in an isolation yield of $87 \pm 4\%$, and an albumin depletion of about 92%. Estimated purity in SDS-PAGE of the final product is about 40%, due to the remaining ligand in solution.
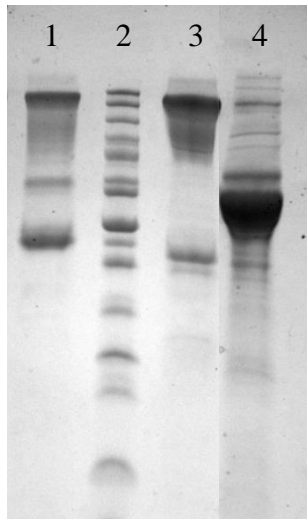

Figure 5.4 HRP purified from CHO cell extracellular media by RSL-ELP40.
Lanes 1: Purified fraction 2: Molecular weight marker 3: Isolated RSL-ELP40 4: Unbound fraction.

5.3.3 Sequential Purification

The multi lectin affinity precipitation process requires sequential lectin addition and ITC. I began to test this method with the model glycoproteins before attempting it with the Epogen acquired from Amgen. I spiked the used CHO media with both fetuin and HRP. I then purified them sequentially, using RSL followed by VCNA.

The sequential purification was done with RSL, followed by VCNA. The first SDS-PAGE panel below shows the ladder, unbound fraction, pelleted ligand, and purified product. Again, there was some residual RSL-ELP40 in the final product, so I performed another ITC round. The isolate is depicted in the 2nd SDS-PAGE panel, which shows just HRP and a band of albumin. Also in the second panel is the purification of the fetuin from the same sample of CHO media using VCNA-ELP40.
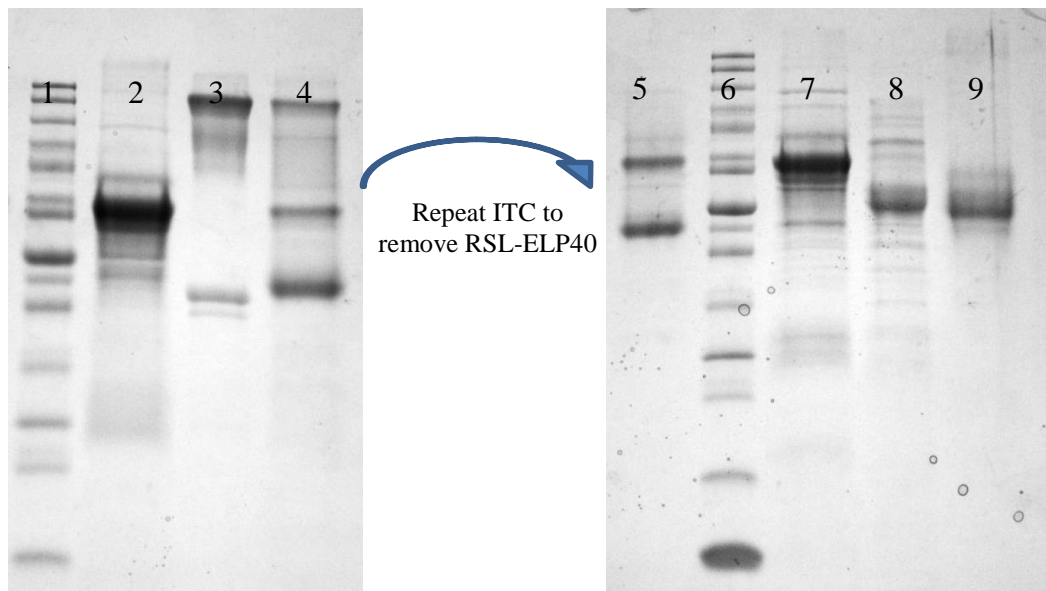
Figure 5.5 Sequential purification of model glycoproteins from CHO media.
Lanes 1 and 6: Molecular weight marker. Lanes 2 and 7: Unbound fraction.
Lane 3: Isolated RSL-ELP40. Lanes 4 and 5: Purified HRP.
Lane 8: Isolated VCNA-ELP40.Lane 9: Purified fetuin.

5.3.4 Over-abundant Contaminant Depletion

As visualized in the sections above, an unanticipated difficulty in isolating EPO from a CHO extracellular media is the large amount of albumin, both BSA and HSA, present in solution. Over abundant glycoprotein depletion is a crucial step in many purification stages, specifically in the area of diagnostics. Albumin in blood samples range from 34-54 g/L, and is the dominant protein, often obscuring the investigation into other lower abundant proteins.

In regards to the current experiment, albumin is present in both the CHO extracellular media as well as the final EPO formulation. The CHO media is supplemented with 10% fetal bovine serum (FBS). Of the total protein content of FBS (average 38 mg/mL), albumin is again the dominant protein (average 23 mg/mL, range of 20-36 mg/mL). As this is added to a final concentration of 10% FBS in media, the final media likely contains about 2.3 mg/mL of albumin. The final formulation of EPO, according to the MSDS, contains 2.5 mg/mL of albumin, but only contains 0.084 mg/mL of EPO. This is about 30 times more albumin than EPO present in the sample, and must be addressed before continuing on with purification. An SDS-PAGE of the received EPO product is shown below in Figure 5.6.
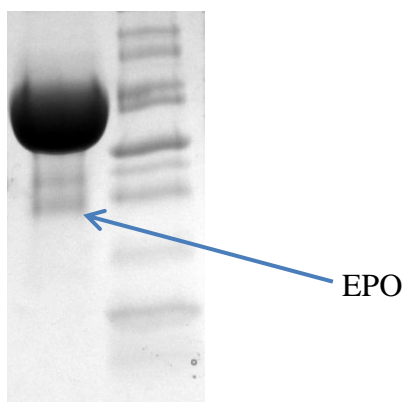
**Figure 5.6 Amgen Epogen visualized on SDS-PAGE.**

According to the MSDS for the product, a 1 mL vial contains 10,000 units, or 0.084 mg of total EPO. The MSDS further outlines the excipient list, which is as follows: 10,000 Units epoetin alfa, albumin (human) (2.5 mg), benzyl alcohol (1%), sodium chloride (8.2 mg), and sodium citrate (1.3 mg) per 1 mL water for injection.[19] So, again, an over-abundant amount of albumin was visualized and perceived as a particular obstacle.

In order to simulate this complication without wasting valuable product, the over abundant protein depletion studies were performed with 40 ug of target protein, either HRP or fetuin, which in 400 µL samples gives a final concentration of about 0.1 mg/mL, similar to that of the EPO presence. The 400 µL sample contained increasing amounts of albumin, and the total recovery as well as albumin depletion was monitored with increasing albumin in both the RSL and VCNA fusion cases.

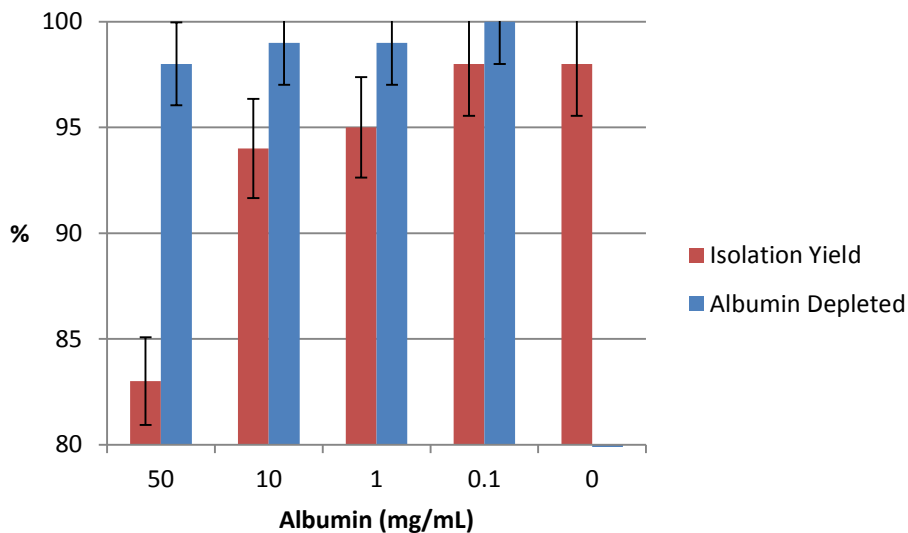5.3.4.1 RSL performance in high albumin

Using RSL-ELP40, the HRP activity could be monitored to determine a more quantitative isolation yield in the presence of excess albumin. A ratio of 12:1 of

RSL:HRP was added to the high albumin solution, as this is a sufficient amount to target the majority of the HRP. This purification was done in 5 solutions of varying albumin content containing 0, 0.1, 1, 10, and 50 mg/mL albumin respectively. This range spans from typical culture concentrations up to the excipient concentrations to seek a point of optimal or failure of purification performance. ITC was performed at 45˚C, and eluted with 1 mM fucose, per the originally defined optimal conditions. The final purified forms from 1 round of ITC are depicted via SDS-PAGE analysis below in Figure 5.7.
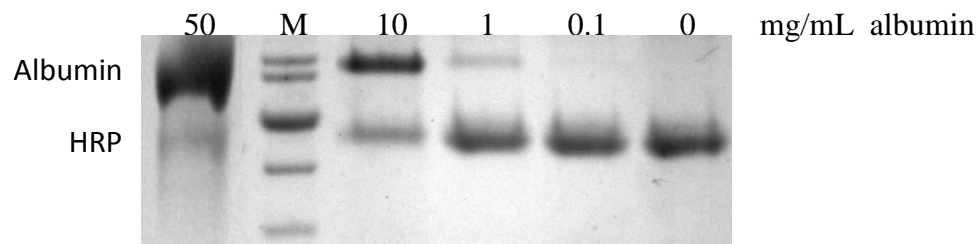
Figure 5.7 Purified HRP from high albumin solutions.

It's apparent that a visible amount of residual albumin is still present as low as with the 1 mg/mL albumin solution, however it is greatly depleted. To further inspect the purification of the HRP, the activity of this solution was measured, and is reported below. An estimate of reduced albumin is reported concurrently in Figure 5.8 to illustrate depletion and enrichment performance.

Figure 5.8 Isolation yield and % albumin depleted as a function of initial albumin concentration (HRP)

5.3.4.2 VCNA performance in high albumin

The same experiment was repeated with the VCNA-ELP40 and accompanying model protein, fetuin. A ratio of 10:1 VCNA ligand to fetuin was used. ITC was performed by the addition of 0.5 M NaCl and heating to 37°C, and elution was done with 2 mM sialic acid in PBS. The final purified forms are depicted via SDS-PAGE analysis below in Figure 5.9.

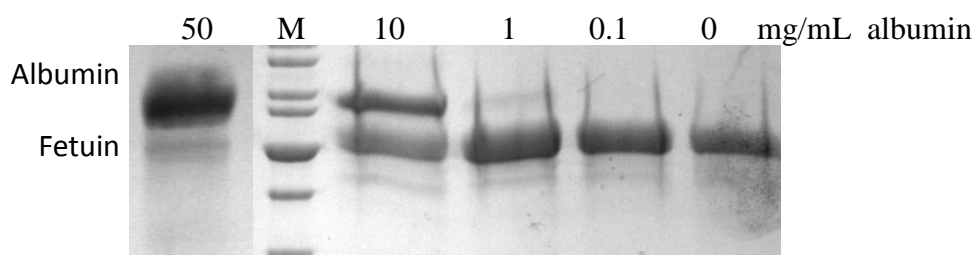**Figure 5.9 Purified fetuin from high albumin solutions.**

Aside from a Bradford protein quantification assay, there was not quantitative method with which to measure the fetuin present in the final isolate. To corroborate estimates, the band intensity and Bradford assay results were compared to estimate the amount of fetuin isolated as well as residual albumin. Results are summarized in Figure 5.10, illustrating some loss of isolation yield with increasing albumin content. However, the albumin depletion remained fairly high at about 95% depletion even in the highest case.
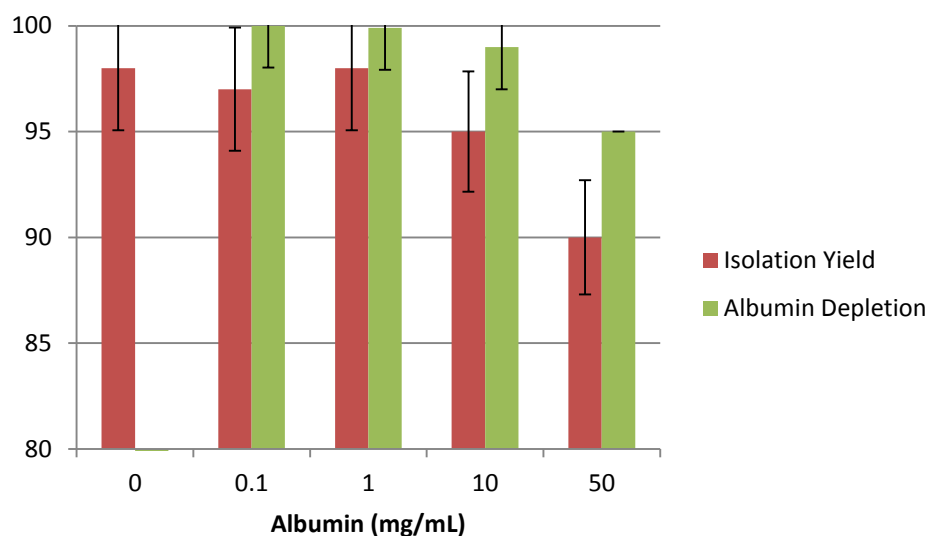
Figure 5.10 Isolation yield and % albumin depleted as a function of initial albumin concentration (Fetuin)

5.3.5 Purification of EPO

After showing a combination of fairly successful purification yields (>80%) and high albumin depletion (>95%) using the lectin fusion proteins in the previous section, an attempt to directly purify EPO from a CHO cell media mix was performed with each lectin. Unfortunately, with such limited sample, various constraints remain to be tested concerning this method. But as a proof of concept, the experimentation was performed.

154

A 50/50 ratio of the EPO with the CHO cell used media was mixed to a final solution with volume 600 μL.

RSL-ELP40 was added in excess, and standard purification conditions were used for ITC and elution. The gel below in Figure 5.11 depicts two dilutions of the final purified fraction, after elution and isolation from the ELP construct. While the albumin is still a dominant band, there is a visible EPO band below the 37 kD marker. Based on band intensity and corroborated with Bradford Assay measurements, 60 ± 3% purification yields were estimated. Furthermore, monitoring the albumin intensity, the albumin depletion reached about 85%, similar to previous experiments using HRP as a target in high albumin solution.
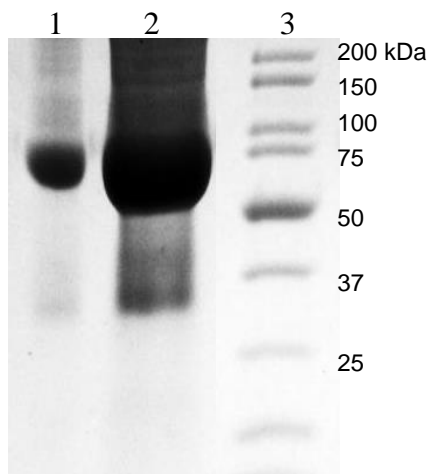
**Figure 5.11 Purified fraction of EPO from CHO cell media**
Lane 1: EPO isolated fraction low sample loading. Lane 2: EPO isolated fraction high sample loading. Lane 3: Molecular weight marker.

Similarly, the same CHO cell media solution combined with EPO was prepared and purified using the VCNA construct. The VCNA-ELP40 construct was added in

excess, and the standard purification conditions were used for ITC and elution. The SDS-

PAGE analysis in Figure 5.12 does not show a purified fraction, as isolation of the EPO

was not successful. Instead, the EPO band seems to be located in the final VCNA

section. The lanes in Figure 5.12 are, respectively, the crude fraction, the first resulting

supernatant of untargeted protein after 1 round of purification, and the final VCNA

isolate. Typically, the final VCNA isolate is a single band seen just below 50 kD;

however, the discernable bands at the EPO molecular weight indicate that the final

product remained with the construct. The EPO was either not sufficiently eluted with

SA, or it's possible that addition of SA non-specifically precipitated the EPO product.
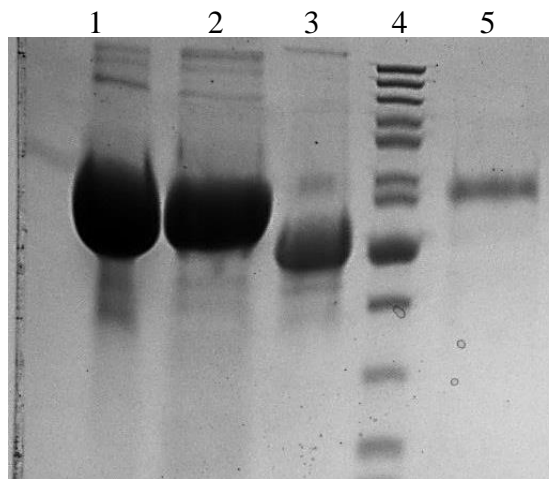
Figure 5.12 SDS-PAGE analysis of VCNA fractions.
Lane 1: EPO in CHO media. Lane 2: Unbound fraction. Lane 3. VCNA-ELP40 and visible EPO.
Lane 4: Molecular weight marker. Lane 5. Failed purified EPO (only visible albumin).

Further indication that EPO was indeed targeted is the reduced visibility of the

EPO band from crude to first supernatant. The samples are identical in concentration and

loaded amount, except for the second lane has been treated with a VCNA-ELP40

purification round. The reduced EPO in the supernatant suggests that is was targeted and

removed from the crude. Based on protein quantification estimates, it seems about 42 ±

5% of the EPO was removed and accounted for in the VCNA stream. The isolated

fraction, as seen in the final column, clearly portrays albumin depletion, upwards of 98%.

However, there is no EPO visible in this fraction.

5.4 Discussion

The goal of this work was to illustrate the efficacy of the designed purification method using relevant therapeutic glycoproteins of clinical interest. To give a good representative test scenario, the therapeutic glycoprotein, EPO, was obtained in final formulation, but submerged in a CHO cell used media solution. Ideally, this would have acted as a realistic case for industrial EPO purification, with similar protein contaminants and concentrations.

However, it's clear that obtaining the EPO prepared as a medicinal deliverable yielded more difficulties than predicted. First of all, the extreme excess of human serum albumin had not been accounted for in preliminary experiments, and although resulting tests in high albumin solutions were relatively successful, the albumin used in those experiments were bovine derived and less likely glycated. The human albumin in the final formulation of EPO likely masked the target, and even with significant depletion, represented a large portion of protein in the putatively purified product. Secondly, the small amount of EPO provided by Amgen only allowed for a few experiments, so continued optimization and varied conditions were not feasible. The preferred method of analysis, SDS-PAGE, did not always yield an amenable detection limit for smaller scale purifications. Limited experiments were done, resulting in merely a proof of concept design and purification.

In spite of these obstacles, fairly promising results were obtained. With regards to the preliminary tests of isolating HRP from high albumin solutions, auspicious yields were achieved. It's clear some purification efficiency is lost due to increased contaminants, but isolation yield was still measured at 83.3% in the highest albumin

158

concentration sample containing 50 mg/mL of albumin. About 98% of the albumin was depleted in this sample. Since the final isolate is in a concentrated solution, the presence of the final albumin is still very significant, albeit greatly reduced. In the case of using VCNA and fetuin in high albumin scenarios, similar results were obtained. Again, isolation yield is affected, but less so than the RSL-HRP system. However, the VCNA does a poorer job of albumin depletion, especially at the higher albumin concentration.

Hopefully, by performing a sequential multi-lectin purification, using both RSL and VCNA sequentially would reduce the amount of albumin dramatically. Assuming a modest reduction of 85% each time, a second round would leave final albumin concentrations at only 2.3% of their original levels. However, these sequential methods were done only with the model glycoproteins, HRP and fetuin. There was not enough EPO sample to sustain further MLAP experimentation.

Finally, the small amount of EPO obtained was used in a proof of concept type purification. EPO formulation was mixed with the used CHO medium and was purified using each of the constructs separately. Purification was acheived at levels of $60 \pm 2.9\%$ and $42 \pm 5\%$ by the RSL and VCNA constructs, respectively, and resulted in albumin depletion of about 85% and 99%, respectively. While this is somewhat below the anticipated isolation ability, it is not uncommonly low. Works often report 15-16% recovery, and some methods with chromatography yield 42% recovery.[20-22] This low recovery is often accepted with the assumption the purity is high, as purity is more valuable for a therapeutic product. While the purity using this purification system does not reach industrial specifications, it is also outside the scope of the method. The method

is designed for targeted, quick, and complete capture from high contaminant solutions, and relative to other published work, the lectin fusion method could be competitive.

5.5 Materials and Methods

Materials

2 vials (10,000 Units/vial) of Epogen were generously donated by Amgen per Amgen Agreement #2014584977. CHO cells and accompanying media was generously cultured and donated by Travis Meyer and Dr. Bao's lab from Georgia Institute of Technology.

Media Preparation

Warmed 1x F-12 Media (with L-glutamine) to 37°C for ~ 1 hour along with 50mL of frozen FBS. Added 50mL of FBS to 500mL of F-12 Media in cell culture hood to prepare 10% FBS CHO Growth Media. F12 media was stored at 4°C, and was warmed to 37°C 30 minutes before use in cell culture experiments.

Thawing Cells

Removed 1mL vial of CHO cells (Brian Wile – CHO 5 – 105 cells/mL, frozen in 2012) and thawed in 37°C water bath for ~10 minutes. Added cells to 9 mL of CHO media and centrifuged at 300xg for 5 minutes at 23°C. Aspirated the supernatant, resuspended in 12mL of CHO media, and transferred cells to a T75 flask. Moved to cell culture incubator (37°C, 5% CO_2).

Passaging Cells

After 48 hours, cells were at ~90% confluence. Aspirated off media, and washed with 10mL of warm PBS (without CaCl2 and MgCl2). Aspirated off the PBS, and added 1 mL of .25% Trypsin-EDTA. Let the cells sit in the trypsin solution for ~5 minutes, and verified the detachment via bright field microscopy. Added 9mL of F12 media and pipetted up and down to remove any weakly attached cells from the surface of the plate. Transferred all 10mL to 15mL conical tube and centrifuged for 5 minutes at 300xg at 25°C. Aspirated off the media, resuspended in 9mL of media, and added 3mL of the cells to 3 different T75 flasks – this made CHO Passage 6. Added 9mL of media to each flask, and then placed in incubator (37°C, 5% CO2).

After 48 hours, cells were at 100% confluence. All of the steps below were completed in triplicate (once for each flask of CHO Passage 6). Aspirated off media and washed with 10mL of warm PBS (without CaCl¬2 and MgCl-2). Aspirated off the PBS and added 1.3 mL of .25% Trypsin-EDTA, and let sit for ~5 minutes at room temperature. Added 9 mL of warm F12 media, and pipetted up and down to remove weakly attached cells. Transferred to 15mL conical tube and centrifuged for 5 minutes at 300xg at 25°C. Aspirated off the media, and resuspended in 8 mL of warm F12 media. Added 2 mL of resuspended cells to 18mL of warm F12 media in new T75 flasks, and placed in incubator (37°C, 5% CO2) to make CHO Passage 7.

After another 72 hours, cells were at 100% confluence. Removed 20mL of media from each (N = 12) of the CHO Passage 7 T75 flasks and transferred to glass container. Transferred conditioned media to conical tubes (4x 50mL and 2x11mL) and centrifuged

at 1000xg for 10 minutes (25°C) to remove dead cells and cellular debris. Decanted the supernatant (leaving 1 mL left at bottom) and transferred to glass bottle.

Freezing Cells

After harvesting conditioned media, took one T75 flask, and rinsed with 10 mL of warm PBS, followed by 2 mL of .25% Trypsin-EDTA, and let sit at room temperature for ~5 minutes. Added 8mL of warm F12 media, transferred to a 15 mL conical tube, and counted cells using automatic cytometer. Centrifuged at 300xg for 5 minutes at 25°C, aspirated media and resuspended in 10% DMSO in F12 media. Transferred cells to cryovial, placed in freezer container, and stored at -80°C overnight. The next day, moved the cells to liquid nitrogen.

5.6 References

1.   Lightfoot, E. N. & Cockrem, M. C. M. Complex Fitness Diagrams: Downstream Processing of Biologicals. Sep. Sci. Technol. 48, 1753–1757 (2013).

2.   Skibeli, V. Sugar profiling proves that human serum erythropoietin differs from recombinant human erythropoietin. Blood 98, 3626–3634 (2001).

3.   Jelkmann, W. Molecular biology of erythropoietin. Intern. Med. 43, 649–59 (2004).

4.   Jacobs, K. et al. Isolation and Characterization of Genomic and cDNA Clones of Human Erythropoietin. Nature 313, 806–810 (1985).

5. Government Accountability Office, U. Testimony Before the Subcommittee on Health, Committee on Energy and Commerce, House of Representatives. Medicare. Inf. Highest-Expenditure Part B Drugs. June 28 (2013).

6. Misaizu, T. et al. Role of antennary structure of N-linked sugar chains in renal handling of recombinant human erythropoietin. Blood 86, 4097–104 (1995).

7. Darling, R. J. et al. Glycosylation of erythropoietin affects receptor binding kinetics: role of electrostatic interactions. Biochemistry 41, 14524–31 (2002).

8. Linsley, K. B. et al. Applications of Electrospray Mass Spectrometry to Erythropoietin N- and O-Linked Glycans. Anal. Biochem. 219, 207–217 (1994).

9. Hermentin, P. et al. The hypothetical N-glycan charge: a number that characterizes protein glycosylation. Glycobiology 6, 217–30 (1996).

10. Yang, M. & Butler, M. Effects of ammonia and glucosamine on the heterogeneity of erythropoietin glycoforms. Biotechnol. Prog. 18, 129–38 (2002).

11. Yang, M. & Butler, M. Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. Biotechnol. Bioeng. 68, 370–80 (2000).

12. Plavina, T., Wakshull, E., Hancock, W. S. & Hincapie, M. Combination of Abundant Protein Depletion and Multi-Lectin Affinity Chromatography (M-LAC) for Plasma Protein Biomarker Discovery research articles. 662–671 (2007).

13. Wang, P. et al. Erythropoietin derived by chemical synthesis. Science 342, 1357–60 (2013).

14. Toyoda, T., Itai, T., Arakawa, T., Aoki, K. H. & Yamaguchi, H. Stabilization of human recombinant erythropoietin through interactions with the highly branched N-glycans. J. Biochem. 128, 731–7 (2000).

15. Hinderer, W. & Arnold, S. Process for the Purification of Recombinant Human Erythropoietin (EPO), EPO Thus Purified and Pharmaceutical Compositions Comprising Same. (2012).

16. Lee, L. Y. et al. An optimized approach for enrichment of glycoproteins from cell culture lysates using native multi-lectin affinity chromatography. J. Sep. Sci. 35, 2445–52 (2012).

17. Yang, Z. & Hancock, W. S. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. J. Chromatogr. A 1053, 79–88 (2004).

18. Arakawa, T., Philo, J. & Kita, Y. Kinetic and Thermodynamic Analysis of Thermal Unfolding of Recombinent Erythropoietin. Biosci. Biotechnol. Biochem. 65, 1321–13 (2001).

19. Amgen. Epogen. MSDS 1–9 (2013).

20. El Khoury, G., Wang, Y., Wang, D., Jacob, S. I. & Lowe, C. R. Design, synthesis, and assessment of a de novo affinity adsorbent for the purification of recombinant human erythropoietin. Biotechnol. Bioeng. 110, 3063–9 (2013).

21. Jeong, T.-H. et al. Soluble expression and partial purification of recombinant human erythropoietin from E. coli. Protein Expr. Purif. 95, 211–8 (2014).

22. Surabattula, R., Rao, K. R. S. S. & Polavarapu, R. An Optimized Process for Expression , Scale-Up and Purification of Recombinant Erythropoietin Produced in Chinese Hamster Ovary Cell Culture. 2, 58–74 (2011).

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE DIRECTIONS

6.1 Conclusions

The work presented in this dissertation accomplished three objectives stated in the introduction: 1) thermo-responsive ligands were designed and expressed extracellularly by *E. coli*, 2) the ligand and accompanying affinity precipitation method were characterized and optimized, and 3) the process was applied to relevant glycoproteins.

6.1.1 Design and Production (Chapter 2)

The design and production of the thermo responsive affinity ligands was achieved. First, a fusion protein with an ELP peptide fused to the RSL lectin was successfully cloned and expressed into *E. coli* for production. Subsequently, two additional constructs with elongated ELP regions were designed and cloned using recursive oligomerization. These were also successfully expressed in the bacterial host. Additionally, one fusion of VCNA lectin to the ELP peptide consisting of 40 pentapeptide repeats was successfully constructed and expressed. Of the four constructs, three were successfully secreted by utilizing the Tat dependent pathway in tandem with a mutant E. coli strain, E609Y. Fractionation of the *E. coli* cell periplasm revealed localization of the largest construct, RSL-ELP60, to the periplasmic fraction, likely due to the large hydrophobic region characteristic of the ELP peptide. However, the other

constructs were successfully produced, and secretion by this method reached production levels of about 20-100 mg/L of recombinant protein over the course of 48 hours. This allows for easy isolation of the desired protein, avoiding inclusion bodies, large scale homogenization steps, and the need for additional protease inhibitors.

6.1.2 Characterization (Chapter 3)

After production of the 3 different length RSL-ELP ligands and VCNA-ELP ligand, various methods were used to ensure the retained functionality of each domain. First, the ELP domain was tested to ensure inverse solubility characteristics were still functional. This was done first by performing inverse temperature cycling on the crude protein solution and confirming the presence of isolated fusion protein. To more accurately determine the transition temperature, turbidity tests were employed, defining the transition temperatures of each of the constructs. The RSL-ELP fusions exhibited phase transitions at 80°C, 45°C, and 36°C respectively to 20, 40, and 60 repeats of ELP pentapeptide. The VCNA-ELP40 fusion exhibited phase transition at 51°C. Further perturbations revealed this transition temperature could be depressed by the addition of salt to the ELP solution to a manageable 37°C.

Lectin functionality was tested using one of two techniques. For the RSL-ELP fusions, functionality was tested using affinity chromatography with a fucose agarose column. Ability to purify the constructs based on their affinity to fucose suggests the binding domain of the lectin is still active and was confirmed. The VCNA was tested by purification of a sialylated glycoprotein, fetuin. An active binding site would bind the

166

glycoprotein, and the isolation of the fusion ligand would result in isolation of the glycoprotein, displaying a functional lectin domain.

The multidomain fusion proteins were tested for the intended use: use as a method of glycoprotein purification. The constructs were used to isolate horseradish peroxidase and fetuin, fucose and sialic acid containing glycoproteins, respectively. Allowing for binding using the lectin functionality, the ELP domain was then exploited using inverse temperature cycling. This resulted in the isolation of the glycoprotein along with the fusion construct. This method was tested with competing glycoproteins and high contaminant solutions to ensure robustness. Recovery yield of the target glycoproteins reached 98% and 95% for HRP and fetuin, respectively.

After characterization of the construct, the method itself was further investigated to optimize operating parameters and minimize energy and materials usage. By using a determined ligand to taget ratio, 0.5 M NaCl in the binding buffer, and 1 mM fucose in the elution buffer, purification of a low abundance glycoprotein could still reach a yield of 95%. Recycle studies were performed as well, maintaining an average of 81.4% recovery over the course of 10 recycles.

6.1.3 Applications (Chapter 4 and 5)

Once the construct and method had been established, this purification system was tested using relevant glycoproteins to portray range of application. First, purification of a robust peroxidase from soybeans from a crude, natural protein extract was investigated. By merely soaking the hulls in an aqueous buffer, adding the fusion constructs, and using the affinity precipitation method, 22.7 fold purification was achievable. Furthermore,

recovery yield was about 95%, and the quality specifications were comparable to the reagent grade product previously sold by Sigma. This purification method resulted in little to no inactivation of the target enzyme, and was performed using simple process steps, such as heating and centrifugation.

The method was then applied to a more industrial scenario to target a medically relevant glycoprotein. Industrial production was mimicked by spiking used CHO cell media with Epogen obtained from Amgen. The ability of the constructs to target EPO in a high contaminant solution was tested, while depletion of contaminants was also monitored. EPO in the replicated used CHO medium was targeted at levels of $60 \pm 2.9\%$ and $42 \pm 5\%$ by the RSL and VCNA constructs, respectively, and resulted in albumin depletion of about 85% and 99%, respectively. A method of multi-lectin affinity precipitation was introduced and tested with model glycoproteins to show enhanced specificity and contaminant depletion.

6.2 Significant Contributions

This dissertation provides significant contributions to the field of protein purification for improving glycoprotein targeting methods. Major contributions were made in the development of affinity-ELP fusion ligands. Whereas previous work employs direct fusion of the ELP to the target protein, the work presented in this thesis suggests a platform for a multi-use construct. Previous use of ELP's requires that the protein of interest be recombinant and transcriptionally fused. Furthermore, this previous approach requires an extra cleavage step, after which the ELP domain is discarded[1–3].

Alternatively, the work described in this thesis advocates the use of a separate, reusable ligand. This not only circumvents the issue of a cleavage step, but allows this construct to be used with natural, unaltered target proteins. Furthermore, the ligand can be recycled; thus extending the use and minimizing total cost.

In regards to the affinity domain, the development of the bacterial lectin was also a novel approach. The vast majority of commercialized lectins are plant derived, with nominal consideration given to bacterial lectins. However, given their small size respective to their plant and mammal parallel, they make an efficient substitute for this affinity ligand construct. Moreover, use of a sugar binding affinity ligand expands the range of applicability of the ligand. In other fusion work, effort has been limited to specific binding domains targeting a single protein of interest[4,5]. But by utilizing a sugar binding module, the construct can be applied to a whole range of glycoproteins. Again, this sets the stage for further development of a library of lectin affinity constructs, used to target a variety of glycoproteins.

Finally, the process itself demonstrates an elegant, simple purification process. With this work, affinity precipitation is extended to the purification of an important category of proteins: glycoproteins. More importantly, the separation is based on the presence of a specific monosaccharide. Further development of the method could lead to a large scale alternative for affinity chromatography. Expression of the ligand by a bacterial host is a start to evolving an economic method of production, should this method be applied industrially. It supplies a substitute to isolation of expensive binding ligands and the immobilization required for chromatography. Research performed in this

169

dissertation could be applied to further advancement of this affinity precipitation method, yielding a simple, cost effective purification system.

6.3 Recommendations for Future Directions

6.3.1 Incorporation of smaller affinity domains

As this method is in the same vein as affinity chromatography, affinity precipitation could stand to benefit from some of the advances that have evolved with chromatographic needs. Specifically, smaller ligands have been designed, and these ligands have been engineering to have a higher affinity toward the target. For example, protein A is notable for its binding capacity for IgG, but a truncated form of this ligand reduces size and non-specific binding while increasing long term stability[6,7]. In this same fashion, further investigation into binding kinetics and protein functionality could result in a smaller, higher affinity ligand. This reduced size would decrease the metabolic burden on the cell during production and promote better binding during the purification process. Furthermore, even though the number of binding sites increases when the RSL forms trimers, the trimer formation is just another parameter to be concerned about. It would be more advantageous to engineer a single protein with multiple binding sites to increase binding efficiency while avoiding the steric hindrance of associated trimers or dimers. This approach from a protein engineering aspect would greatly benefit the field.

6.3.2 Expansion to library of lectins

The work in the thesis pertained to the development of two ligands: a fucose binding lectin and a sialic acid binding lectin. With merely these two designs, a variety of applications were addressed. However, to further expand the range of applicability, an entire library of lectin fusions could be developed. Investigation of different lectins has led to the discovery of their immense variety of diverse carbohydrate targets. Not only are different monosaccharide sugar moieties targeted, but also branched and linear oligosaccharides with differing compositions[8,9]. This would allow for specific isoform isolation capability. With the ability to target based on glycan recognition, a wide array of applications could be pursued. Furthermore, this expanded library of lectin fusions could participate in the MLAP purification scheme. In the case of many MLAC applications, as many as five different lectins have been employed for glycan targeting.[10] This gives the ability to screen for glycosylation changes in biological samples, giving a more comprehensive look at changing glycomes.[11] This could be achieved with MLAP as well, providing a new library of fusion constructs were designed for this application.

6.3.3 Optimization using EPO

Unfortunately, due to limited amount of product, optimization and constraints with the lectin constructs and proposed MLAP method could not be performed with EPO. Only a quick proof of concept, resulting in less than predicted purification yields, could be performed with the obtained samples. While the ratio optimizations for the constructs and their targets were done on model glycoproteins, not enough sample was obtained to optimized these ratios with EPO as the targeted glycoprotein. Furthermore, it appeared as if the optimized elution method with regards to VCNA and fetuin were not amenable

to the VCNA: EPO combination. If more EPO sample could be obtained, further experimentation, as well as multi-lectin targeting, would be possible, hopefully with improved results.

6.3.4 Scale-up

Finally, in order for this technique to become industrially relevant, the process will need to be scaled appropriately and constrained using industrial scale parameters. First of all, the constructs would need to be produced on the order of grams instead of the current milligram scale presented in this thesis. A larger culture volume would need to be employed, as well as some method for controlled temperature ramping and cooling to perform the ITC to collect the construct. By design, the scalability should be relatively straight forward when compared to its counterpart, affinity chromatography. By employing this in-solution method, flowrate constraints and diffusion constraints are avoided. In tandem with these chromatographic drawbacks is an extended time scale, which is limited by the volumetric throughputs[12]. The affinity precipitation method reduces this, though by how much would need to be investigated at a larger scale. The recycling of the construct, though effective, does diminish over time, which would also need to be addressed at the larger scale. While affinity precipitation provides a promising alternative to chromatography, further investigations need to be pursued.

6.4 References

1. Shi, C., Meng, Q. & Wood, D. W. A dual ELP-tagged split intein system for non-chromatographic recombinant protein purification. Appl. Microbiol. Biotechnol. 97, 829–35 (2013).

2. Wood, D. W., Wu, W., Belfort, G., Derbyshire, V. & Belfort, M. A genetic system yields self-cleaving inteins for bioseparations. Nat. Biotechnol. 17, 889–92 (1999).

3. Banki, M. R., Feng, L. A. & Wood, D. W. Simple bioseparations using self-cleaving elastin-like polypeptide tags. Nat. Methods 2, 659–661 (2005).

4. Sheth, R. D., Madan, B., Chen, W. & Cramer, S. M. High-throughput screening for the development of a monoclonal antibody affinity precipitation step using ELP-z stimuli responsive biopolymers. Biotechnol. Bioeng. 110, 2664–76 (2013).

5. Madan, B., Chaudhary, G., Cramer, S. M. & Chen, W. ELP-z and ELP-zz capturing scaffolds for the purification of immunoglobulins by affinity precipitation. J. Biotechnol. 163, 10–6 (2013).

6. Graille, M. et al. 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. Proc. Natl. Acad. Sci. U. S. A. 97, 5399–404 (2000).

7. Gottschalk, B. U. Downstream Processing of Monoclonal Antibodies : from High Dilution to High Purity. BioPharmInternational.com 1–17 (2010).

8. Wu, A. M., Lisowska, E., Duk, M. & Yang, Z. Lectins as tools in glycoconjugate research. Glycoconj. J. 26, 899–913 (2009).

9. Singh, R. S., Tiwary, a. K. & Kennedy, J. F. Lectins: Sources, Activities, and Applications. Crit. Rev. Biotechnol. 19, 145–178 (1999).

10. Yang, Z. & Hancock, W. S. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. J. Chromatogr. A 1053, 79–88 (2004).

11. Yang, Z. & Hancock, W. S. Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. J. Chromatogr. A 1070, 57–64 (2005).

12. Sheth, R. D. et al. Affinity precipitation of a monoclonal antibody from an industrial harvest feedstock using an ELP-Z stimuli responsive biopolymer. Biotechnol. Bioeng. 111, 1595–603 (2014).