

**ENGINEERING THE ASLOV2 DOMAIN FOR
OPTICALLY CONTROLLED PROTEIN CONJUGATION**

by

Justin Michael Terr

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Bachelor of Chemical Engineering with Distinction

Spring 2019

© 2019 Justin Michael Terr
All Rights Reserved

**ENGINEERING THE ASLOV2 DOMAIN FOR
OPTICALLY CONTROLLED PROTEIN CONJUGATION**

by

Justin Michael Terr

Approved: _____
Wilfred Chen, Ph. D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Eleftherios T. Papoutsakis, Ph. D.
Committee member from the Department of Chemical & Biomolecular
Engineering

Approved: _____
Susan E. Groh, Ph. D.
Committee member from the Board of Senior Thesis Readers

Approved: _____
Earl Lee II, Ph.D.
Deputy Faculty Director, University Honors Program

ACKNOWLEDGMENTS

During my four years of undergraduate study at The University of Delaware, I have been lucky enough to meet amazing friends and colleagues of whom I am very thankful for. I would also like to thank the University for allowing me to pursue interests outside of my field of study, such as a minor in political science, and providing an outstanding platform to grow both professionally and socially. In addition to undergraduate research, I was given the opportunity to study abroad in Australia, become a Blue Hen Ambassador, and hold leadership positions in various organizations. These experiences greatly supplemented my education and will have a lasting impact on me.

First, I must acknowledge the parties that made my senior thesis possible. This would be the Office of Undergraduate Research & Experimental Learning (UREL) and my committee. The UREL has been very helpful in my research since the beginning. The Summer Scholar Program supported my summer research for two years and allowed me to gain practice in presenting and making posters. This experience prepared me to present my research at the American Institute of Chemical Engineering (AIChE) student conference, where I placed 2nd in my group. Dr. Barsky and the rest of the staff at the UREL were extremely helpful through the entire thesis process as well. In addition to being there during times of stress, they made the program's deadlines and expectations very clear from the beginning. I was also lucky enough to be assigned to Dr. Susan Groh, my freshman chemistry professor, as my

third reader. It was satisfying to have my college experience come full circle with Dr. Groh.

None of the work presented in this thesis would have been accomplished without the guidance of Dr. Wilfred Chen, Emily Harztel, and all other members of the Dr. Chen's research group. After I decided to pursue chemical engineering research in a biological setting, I reached out to a few professors whose research interested me. After I met with Dr. Chen, he took me into his group as a rising sophomore and paired me with one of his Ph. D. candidates, Emily Harztel. From there, Emily took me under her wing and began to teach me various lab molecular biology laboratory techniques. These skills were applied to a basic recombinant DNA project to produce a snake anti-venom peptide in *E.coli*. Emily was always there and took time out of her day to answer any questions and address any concerns I could not have asked for a better graduate student mentor. After sparking my interest in protein engineering, the following summer I took on the initial LOV-Spy project presented in Chapter 2. This work would be continued through the end of my senior year and was the basis of my thesis.

In addition to my research, as a whole, the Chemical & Biomolecular Engineering department has played a pivotal role in my undergraduate career. Dr. Joshua Enszer and Ms. Megan Argoe provided great undergraduate advisement in class choices. Specifically, I greatly enjoyed my biochemical engineering elective classes taught by Dr. Wilfred Chen, Dr. Eleftherios Papoutsakis, and Dr. Millicent Sullivan. These classes maintained my interest in the subject and excites me for what I can do in the future.

Last, but not least, I would like to thank my friends and family for their daily support and companionship. In addition to financially supporting me up until this point, my parents have acted as my rock and have always been there for me. I appreciate all of the sacrifices my parents have made for me, and hope to be half as good of a parent one day. As my sister Lauren graduates high school next year, I hope to be a good role model for her as she begins to determine what interests to pursue. To my roommates, classmates, close friends, and fraternity brothers, it has been a pleasure to share these past four years with you. Filled with laughter and sometimes tears, I can't imagine this journey going any other way.

Thank you all for everything. I cannot wait to see what the future holds for myself and all those that I have encountered!

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES.....	x
ABSTRACT	xv
1 BACKGROUND & SIGNIFICANCE.....	1
1.1 Protein Engineering & Conjugation	1
1.2 SpyCatcher:SpyTag Peptide Conjugation System	3
1.3 Sortase A Protein Conjugation System	6
1.4 Light in Biological Systems	8
1.5 AsLOV2: Incorporation of Light Dependence.....	10
1.6 Thesis Overview	11
2 LOV IS BLISS: CREATING A BLUE LIGHT INDUCED SPYCATCHER SYSTEM	13
2.1 Introduction	13
2.2 Methods & Materials.....	15
2.2.1 Molecular Cloning.....	15
2.2.1.1 Obtaining Template Plasmid & PCR Design	15
2.2.1.2 Polymerase Chain Reaction (PCR)	16
2.2.1.3 Restriction Enzyme Digest & Agarose Gel Electrophoresis	16
2.2.1.4 DNA Ligation.....	17
2.2.1.5 Bacterial Transformation for Plasmid Preparation.....	18
2.2.1.6 Cloning Confirmation and Plasmid Extraction	18
2.2.2 Protein Expression & Purification.....	19
2.2.2.1 Bacterial Transformation for Protein Expression.....	19
2.2.2.2 Protein Expression.....	19
2.2.2.3 Protein Characterization	20
2.2.2.4 Protein Purification.....	21
2.2.3 Light Sensitive Experimentation	22
2.3 Results & Discussion.....	22
2.3.1 Generation of BLISS Library	22
2.3.2 Protein Expression & Purification.....	24

2.3.3	Blue Light Dependent Experiments	27
2.4	Summary and Conclusions	31
3	SPATIAL AND TEMPORAL CONTROL OF REVERSIBLE SORTASE A TRANSPEPTIDATION WITH THE LOV DOMAIN	33
3.1	Introduction	33
3.2	Methods & Materials	34
3.2.1	Molecular Cloning	34
3.2.1.1	PCR Design	34
3.2.1.2	Polymerase Chain Reaction (PCR)	35
3.2.1.3	Restriction Enzyme Digest & Agarose Gel Electrophoresis	35
3.2.1.4	DNA Ligation	35
3.2.1.5	Bacterial Transformation for Plasmid Preparation	35
3.2.1.6	Cloning Confirmation & Plasmid Extraction	35
3.2.2	Protein Expression & Purification	35
3.2.2.1	Bacterial Transformation for Protein Expression	35
3.2.2.2	Protein Expression	36
3.2.2.3	Protein Characterization	37
3.2.2.4	Protein Purification	37
3.2.3	Functionality Experiments	38
3.2.3.1	Blue Light Independent Sortase Experiments	38
3.3	Results & Discussion	38
3.3.1	Generation of Constructs for Blue Light Induced Sortase A Transpeptidation	38
3.3.2	Protein Expression & Purification	39
3.3.3	Optimization of Sortase Reaction	42
3.4	Summary & Conclusions	43
4	CONCLUSIONS	45
4.1	Conclusions	45
	REFERENCES	47

A	DETAILED PROCEDURE AND RESULTS FOR GENERATION OF BLISS SYSTEM CONSTRUCTS	50
B	DETAILED PROTEIN EXPRESSION & CHARACTERIZATION RESULTS FOR BLISS SYSTEM	57
C	DETAILED PROCEDURE AND RESULTS FOR GENERATION OF LOV-CONSTRUCTS FOR SORTASE TRANSPEPTIDATION.....	61
D	ADDITIONAL RESULTS FOR EXPRESSION AND PURIFICATION OF LOV-CONSTRUCTS FOR SORTASE TRANSPEPTIDATION.....	67

LIST OF TABLES

Table 1:	Expected Size of SpyTag:SpyCatcher Reactions.....	27
Table 2:	Expected Size bands of Engineered Proteins	39
Table 3:	Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-SpyTag construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow), Sall (blue), HindIII (green). The DNA to code for the SpyTag peptide is also highlighted in pink.	50
Table 4:	Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-SpyCatcher construct. Regions homologous to the template plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow) & HindIII (green).	54
Table 5:	Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-LPXTGGSG construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow) & HindIII (green). The DNA to code for the LPXTGGSG (red) peptide and the STOP (black) codon is also highlighted.	61
Table 6:	Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-LPXTGGSG construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow). HindIII (green), SacI (royal) , EcoR1 (navy), NdeI (purple), & XhoI(grey).	65

LIST OF FIGURES

- Figure 1: Cartoon illustration of the splitting of CnaB2 domain. After further optimizations, the split domain resulted in a large N-terminal protein (SpyCatcher) and a small C-terminal peptide (SpyTag). Reactive residues are highlighted in red. Image reproduced from Reference (10). 4
- Figure 2: Illustration of isopeptide bond formation between SpyTag and SpyCatcher proteins. The reactive Aspartic Acid group of SpyTag reacts with the Lysine group of SpyCatcher. Image reproduced from Reference (11). 5
- Figure 3: Illustration of Sortase A mediated transpeptidation in its native form. A reactive thiol group in the Sortase Enzyme reacts with the C-terminal carboxy group of Protein A's conserved LPXTG sequence. Sortase cleaves the amide bond between threonine and glycine to form an intermediate. This intermediate then undertakes a nucleophilic attack from an oligoglycine substrate and forms a stable amide bond. Reproduced from Reference 13..... 7
- Figure 4: Previously engineered example of the AsLOV2 domain being used to block an active site. The LOV domain was used to successfully control the activity in GTPase Rac1 by blocking its effector domain (star) in LOV's dark state. A 10-fold difference in binding to the domain in the lit state was reported. Reproduced from Reference (17). 11
- Figure 5: Illustration of discussed constructs. In the top image, SpyTag is genetically fused to the $J\alpha$ -helix and in the bottom image SpyCatcher is fused. Once irradiated by blue light, a conformational change should occur within the $J\alpha$ -helix which allows undocking, allowing the reaction to proceed. 14
- Figure 6: 2% agarose gel of digested LOV-540-SpyTag constructs. Constructs were digested with HindIII & BamHI. Lane Profile: (A) 2 log ladder, (B-H) Ligated clones of LOV-540-SpyTag. Positive clones with a band at .465 kb can be seen in lanes B, D, E, F, & G. 23
- Figure 7: 2% agarose gel of digested LOV-541-SpyCatcher constructs and iLID (cloning template). Constructs were digested with HindIII & BamHI. Lane Profile: (A) 2 log Ladder, (B-E) LOV-541-SpyCatcher ligated clones, (F) iLID control. It is observed that the positive clones in this screening are in lanes C, D, & E with a band corresponding to .675kb. iLID control corresponds to correct band at .456kb..... 24

Figure 8:	12.5% SDS PAGE of results iLID expression and purification via conditions reported in literature. Lane profile: (A)Ladder, (B)Soluble Lysate, (C) Elution Sample.	25
Figure 9:	12.5% SDS PAGE of expression and purification results of LOV-X-SpyTag constructs. Lane profile: (A) Ladder, (B-D) LOV-537-SpyTag, (E-H) LOV-540-SpyTag, (I-K) LOV-541-SpyTag. Color key: Green (Insoluble Lysate), Blue (Soluble Lysate), Black (Elution). ..	26
Figure 10:	12.5% SDS PAGE of expression and purification results of LOV-541-SpyCatcher constructs. Lane profile: (A) Ladder, (B) Insoluble Lysate (C) Soluble Lysate, (D) Flow Through 1, (E) Flow Through 2, (F) Elution.	26
Figure 11:	12.5% SDS PAGE of LOV-X-SpyTag + SpyCatcher light dependent conjugation experiment. Lane Profile: (A) Standard, (B-D) LOV-537-SpyTag, (E-G) LOV-540-SpyTag , (H-J) LOV-541-SpyTag. Color Key: White (Time 0), Black (30 minutes in dark), Blue (30 minutes in blue light).	28
Figure 12:	12.5% SDS PAGE of LOV-X-SpyCatcher + SpyTag light dependent conjugation experiment. Lane Profile: (A) Standard, (B-D) LOV-537-SpyCatcher, (E-G) LOV-540-Spy-Catcher , (H-J) LOV-541-SpyCatcher. Color Key: White (Time 0), Black (30 minutes in dark), Blue (30 minutes in blue light).	28
Figure 13:	12.5% SDS PAGE of LOV-537-SpyTag + SpyCatcher light dependent conjugation experiment. Lane Profile: (A) Standard, (B-N) LOV-537 + SC reaction. Color Key: White (Time 0), Black (No Blue Light), Blue (Constant Blue Light).	30
Figure 14:	Reconstitution of LOV-537-SpyTag + SpyCatcher complex. Reconstitution was determined by averaging the amount reacted of the LOV-537-SpyTag & Spy-Catcher constructs. Time points were taken at 1 minute, 10 minutes, 30 minutes, 1 hour, 3, hours, 5 hours.....	31

Figure 15:	Proposed schematic of how Sortase transpeptidation system would be controlled with blue light irradiation. When LOV-LPXTG undergoes blue light irradiation, the Ja (orange) should become undocked. When in solution with SrtA (black) and an oligoglycine substrate (GGG-mRuby2), transpeptidation should occur. To test reversibility, if light is removed, and then reapplied with the addition of another substrate to solution, the substrates should become interchangeable. Reversibility would be tested with FRET. (Right, Upper) A FRET response would be expected with an mRuby2 conjugation. (Right, Lower) FRET response would weaken with the addition of a nonfluorescent substrate as it replace mRuby2.	34
Figure 16:	2% agarose gel of digested GGG-mCherry, GGG-mRuby2, & GGG-ELP constructs. Constructs were digested with EcoRI & XhoI. Lane Profile: (A) 2 log ladder, (B-D) Ligated GGG-mCherry clones, (E-G) Ligated mRuby2 clones, (H) GGG-ELP. Positive clones are observed for all constructs screened.	39
Figure 17:	12.5% SDS PAGE of expression & purification results. Color key: Blue (Soluble Lysate), Black (Elution/Purified). Lane Profile (Left): (A) Ladder, (B-C) GGG-mRuby2, (D-E) GGG-ELP, (F-G) ELP-LPETG. Lane Profile (Right): (A) Ladder, (B-C) GGG-mCherry, (D-E) LOV-537-LPITG, (F-G) LOV-540-LPETG, (H-I) LOV-541-LPETG.....	41
Figure 18:	Light Independent Sortase A Transpeptidation. Color key: Green (4 hr Sortase A reaction) Grey (no calcium negative control). Lane profile: (A) Ladder, (B) GGG-mRuby2, (C) GGG-ELP, (D) ELP-LPETG, (E) Sortase A, (F-H) Sortase A + GGG-mRuby2 + ELP-LPETG, (I-K) Sortase A + GGG-ELP + ELP-LPETG. Transpeptidation band observed in lane G at about 49kDa.....	43
Figure 19:	2% agarose gel of LOV-Tag 537 PCR confirmation. Lane Profile: (A) 2-Log ladder, (B) PCR Sample, (C) PCR control, contaminated by plasmid	51
Figure 20:	2% Agarose gel of digested LOV-X-SpyTag Constructs. Lane Profile: (A-B) 2 Log Ladder, (C) LOV-537-SpyTag, (D) LOV-540-SpyTag, (E-F) LOV-541-SpyTag. It is observed that the only positive clone in this screening is LOV-541-SpyTag in lane 5.	53

Figure 21:	2% agarose gel of PCR amplification of products of Gibson Assembly. Lane Profile: (A) 2-Log Ladder, (B) LOV-SC-537 Gibson Insert, (C) LOV-SC-537, (- control, no template), (D) LOV-SC-540 Gibson Insert, (E) LOV-SC-540, (- control, no template), (F) LOV-SC-541 Gibson Insert, (G) LOV-SC-541, (- control, no template).....	55
Figure 22:	2% Agarose gel of digested LOV-X-SpyCatcher Constructs. Lane Profile: (A) 2 Log Ladder, (B-D) LOV-537-SpyCatcher A-C, (E-G) LOV-540-SpyCatcher A-C, (H-J) LOV-541-SpyCatcher A-C. Positive clones are observed in LOV-537-SpyCatcher A & LOV-540 A, B, & C.....	56
Figure 23:	SDS PAGE of LOV-X-SpyCatcher purification process. Lane profile: (A) Precision Plus Ladder, (B-F) LOV-537-SpyCatcher (Insoluble Lysate, Soluble Lysate, Flow Through 1, Flow Through 2, Elution 1, Elution 2), (H-N) LOV-540-SpyCatcher (Insoluble Lysate, Soluble Lysate, Flow Through 1, Flow Through 2, Elution 1, Elution 2).	58
Figure 24:	12.5% SDS PAGE of LOV-540-SpyTag + SpyCatcher-ELP Experiments. Blue lanes show samples irradiated in blue light and black lanes show samples hidden from light.....	59
Figure 25:	12.5% SDS PAGE of LOV-541-SpyTag + ELP-SpyCatcher Experiment. Blue lanes show samples irradiated in blue light and black lanes show samples hidden from light.....	59
Figure 26:	LOV-X-SpyTag Blue Light Experimental controls. Lane Profile: No ladder shown, (A-E) Negative control reacting SpyCatcher + iLID, (F-J) Positive control reacting SpyTag + SpyCatcher.....	60
Figure 27:	2% Agarose gel for PCR results for amplification of inserts for LOV-540-LPETG & LOV-541-LPETG. Lane Profile: (A) 2 Log Ladder, (B-D) LOV-540-LPETG [sample, no plasmid control, no primer control], (E-G) LOV-541-LPETG [sample, no plasmid control, no primer control].....	62
Figure 28:	2% Agarose gel for PCR results for amplification of inserts for LOV-537-LPETG Lane Profile: (A) 2 Log Ladder, (B) sample, (C) no primer control, (D) no plasmid control.	62

- Figure 29: DNA Sequencing Results for LOV-541-LPETGGSG. Alignment of sequences used via Serial Cloner software. The top sequence represents the sequenced DNA and the bottom is the theoretical construct. Green region represents the N-terminal Histidine Tag and the Purple represents the LPETGGSG motif. 64
- Figure 30: 2% Agarose gel for PCR results. After the ladder all lanes go in the following order: sample, no plasmid control, no primers control Lane Profile: (A) 2 Log Ladder, (B-D) LOV-537-LPETG, (E-G) GGG-mCherry, (H-J) GGG-mRuby2, & (K-M)GGG-GFP. Expected sizes are confirmed for all constructs. 66
- Figure 31: Expression and purification of Z-Ct-LPETG. Lane profile: (A) Ladder, (2) Soluble Lysate, (3) Insoluble Lysate, (4) Purified, (5) Resolubilized Pellet. Protein can be seen at about 29.7 kDa. 67
- Figure 32: Sortase A reaction at 60mM GGG-X substrate and 20mM ELP-LPETG & Sortase A. Lane profile: (A) Ladder, (B) GGG-mRuby2, (C) GGG-ELP, (D) ELP-LPETG, (E) Sortase A, (F-H) Sortase A + GGG-mRuby2 + ELP-LPETG, (I-K) Sortase A + GGG-ELP + ELP-LPETG. Transpeptidation band observed in lane G at about 49kDa. 67

ABSTRACT

The ability to specifically conjugate proteins together has been a useful tool in cellular monitoring and controlling expression. The SpyTag:SpyCatcher systems has been developed as a robust, irreversible peptide tagging system used *in vitro* and *in vivo*. The Sortase A transpeptidation system, derived from Gram-positive bacteria, is a reversible, enzyme mediated conjugation system with applications in purification and tagging. We can imagine that if we can control the conjugation event in time and space, the system could become more versatile with a metaphorical “on/off” switch.

Learning from photoreceptors found in nature, constructs were generated to control protein conjugation events. Blue light dependence was introduced via the light-oxygen-voltage 2 domain from *Avena sativa* (AsLOV2). When irradiated with blue light the domain undergoes a dramatic conformational change that renders the C-terminus “undocked”. Blue light dependent conjugation was tested by genetically fusing the AsLOV2 domain to a reactive partner of either the SpyTag:SpyCatcher system or Sortase A system. Fusion of SpyTag to LOV showed the best results for the irreversible system. After 5 hours, the conjugation reaction was 3-fold more complete with blue light irradiation. The Sortase A system has been generated and expressed but has not been tested. If deemed feasible, application of these types of systems include but are not limited to intracellular protein localization and the dynamic decoration of hydrogels.

Chapter 1

BACKGROUND & SIGNIFICANCE

1.1 Protein Engineering & Conjugation

In nature, proteins serve as versatile biomolecules in both function and application. They provide structure and storage, catalyze chemical reactions, and act as sensory messengers of internal and exterior stimuli.¹ With the advent of modern biological methods, scientists are starting to rationally engineer proteins to do specific functions. For example, protein engineering has been used to improve the catalytic activity of certain enzymes,² improve the kinetics of multienzyme pathways,³ and overcome drug delivery challenges.⁴

Proteins are the final biomacromolecule in the central dogma of biology. In protein engineering, the desired product is the fully folded protein, but the engineering occurs at the DNA level. Proteins are the desirable products due to their various functions. The structure of proteins is analyzed on four levels, from a primary to a quaternary structure. These levels of structure ultimately dictate the function of the protein. The order of the monomeric subunits, or amino acids, in a protein determines its primary structure. Folding occurs at the secondary and tertiary levels to form stable domains. The interactions of these domains with each other form fully active complexes. Scientists became interested in studying protein complexes because of their versatility in function. Since the understanding of protein structure has been widely developed, engineers can now rationally tune the structure to impart desirable properties on the resultant protein.

In addition to a protein's individual role, nature often utilizes assemblies of distinct proteins to expand the functionality of the unit to take on more complex task. In nature, methods have been evolved to conjugate the monomers with high precision and efficiency.⁵ The assembly of proteins is essential to the formation of antibody-protein complexes⁶ and fatty acid synthase (FAS) complexes.⁵ Following nature's lead, scientists began to develop methods to synthetically assemble proteins with multi-component architecture.

In creating these synthetic biological assemblies, stable, covalent linkages are often desired. Classical recombinant methods involve designing DNA sequences that would encode the desired protein complex as a genetic fusion. For example, fusing Green Fluorescent Protein (GFP) to a protein of interest allowed for the study of protein expression and localization in living systems. While genetic fusions work in many examples, these do not always result in successful protein expression due to issues with proper folding. Since many protein complexes are naturally formed by post-translational mechanisms, more sophisticated protein modification methods have been developed to mimic nature's solution. These include chemical methods, which use conjugation chemistries on reactive side chains, and biological methods, which use biocompatible enzymatic conjugation methods on peptide and protein motifs.⁷

The ideal chemical modification for this method uses chemioselective, stereoselective, and site-selective chemistries. This typically involves reacting the amine groups in a lysine or thiol in cysteine to specific target molecules.^{8, 5} Unfortunately, chemical modifications, by nature, tend to be residue specific but not site specific, which leads to lower yields when compared to biological methods. The amino acids that contain the reactive side groups are found with low frequency on

protein surfaces.⁵ When treating an entire protein with a set of reaction conditions, all reactive residues will react. This can also lead to unwanted off-target reactions that could impact the folding and function of the entire protein.

Unnatural amino acid (UAA) incorporation has been utilized to address these issues by uniquely hijacking the cell's translational machinery.⁵ UAAs get around the specificity problem presented with chemical methods by incorporating a UAA with a biorthogonal side chain, such as an azide or alkyne, at a specific site for "click-chemistry".⁵ While these methods solve issues with site-specificity, the efficiency of UAA incorporation leads to lower yields in protein expression. Another downside to UAA protocols is that their incorporation into mammalian cells has also proved challenging, limiting the applications available.⁹

Enzyme mediated conjugations improve on chemical methods in that they provide site-specificity using small genetically encoded peptide motifs and react with high selectivity and yields under biologically favorable conditions.⁷ The wide scope of product possibilities makes this conjugation technology promising. Following the development of GFP, better self-modifying enzymes were developed to act as protein markers with more functionality. These enzyme-based approaches showed high specific labeling with reliable long-term experiments.⁷ An example of this is the Sortase A conjugation system, which is described in detail later on.

1.2 SpyCatcher:SpyTag Peptide Conjugation System

A commonly used, recently developed, peptide tagging system in literature is the SpyCatcher:SpyTag conjugation system. The SpyCatcher:SpyTag system was developed from an extracellular protein in Gram-positive bacteria *Streptococcus pyogenes*. It has been shown that many Gram-positive bacteria contain extracellular

proteins that are stabilized by spontaneous intramolecular isopeptide bonds.¹⁰ The SpyCatcher:SpyTag system is derived from the splitting of the second immunoglobulin-like collagen adhesion domain (CnaB2). This is part of the fibronectin binding protein FbaB necessary for phagocytosis-like events. After identification of reactive functional groups, CnaB2 was split into a C-terminal peptide tag containing the reactive aspartic acid and its covalent reaction partner with a reactive lysine. With the addition of further rational modifications, a highly robust and irreversible peptide tagging system was created.¹⁰ See Figure 1 below for an illustration of how the CnaB2 domain was split. In addition, Figure 2 shows an illustration of the covalent reaction of the two partners.

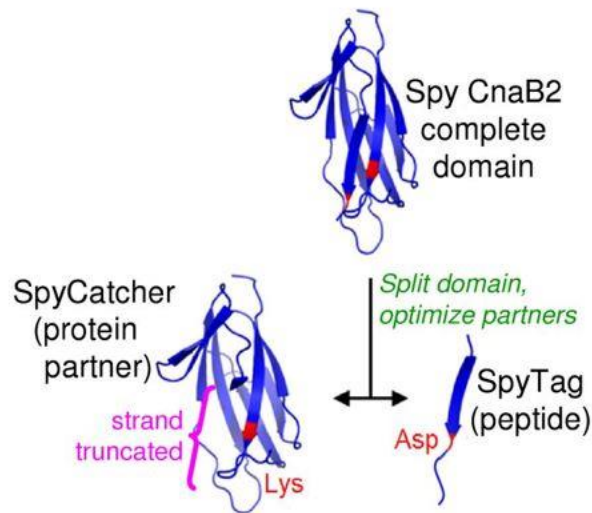


Figure 1: Cartoon illustration of the splitting of CnaB2 domain. After further optimizations, the split domain resulted in a large N-terminal protein (SpyCatcher) and a small C-terminal peptide (SpyTag). Reactive residues are highlighted in red. Image reproduced from Reference (10).

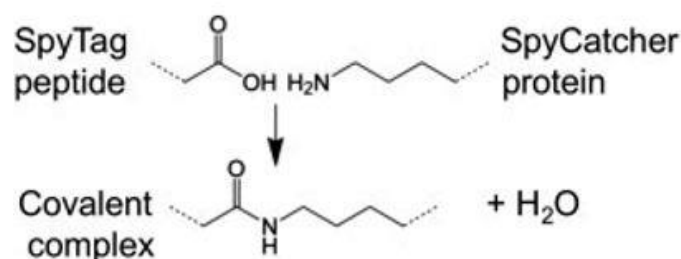


Figure 2: Illustration of isopeptide bond formation between SpyTag and SpyCatcher proteins. The reactive Aspartic Acid group of SpyTag reacts with the Lysine group of SpyCatcher. Image reproduced from Reference (11).

The isopeptide bond formation has since been further studied and characterized. The system showed success both *in vitro* and *in vivo* environments. The isopeptide bond was shown to be robust over a range of pH values and temperatures.¹⁰ The reaction was also shown to happen with very quick kinetics, with a half-time of 74s when 10 μ M of each partner were incubated with each other.¹¹ It has also been shown that SpyTag can be placed at the N-terminus, C-terminus, or at an internal position of a protein, giving it additional versatility over other protein tagging systems.¹¹

This has been an impactful technology. The founding group has shown the SpyCatcher:SpyTag complex useful in stabilizing enzymes and creating modular “plug and play” vaccines. The addition of the highly durable and specific covalent bond between two partner proteins gives them higher probability of remaining attached in denaturing conditions. New biomaterials have also been developed using the SpyCatcher:SpyTag system. The system allows for the creation of large multi-dimensional, multi-component architectures. Adding multiple SpyTag or SpyCatcher sites or incorporating the system with other conjugation pairs has allowed for the

creation of complex self-assembling hydrogels and catalytic biofilms. These technologies show great promise in the fields of industrial enzyme application, tissue engineering, and dynamic protein assemblies.¹²

1.3 Sortase A Protein Conjugation System

Sortase transpeptidases are essential enzymes found in most Gram-positive bacterium. This classification of enzyme derives its name from its ability to sort different useful peptides to the peptidoglycan wall of its host. These enzymes can be generalized into three different isoforms, each with their own distinct properties and functions. Once the correct cell wall sorting signal (CWSS) is recognized by the Sortase, it facilitates a covalent attachment to a protein on its peptidoglycan wall. Although each isoform of Sortase recognizes a different CWSS's, all the isoforms have a highly conserved active site with several charged residues and a terminal motif of TLXTC. Sortase A (SrtA) is the most abundant of the isoforms in nature and of the isoforms and is considered the “housekeeping Sortase”.¹³

The mechanism that Sortases use to facilitate transpeptidation has only been elucidated for SrtA at this point. Sortases must be able to recognize both peptide substrates for it to facilitate the amide bond formation. The donor substrate, containing the CWSS motif, is cleaved at a specific location to expose a C-terminal carboxyl group. The acceptor substrate's N-terminal amine group is then used as a nucleophile to attack the carboxyl group and form the amide bond.¹³ In SrtA specifically, the enzyme recognizes the amino acid sequence of LPXTG (where X is any amino acid other than proline) and cleaves the bond between threonine and glycine. This results in a thio-acyl intermediate. It is resolved with an oligoglycine subunit that results in an amide bond formation between the substrate and the

intermediate.¹⁴ See Figure 3 below for an illustration of the mechanism Sortase A undergoes to complete transpeptidation.

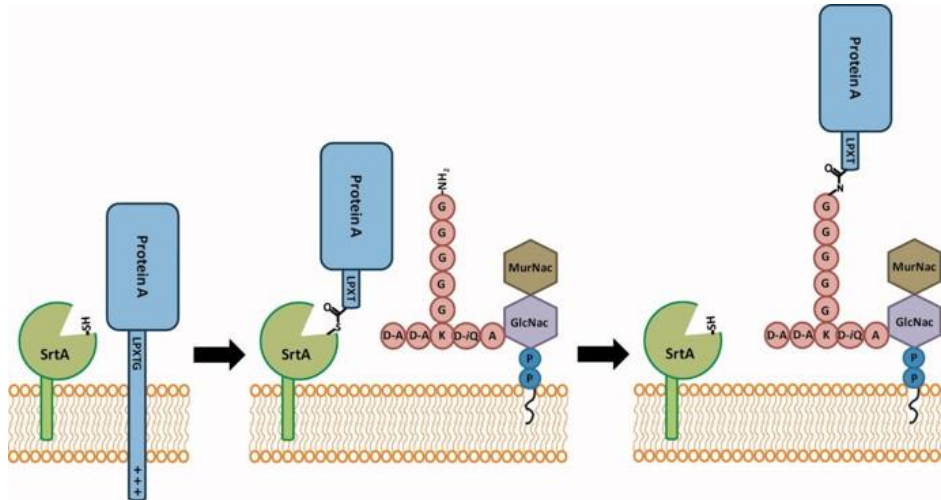


Figure 3: Illustration of Sortase A mediated transpeptidation in its native form. A reactive thiol group in the Sortase Enzyme reacts with the C-terminal carboxy group of Protein A's conserved LPXTG sequence. Sortase cleaves the amide bond between threonine and glycine to form an intermediate. This intermediate then undertakes a nucleophilic attack from an oligoglycine substrate and forms a stable amide bond. Reproduced from Reference 13.

Modifying Sortase A's biological function, the enzyme is currently used for the site-specific conjugation of various markers and functional molecules onto proteins of interest. It differs from the SpyCatcher:SpyTag system in that the amide bond formed is reversible; this allows for more dynamic systems, such as drug delivery and degradation. The Sortase A technology has been successful in site specific protein labeling,¹⁵ protein purification¹⁶ and immobilization.¹⁵ It is a

relatively small enzyme, allowing for easy diffusion and access to reactive sites in proteins.¹⁴

1.4 Light in Biological Systems

Light can serve as a powerful “reagent” in a chemical or biological process. It is easily attainable at various wavelengths and intensities. Differing from a classical chemical stimulus, light can be used with high precision of space and time. Light retains the ability to permeate most biological membranes while remaining noninvasive at most wavelengths.¹⁷ Additionally, light is orthogonal to most biological systems, and leaves the samples traceless, without contamination.¹⁸ These properties are especially important in protein chemistry because of their sensitivity to denaturation.

Evolution has afforded many organisms the ability to incorporate light into some of their biological processes. For example, light is used as a source for energy in photosynthesis. It is also used to relay information in the form of photoreceptors that can elicit a physiological response.¹⁷ In native biological processes, chromophores imbedded into a protein absorb a photon that may induce a conformational change caused by a chemical transformation.¹⁸ Chromophores are small organics or peptides that can be bound by both covalent and noncovalent forces.¹⁸ Light sensitive proteins can be characterized as either fluorescent proteins or photoreceptor proteins. Fluorescent proteins, like GFP, have been extensively studied due to their abundance in certain organisms. Photoreceptors, which serve as information carriers and control cellular behavior, have lagged in understanding. While fluorescent proteins have improved cellular monitoring, photoreceptor proteins present an opportunity in synthetically engineering cellular dynamics and physiology.¹⁷

Due to advantages light has in controlling chemical reactions, photochemistry has developed many successful strategies for different applications. Light sensitive chemistries have been used to cage a small molecule with group that would cleave when irradiated, delivering the attached small molecules.¹⁹ This process of photolysis is usually irreversible and although the initial cleavage may be ultrafast, downstream effects may require additional time.¹⁷ Reversible chemical methods have also been observed in light dependent conversion between *cis* and *trans* isomers of different organic molecules. Although these chemical methods are ultrafast and highly versatile, they have proven difficult for *in vivo* use and can cause the formation of reactive intermediates.

The challenges of broadening these chemical methods to biological systems has been addressed by incorporating natural and engineered photoreceptors. The advantages of these proteins include their native reversibility in their response to irradiation and their ability to be genetically encoded rather than delivered directly. Commonly used photoreceptors are the LOV (light-oxygen-voltage) domains and phytochromes. Both the LOV domain and phytochromes are derived from plant, fungal, and bacterial hosts. Phytochromes absorb red & far-red light. Behaving somewhat similarly to LOV domains, which are utilized in this work, the domain transitions between a red-light absorption state (P_R) and a far-red light absorption state (P_{FR}). Like the LOV domain, phytochromes have been successfully implement for light dependent protein expression and other cellular functions.¹⁷ However, LOV's on and off dynamics are much better than phytochrome, which require infrared light to be returned to ground state.¹⁷

1.5 AsLOV2: Incorporation of Light Dependence

LOV domains are a part of the broader group of Pert-Arnt-Sim (PAS) domains. The PAS domain is observed in all kingdoms of life, and not just in organisms with light dependent physiology. The purpose of these proteins is to facilitate downstream responses to environmental stimuli.²⁰ The PAS core in LOV domains is made up of five antiparallel beta sheets and several alpha helices. An essential cofactor for proper function is it must be noncovalently bound in the form of either flavin mononucleotide (FMN) or flavin adenosine dinucleotide (FAD).¹⁷ In signaling, when irradiated with blue light, a photon is absorbed by the flavin nucleotide cofactor. This leads to a thioether bond between an aromatic carbon in the nucleotide with a conserved cysteine in the PAS core.

The light dependent protein utilized in this study is derived from the plant species *Avena sativa*. The AsLOV2 domain is the most well understood of the LOV domains identified so far. In this specific form, light absorption via FMN promotes the unfolding of the C-terminal J α helix. In the dark state this helix is tightly packed onto the Beta sheet.¹⁷ The reversible docking ability of the J α helix has been advantageous in various optogenetic tools. By appending a peptide of interest to the C-terminus, the AsLOV2 domain maintains the ability to allosterically block the peptide in the dark and releasing it in the light. This technique has been used successfully to control the reactivity of peptides of interest of various sizes by being able to block the key reactive amino acids in the dark state. Protein conjugation can be initiated with the irradiation of blue light.



Figure 4: Previously engineered example of the AsLOV2 domain being used to block an active site. The LOV domain was used to successfully control the activity in GTPase Rac1 by blocking its effector domain (star) in LOV's dark state. A 10-fold difference in binding to the domain in the lit state was reported. Reproduced from Reference (17).

1.6 Thesis Overview

The goal of the work presented is to show whether constructs engineered with the AsLOV2 protein domain would gain the ability to optically control commonly used protein conjugation methods. Using molecular cloning methods, a variety of constructs was created that genetically fused a reactive conjugation partner to the LOV2 domain. Five key amino acid residues have been identified as suitable truncation points into the $J\alpha$ helix for a peptide of interest, three of which were identified to be studied.²¹ Truncation sites were chosen to incorporate the new peptide while still retaining necessary residue interactions between the PAS core and the $J\alpha$ helix. It was hypothesized that in the dark state, the reactive site of the C-terminal protein would be sterically blocked, but when irradiated, the appended protein would be uncaged and free to react with its partner.

In Chapter 2, this concept is applied to the SpyCatcher:SpyTag conjugation system. LOV constructs are cloned with either the protein (SpyCatcher) or peptide (SpyTag) partner. Constructs are expressed and light dependent conjugation is observed. In Chapter 3, the same concept is applied to the Sortase A reversible

conjugation system to make the design more dynamic. The Sortase A recognition sequence, LPXTG, is genetically fused to the LOV domain. Constructs are expressed and light independent Sortase A reactions are completed. Plans for testing light dependence and reversibility have been made for future study. In Chapter 4, conclusions are made from both light dependent conjugation systems. Experiments to improve the system are discussed as well as unique application considerations.

Chapter 2

LOV IS BLISS: CREATING A BLUE LIGHT INDUCED SPYCATCHER SYSTEM

2.1 Introduction

As previously discussed, the SpyCatcher:SpyTag peptide tagging system has been used extensively for various applications. In this chapter, the ability to control this protein conjugation system with light is tested via incorporation of the light-oxygen-voltage domain 2 (LOV2) previously discussed. By appending either partner of the SpyCatcher:SpyTag system to the C-terminus of the LOV2 domain, under dark conditions, the LOV2 domain was hypothesized to be able to sterically block the appended protein from reacting from its native partner. Once irradiated with blue light, the LOV2 protein would undergo a conformational change which would allow for the undocking of the appended protein. The undocking would rid the system of steric hindrance and allow for the reaction between the two reactive partners proceed. Figure X below shows an illustration of the blue light dependent docking and undocking for each design.

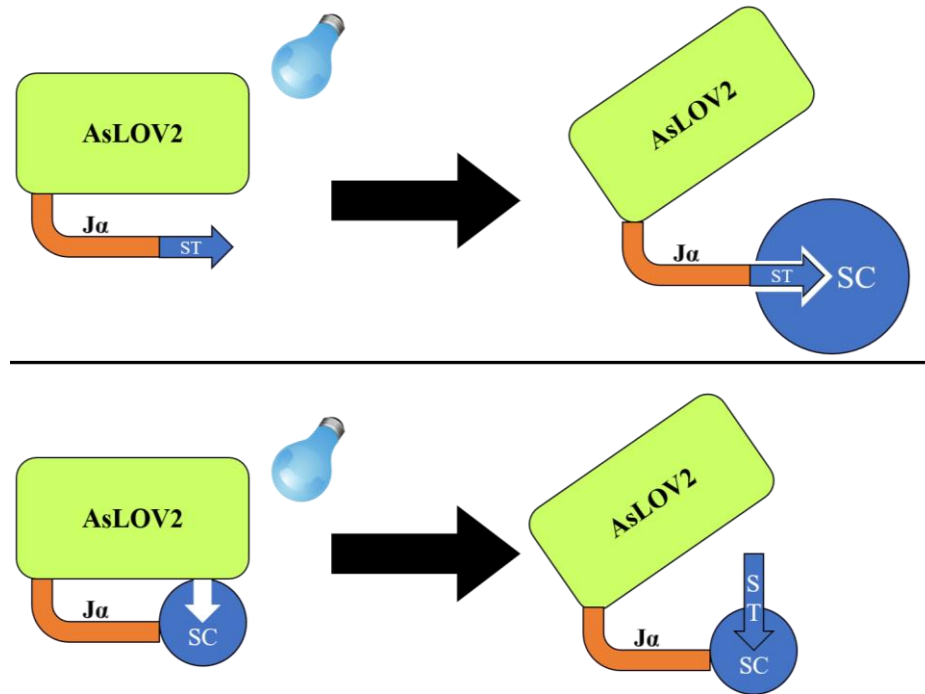


Figure 5: Illustration of discussed constructs. In the top image, SpyTag is genetically fused to the $J\alpha$ -helix and in the bottom image SpyCatcher is fused. Once irradiated by blue light, a conformational change should occur within the $J\alpha$ -helix which allows undocking, allowing the reaction to proceed.

Using appropriate molecular cloning methods, six constructs were made with the LOV2 protein to test their ability to control the SpyCatcher:SpyTag irreversible reaction. Either SpyCatcher or SpyTag was appended to one of three truncation points at the C-terminus of the LOV2 domain. Constructs were expressed in *E.coli* and purified using Ni-NTA His-tag chromatography. After purification, blue light dependent experiments showed that the LOV2 domain was only able to control the reaction when appended to the smaller of the two partners, SpyTag.

2.2 Methods & Materials

2.2.1 Molecular Cloning

2.2.1.1 Obtaining Template Plasmid & PCR Design

An engineered version of the LOV2 protein had been made available by the Kuhlman group named iLID (improved light inducible dimer). The plasmid for this construct was purchased on Addgene in a pQE-80L 5.25kb vector backbone with Ampicillin resistance. The gene of interest was flanked by BamH1 & HindIII restriction sites with a N-terminal histidine tag. pQE-80L iLID (C530M) was a gift from Brian Kuhlman (Addgene plasmid # 60408 ; <http://n2t.net/addgene:60408> ; RRID:Addgene_60408).²²

For creation of the LOV2-SpyTag constructs, PCR primers were designed to directly incorporate the SpyTag sequence at three different truncations of the LOV2 protein. A forward primer contained the already present BamHI cut site and reverse primers incorporated SalI sites in the middle of the SpyTag sequence. Adapters were also designed to insert the rest of the SpyTag and link HindIII, which was conserved in the vector backbone.

For the creation of LOV2-SpyCatcher, Gibson Assembly was utilized. Primers for this process were made to first amplify the SpyCatcher construct and the LOV2 protein separately with overlapping regions for Gibson Assembly. PCR primers for SpyCatcher used a gBlock (Integrated DNA Technologies) that contained the sequence for SpyCatcher. All primers and adapters were ordered from Integrated DNA Technologies.

2.2.1.2 Polymerase Chain Reaction (PCR)

Polymerase chain reactions were conducted via a BioRad Thermocycler. Samples of 25 μ L were prepared with about 0.5ng of template DNA and Phusion High-Fidelity DNA Polymerase purchased from New England BioLabs. The mixture was supplemented with 5x Phusion Buffer, and 1.25 μ L of 10 μ M in forward and reverse primers, and .5 μ L of 10 μ M deoxy nucleotide triphosphates (dNTPs).

Samples were cycled through denaturing, annealing, and elongation steps to amplify a gene of interest. For each reaction an annealing temperature was determined to be +3 of the lowest primer melting temperature and the elongation time to be 30s/1kb of expected product size.

2.2.1.3 Restriction Enzyme Digest & Agarose Gel Electrophoresis

Enzyme digests were set up using buffers and enzymes purchased from New England BioLabs. 50 μ L samples were prepared with approximately 2 μ g of DNA and supplemented with 10x CutSmart buffer and 1 μ L of each enzyme needed. Samples were allowed to react for >4hrs, and usually overnight, in a 37 $^{\circ}$ C water bath.

Agarose gels were used to characterize and confirm steps in the cloning process. Agarose gels were made with a concentration of agarose ranging from 0.5-2% (wt/v) depending on the resolution and supplemented with 1% (v/v) Ethidium Bromide. Agarose would be dissolved in TAE (Tris-Acetate-EDTA) buffer and heated using a microwave. After ethidium bromide was added to the hot solution, the gel was poured in a cast to solidify. The agarose utilized was purchased from Bio-Rad laboratories.

Gels were loaded with 5 μ L of either 2-log or 1kb ladder (New England BioLabs) and 5-25 μ L of DNA samples. DNA samples were composed of DNA and 6x gel loading dye (New England BioLabs).

Gels were run at a constant voltage of 90V for 30 minutes. Gels were then imaged on a UV plate. If digested vector backbone was run, DNA band would be cut out and purified out of the agarose using a gel extraction kit (Zymo).

In preparation of ligations and enzyme digests, Zymoclean™ Gel DNA Recovery Kit was used to purify DNA purchased from Zymo Research. In preparation of ligations, the kit was used to purify the digested vector DNA cut from the agarose gel. The protocol provided by the supplier was followed. DNA concentration was determined via wavelength absorbance at 260nm. Absorbance measurements were taken by a nanodrop spectrophotometer (Thermo Fisher).

2.2.1.4 DNA Ligation

In preparation for the ligation of LOV-Tag constructs, adapters were phosphorylated at the 5' end in preparation for ligation. T4 Polynucleotide Kinase (T4PNK) purchased from New England BioLabs was used. Three-piece ligations were set up in 20 μ L samples with a molar ratio of 10:10:1 (insert:adapter:vector). T4 Ligase and 10x T4 Ligase buffer were utilized and purchased from New England BioLabs. Reactions were allowed to run at room temperature for 10-20 minutes.

Gibson Assembly was utilized to ligate together the amplified LOV2 domain and the amplified SpyCatcher domain. 10 μ L samples containing 0.15 μ mol of each construct and 5 μ L of Gibson Master Mix, purchased from New England BioLabs were incubated at 50°C for about 60 minutes. The resulting products were then amplified using the PCR methods using the LOV2 Forward and SpyCatcher Reverse primers.

Two-piece ligations were set up in 20 μ L samples with a molar ratio of 3:1 (insert:vector). T4 Ligase and 10x T4 Ligase buffer were utilized and purchased from New England BioLabs. Reactions were allowed to run at room temperature for 10-20 minutes.

2.2.1.5 Bacterial Transformation for Plasmid Preparation

DNA was transformed into competent cell lines and plated on antibiotic (100 μ g/mL ampicillin) selective plates via heat shock. For ligated products, 50 μ L of competent cells were incubated with 5 μ L of ligated product for 30 minutes on ice. This solution was then heated at 42 $^{\circ}$ C for 45 seconds and returned to ice for 5 minutes. After 5 minutes, 950 μ L of LB media (1% tryptone, .5% yeast extract, 1% NaCl) was added to solution and the mixture was incubated in a 37 $^{\circ}$ C shaker for an hour. After incubation, the solution was centrifuged for 3 minutes at 3000xg. About half of the supernatant was removed and the cellular pellet was resuspended in the remaining media. 50 μ L of this solution was then plated onto an appropriate antibiotic (ampicillin) resistant plate and grown overnight at 37 $^{\circ}$ C in a plate incubator (Fisher Scientific). The competent cell line used for ligated products was NEB 5-alpha *E.coli.*, purchased from New England BioLabs. This cell line is a derivative of the DH5 α cell line that is used for high-quality plasmid preparation.

2.2.1.6 Cloning Confirmation and Plasmid Extraction

Bacterial colonies grown in NEB-5 alpha were picked and 5mL cultures were grown overnight in LB media and incubated in a 37 $^{\circ}$ C shaker. Plasmids were extracted from the bacterial culture using ZyppyTM Plasmid Miniprep Kit purchased

from Zymo Research. The kit was used according to the manufacture's instructions. Concentrations would be confirmed via a nanodrop spectrophotometer.

2.2.2 Protein Expression & Purification

2.2.2.1 Bacterial Transformation for Protein Expression

The protocol for bacterial transformation for the purpose of protein expression follows the same steps as seen in section 2.2.1.13 with a few deviations. The competent cell line used in BL21(DE3)pLysS provided by New England BioLabs. This cell line is commonly used protein production. In addition, because genomic, unaltered DNA is being transformed only 1 μ L is needed to be incubated with the cells and 50 μ L cells are plated without the spinning down and concentrating of the cellular pellet. Plates were grown overnight in a 37°C plate incubator (Fisher Scientific).

2.2.2.2 Protein Expression

.Because all constructs made contained the same vector backbone, expression conditions were the same. Protein expression was done primarily in 125mL Erlenmeyer flasks on 2xTY media (1.6% Tryptone, 1% Yeast Extract, 0.5% NaCl) supplemented with 100 μ g/mL ampicillin and 25 μ g/mL chloramphenicol for plasmid selection. Precultures of 2mL were started in the morning and allowed to grow at 37°C, 250 rpm to a high enough concentration to inoculate a culture of 25mL to an OD₆₀₀ of 0.1. Cells were grown at 37°C until they reached an OD₆₀₀ of 0.6-1.0. Cellular density was measured using a spectrophotometer (Biowave) and taking an optical density measurement at 600nm (OD₆₀₀). Once this concentration was reached, the culture was induced with 333 μ M of IPTG and moved to a 20°C, 250 rpm shaker (New Brunswick Scientific) for overnight growth.

Cells were harvested in the morning by centrifuge (Thermo Fisher). After the final OD₆₀₀ of the cellular broth was recorded, the suspension was transferred from the 125mL Erlenmeyer flasks to prechilled 50mL centrifuge tubes on ice. Cellular suspensions were spun at 3000xg for 10 minutes at 4°C. The supernatants were discarded, and the cellular pellet was resolubilized in cold TN150 buffer (50mM Tris, 150mM NaCl, pH 8.0) to an OD of 20 or 50.

Cell membranes were disrupted using sonication methods. Suspended in TN150, cells were subjected to 3 second pulses with an amplitude of 35 or 45, dependent on the resolubilized concentration, followed by 9 seconds of rest. As a general rule of thumb, the amount of time (minutes) of sonication was equal to the resuspended volume. The sonicated product was then centrifuged at max speed for 20 minutes at 4°C. The soluble fraction of the lysate was removed for further analysis and the insoluble, pelleted, fraction was resuspended in TN150 buffer.

2.2.2.3 Protein Characterization

A Bradford Assay was completed on soluble lysate fractions to determine the concentration of protein in solution. Bradford standards ranging from 0.05mg/mL to 0.5 mg/mL of Bovine Serum Albumin (BSA) were used to create calibration curves. 10µL of standard and diluted (usually 1:20) protein samples were added to a 96 well plate. 200µL of 1x Bradford dye, purchased from Bio Rad Laboratories, was added simultaneously to each sample using a multi-channel pipette. The plate was kept at room temperature for 5—10 minutes and samples were mixed with a plate shaker. The 96-well plate was then measured for absorbance at 595nm by a plate reader. Calibration curves were made to correlate absorbance values and protein concentrations.

Sodium Dodecyl Sulfate (SDS) Polyacrylamide Gel Electrophoresis (PAGE) was used to characterize the proteins by size. SDS gels are made of two sections, a resolving section and a stacking section. The resolving section differed in SDS concentration depending on the desired resolution. Gels were composed of acrylamide, SDS, Tris base, Ammonium Persulfate (APS), and Tetramethylethylenediamine (TEMED). SDS and Tris base were purchased from Fisher Scientific. Other reagents were purchased from Bio-Rad laboratories.

SDS samples of 80 μ L were composed of 40 μ g protein and 20 μ l of 4x SDS reducing dye. The samples were boiled for about 5 minutes before loading. Gels were loaded with 5 μ L of Precision Plus Protein™ Unstained Standard, purchased from Bio-Rad laboratories. 15 μ L of each boiled SDS protein sample was loaded in the gel. The gel was ran at a constant voltage of 125V for about 100 minutes, or until the samples reached the bottom of the gel. The gels were run in a running buffer containing 25mM Tris, 3.5mM SDS, and 200mM Glycine.

After running the gel, the stacking section was discarded, and the resolving section was stained overnight in Coomassie Brilliant Blue stain (0.1% Coomassie Brilliant Blue R-250, 50% methanol, 10% acetic acid). After staining, multiple rounds of destain (40% methanol, 10% glacial acetic acid) were used to visualize the protein bands. Gels were then imaged on a white tray (BioRad).

2.2.2.4 Protein Purification

Proteins were purified to remove cellular contaminants. In this chapter, all constructs were cloned with 6 N-terminal histidine residues (His-Tag) for Ni-NTA column purification. Tris buffer (50mM Tris, 150mM NaCl, pH 8.0) with varying concentrations of imidazole were prepared for binding (10mM), washing (50mM), and

eluting (250mM). Columns were purchased from BioRad laboratories and column resin were purchased from Thermo-Fisher.

Briefly, gravity columns are initially charged with Ni-NTA slurry. The bed was allowed to pack and was then washed with three column volumes of water, followed by three column volumes of binding buffer. The protein samples being purified were brought to the same imidazole concentration as the binding buffer. The protein samples are then added to the column. Flow through samples are taken. After protein loading, 10 column volumes of binding buffer are added, followed by 6 volumes of washing buffer. The POI (protein of interest) is then eluted with 6 column volumes of elution buffer.

2.2.3 Light Sensitive Experimentation

Light dependent experiments were conducted in the dark behind a blackout curtain. 500 μ L samples of equimolar (10mM) SpyCatcher and SpyTag constructs were prepared inside the curtain and kept at room temperature. After sample preparation, volumes would be split, and half of the samples would be subjected to blue light irradiation at 0.5mW/cm². The blue light source was a lab made blue LED 470 nm. “Light” were kept under a cardboard box, behind the curtain, with the constant light exposure, while “Dark” samples were behind the curtain covered in tin foil. At time points, protein samples would be added to SDS to quench the reaction.

2.3 Results & Discussion

2.3.1 Generation of BLISS Library

.Three truncations of each LOV-SpyCatcher and LOV-SpyTag were successfully prepared using the methods described in chapter 2.2.1; confirmation of

the successful truncations was determined via restriction enzyme digest. Digests were run on 2% agarose gel for visualizations. Figures 6 & 7 below show examples of positive clone for each LOV construct. Once all constructs were confirmed, the DNA was transformed into the production cell line pLysS for protein expression.

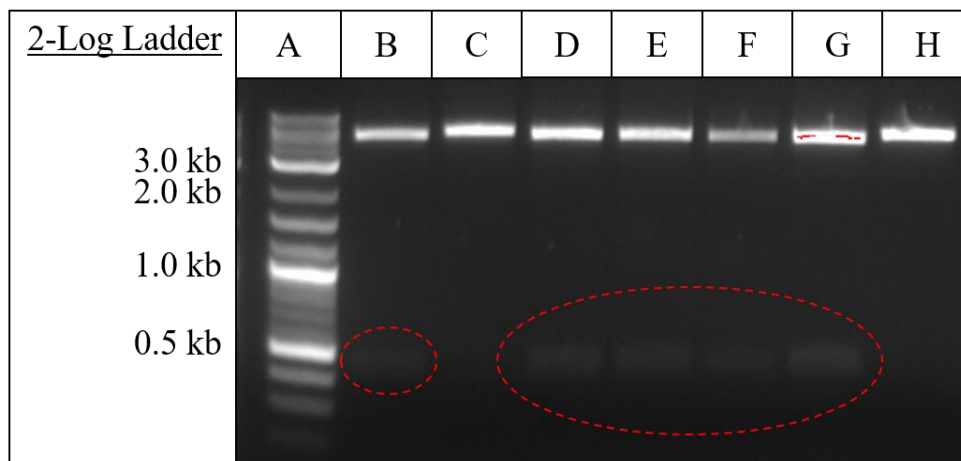


Figure 6: 2% agarose gel of digested LOV-540-SpyTag constructs. Constructs were digested with HindIII & BamHI. Lane Profile: (A) 2 log ladder, (B-H) Ligated clones of LOV-540-SpyTag. Positive clones with a band at .465 kb can be seen in lanes B, D, E, F, & G.

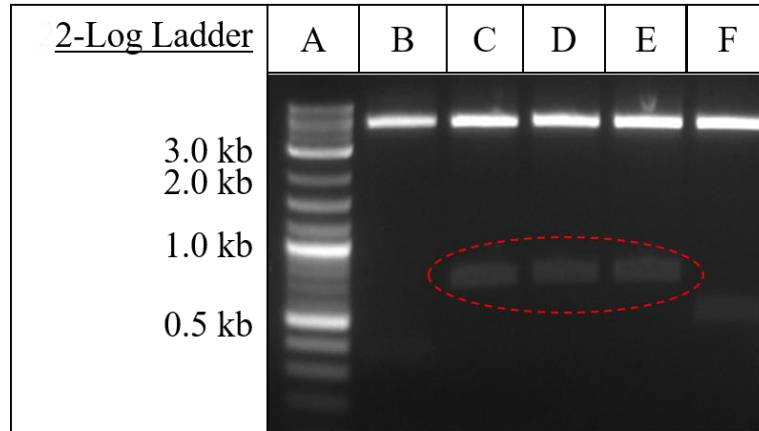


Figure 7: 2% agarose gel of digested LOV-541-SpyCatcher constructs and iLID (cloning template). Constructs were digested with HindIII & BamHI. Lane Profile: (A) 2 log Ladder, (B-E) LOV-541-SpyCatcher ligated clones, (F) iLID control. It is observed that the positive clones in this screening are in lanes C, D, & E with a band corresponding to .675kb. iLID control corresponds to correct band at .456kb.

2.3.2 Protein Expression & Purification

At first, conditions from literature²³ were used to express and purify the original iLID structure. These experiments were conducted, and the results of expression and purification are below in Figure 8 below. In Figure 8, a protein band at the appropriate size to represent the iLID structure (18.58 kDa) is observed in the soluble lysate and showed highly purified in the elution lanes.

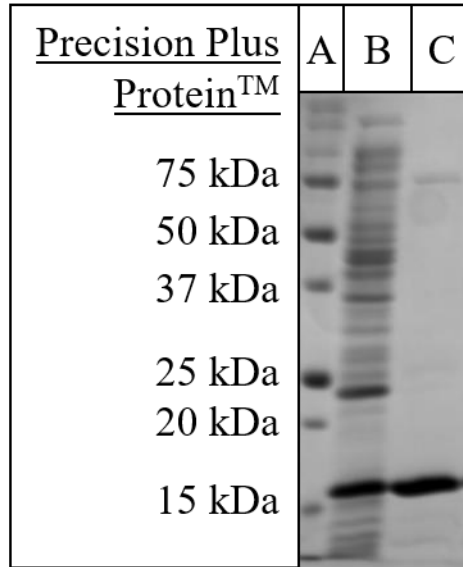


Figure 8: 12.5% SDS PAGE of results iLID expression and purification via conditions reported in literature. Lane profile: (A)Ladder, (B)Soluble Lysate, (C) Elution Sample.

Once the expression conditions reported were confirmed to produce high levels of soluble, easily purified protein, expression of the genetically engineered constructs began. The results can be seen in Figures 9 & 10 below with positive size bands for LOV-X-Tag (18.9 kDa) & LOV-X-Catcher (26.5 kDa), respectively. Additional SDS PAGE gels can be found in the Appendix.

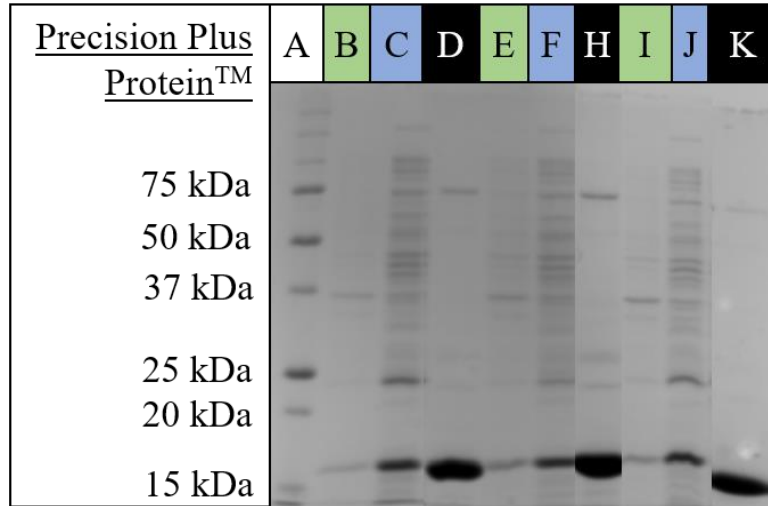


Figure 9: 12.5% SDS PAGE of expression and purification results of LOV-X-SpyTag constructs. Lane profile: (A) Ladder, (B-D) LOV-537-SpyTag, (E-H) LOV-540-SpyTag, (I-K) LOV-541-SpyTag. Color key: Green (Insoluble Lysate), Blue (Soluble Lysate), Black (Elution).

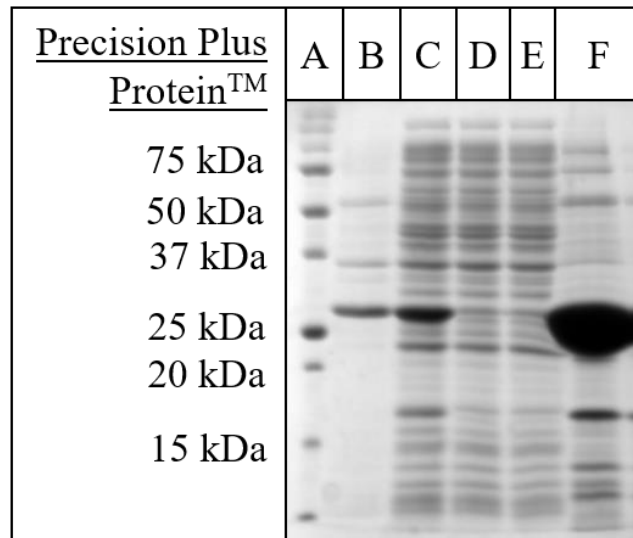


Figure 10: 12.5% SDS PAGE of expression and purification results of LOV-541-SpyCatcher constructs. Lane profile: (A) Ladder, (B) Insoluble Lysate (C) Soluble Lysate, (D) Flow Through 1, (E) Flow Through 2, (F) Elution.

2.3.3 Blue Light Dependent Experiments

Once all cloned constructs were expressed and characterized, light dependent conjugation properties of SpyCatcher:SpyTag system, were screened. For the initial set of experiments, time points were taken at the start and end (30 minutes) of the experiments. Reactions were quenched by SDS as stated in methods. Purified SpyTag-ELP & Z-ELP-SpyCatcher available from students in the lab were used as reaction partners. Elastin like polypeptide (ELP) are useful amino acid tags used for to easily purify recombinant proteins. Table 1 below shows a summary of expected reaction size bands is.

Table 1: Expected Size of SpyTag:SpyCatcher Reactions.

Construct	Expected Size
SpyTag-ELP	22 kDa
Z-ELP-SpyCatcher	35 kDa
iLID (AsLOV2)	18.5 kDa
LOV-SpyTag	18.9 kDa
LOV-SpyCatcher	26.5 kDa
LOV-SpyTag + SpyCatcher	53.5 kDa
LOV-SpyCatcher + SpyTag	48.5 kDa

A summary of the results from initial experiments can be seen below in Figures 11 & 12. 12.5% SDS gels of the control experiments for both constructs can be seen in the Appendix.

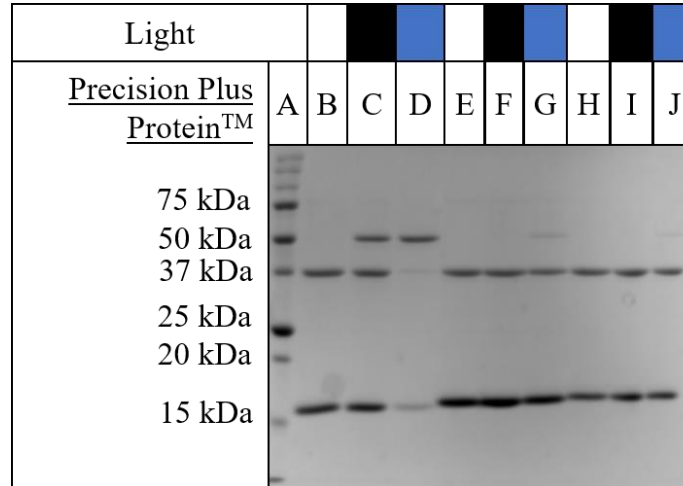


Figure 11: 12.5% SDS PAGE of LOV-X-SpyTag + SpyCatcher light dependent conjugation experiment. Lane Profile: (A) Standard, (B-D) LOV-537-SpyTag, (E-G) LOV-540-SpyTag , (H-J) LOV-541-SpyTag. Color Key: White (Time 0), Black (30 minutes in dark), Blue (30 minutes in blue light).

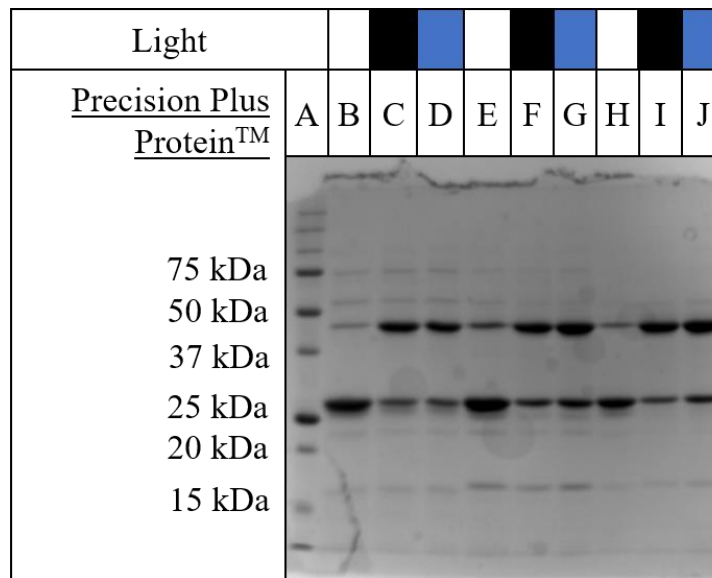


Figure 12: 12.5% SDS PAGE of LOV-X-SpyCatcher + SpyTag light dependent conjugation experiment. Lane Profile: (A) Standard, (B-D) LOV-537-SpyCatcher, (E-G) LOV-540-Spy-Catcher , (H-J) LOV-541-SpyCatcher. Color Key: White (Time 0), Black (30 minutes in dark), Blue (30 minutes in blue light).

. Following the completion of initial experiments, it was determined that the only structures that were able to successfully control the reaction, at least partially, were the LOV-Tag structure. After 30 minutes, it was observed that for the construct that had SpyTag appended deepest into the $J\alpha$ helix (at the 537 location), a majority of the free SpyCatcher reacted with the LOV-SpyTag construct when irradiated in blue light. A substantial noticeable amount of background reaction was observed for this construct as well. Later experiments showed that although background is present, it is not as apparent as seen in the above figure. This was determined to be because errors in mixing and pipetting during initial experimentation. The other two LOV-Tag constructs showed a much smaller amount of conjugation under irradiated conditions but practically no amount of background.

When analyzing conjugation results of the LOV-SpyCatcher constructs, it is observed that the reaction occurred at the same rate regardless of whether the samples were irradiated in blue light. This is most likely because the SpyCatcher was too large for the LOV protein to be able to sterically hinder SpyCatcher's active site.

Further experiments were conducted with the collection of more frequent and longer time points on the LOV-SpyTag constructs to better determine the kinetics of the reaction. Results of these experiments can be seen below in Figure 13 below and the Appendix.

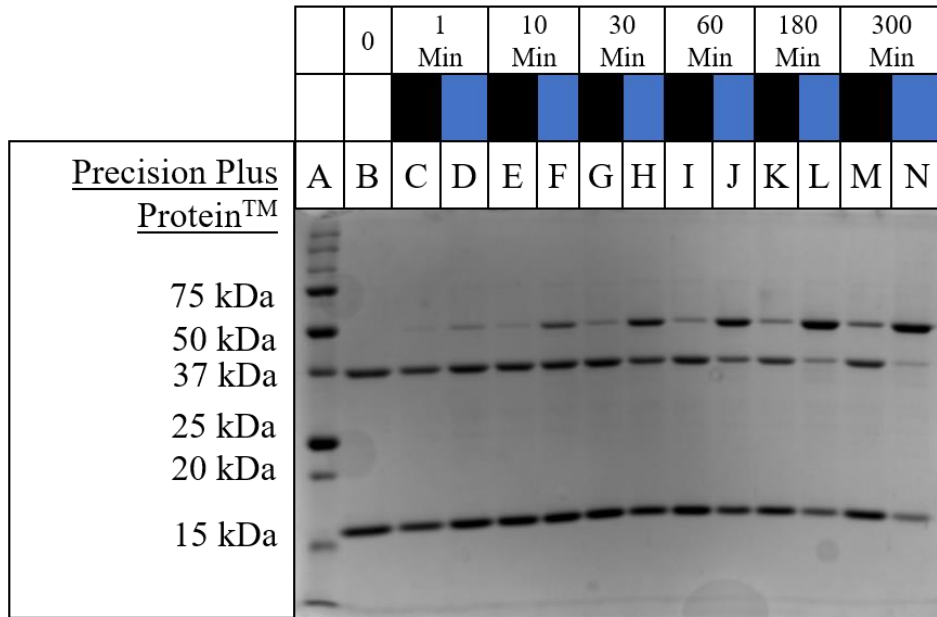


Figure 13: 12.5% SDS PAGE of LOV-537-SpyTag + SpyCatcher light dependent conjugation experiment. Lane Profile: (A) Standard, (B-N) LOV-537 + SC reaction. Color Key: White (Time 0), Black (No Blue Light), Blue (Constant Blue Light).

By analyzing the bands with densitometry, estimates of ligation yield over time could be estimated. Figure 14 below shows how the much of SpyCatcher:SpyTag system was formed for the light and dark samples. Reconstitution percentages were determined by averaging the relative disappearance of the unconjugated partners. It was observed that after 5 hours about 70% of reaction had occurred in the irradiated samples and 17% in the dark samples.

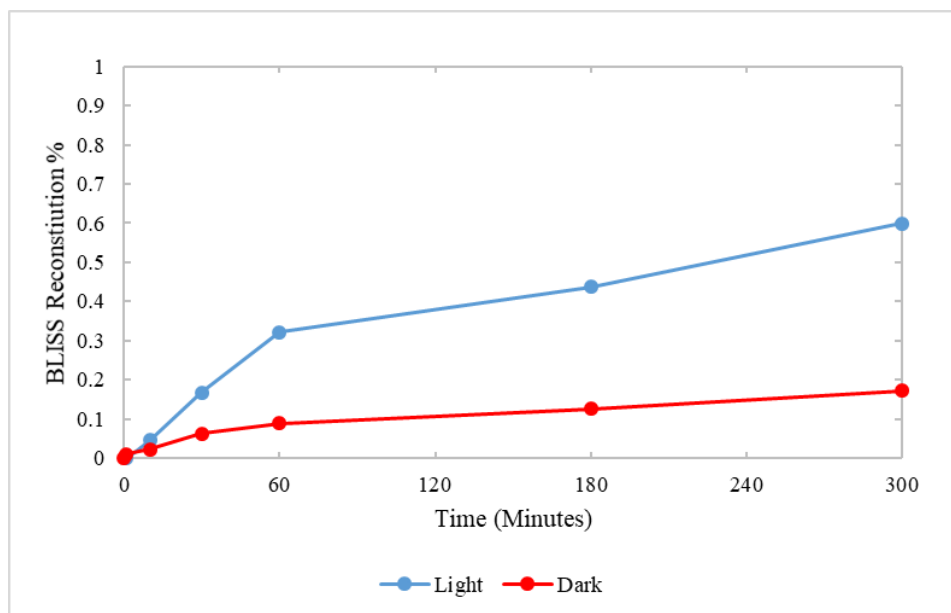


Figure 14: Reconstitution of LOV-537-SpyTag + SpyCatcher complex. Reconstitution was determined by averaging the amount reacted of the LOV-537-SpyTag & Spy-Catcher constructs. Time points were taken at 1 minute, 10 minutes, 30 minutes, 1 hour, 3, hours, 5 hours.

2.4 Summary and Conclusions

In summary, constructs were cloned to test the ability of the AsLOV2 domain to control the SpyCatcher:SpyTag conjugation system. The LOV domain also showed an inability to hinder SpyCatcher's active site in the dark when genetically fused together. However, by appending the SpyTag peptide of the SpyCatcher:SpyTag conjugation system to the C-terminus of the α helix of the LOV domain, blue light dependent protein conjugation was achieved in one engineered construct. Of the three truncations tested, the inner most showed the greatest reactivity. Although the other two constructs showed no background conjugation, an insufficient, amount of conjugation occurred in the lit state. When SpyTag was inserted at the 537th residue of the LOV domain, it was observed that protein conjugation increased over 3-fold

between lit and dark states. This marks a novel use of the LOV2 protein to activate conjugation of the SpyTag:SpyCatcher complex.

Further studies completed by research group members showed that system responded well to pulsating light, turning on and off quickly. Experiments where wavelengths and intensities are varied would benefit our understanding of the system. SpyTag was also fused into other locations in the J α and tested for ability to dynamically cage and uncage. Results showed the same inability to uncage efficiently. Further analysis of the interactions between the core of the LOV2 domain and the J α helix would be needed to elucidate why the 537 truncation is the only one that allows for substantial undocking.

The LOV-SpyCatcher construct may also be revisited and redesigned to improve dark state caging. This could be achieved by changing the orientation of SpyCatcher via a peptide linker. It may also be interesting to consider the addition of a peptide (like SpyTag) to the N-terminus of the LOV domain. This peptide would have an affinity to the SpyCatcher active site but not result in an irreversible bond. The ability to block and unblock an entire protein, and not just a small peptide, would have much broader and more versatile application.

Chapter 3

SPATIAL AND TEMPORAL CONTROL OF REVERSIBLE SORTASE A TRANSPEPTIDATION WITH THE LOV DOMAIN

3.1 Introduction

Chapter 2 provided evidence that the AsLOV2 domain could be engineered to light dependently induce conjugation of the SpyCatcher:SpyTag complex. This was done by genetically fusing the smaller of conjugation pair, SpyTag, to the C-terminus of the AsLOV2 protein. The covalent conjugation of SpyCatcher:SpyTag provides stable binding to a LOV2 based photoswitch; however, the irreversibility of the reaction limits the application of the system. It became a point of interest to develop a reversible covalent conjugation system. The Sortase A transpeptidation system was introduced to provide the system a reversible aspect. Being able to use light to interchangeably conjugate proteins could be a very useful tool.

In this chapter, steps are taken to clone the Sortase A C-terminal recognition site into the $J\alpha$ helix of the LOV domain at the same three previously identified sites. Figure 15 below shows an illustration of the constructs engineered and their intended functionality in proposed experiments. The Sortase A recognition site should become available to Sortase A when the LOV constructs are irradiated with blue light. When this occurs Sortase A should be able to facilitate the transpeptidation reaction. After cloning was completed, proteins were expressed and purified. Work was done to optimize the conditions and concentrations of the Sortase A mediated transpeptidation. Experiments were planned to test light reversibility and expected results are discussed.

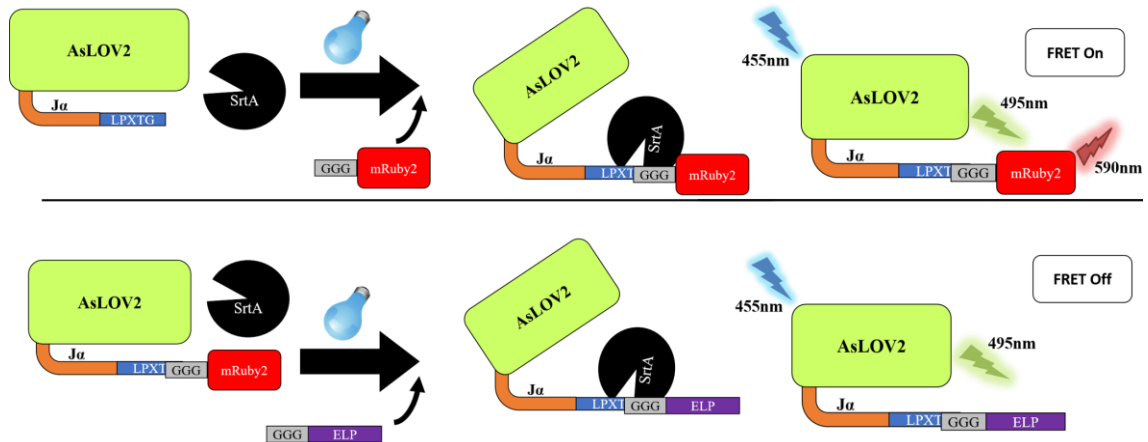


Figure 15: Proposed schematic of how Sortase transpeptidation system would be controlled with blue light irradiation. When LOV-LPXTG undergoes blue light irradiation, the $J\alpha$ (orange) should become undocked. When in solution with SrtA (black) and an oligoglycine substrate (GGG-mRuby2), transpeptidation should occur. To test reversibility, if light is removed, and then reapplied with the addition of another substrate to solution, the substrates should become interchangeable. Reversibility would be tested with FRET. (Right, Upper) A FRET response would be expected with an mRuby2 conjugation. (Right, Lower) FRET response would weaken with the addition of a nonfluorescent substrate as it replaces mRuby2.

3.2 Methods & Materials

3.2.1 Molecular Cloning

3.2.1.1 PCR Design

Primers were designed to incorporate the Sortase A recognition sequence (LPXTGGSG) into the C-terminus of the LOV2 protein. In addition, primers were designed to flank the PCR product with BamHI and HindIII restriction enzyme cut sites for cloning. The original iLID structure, as discussed in 2.2.1.1, was utilized as the PCR template and the ligation vector.

To obtain oligoglycine substrates, a GGG-ELP-His6 vector, was used to incorporate genes of interest. Primers were designed to amplify mRuby2 and mCherry with appropriate enzyme digestion sites. These inserts were designed to replace the DNA coding for ELP in the ligation vector

3.2.1.2 Polymerase Chain Reaction (PCR)

See section 2.2.1.2.

3.2.1.3 Restriction Enzyme Digest & Agarose Gel Electrophoresis

See section 2.2.1.3.

3.2.1.4 DNA Ligation

See section 2.2.1.4.

3.2.1.5 Bacterial Transformation for Plasmid Preparation

See section 2.2.1.5.

3.2.1.6 Cloning Confirmation & Plasmid Extraction

In addition to the methods discussed in section 2.2.1.6. DNA sequencing was utilized to confirm successful integration of the Sortase A recognition sequence. 10 μ L samples of prepared plasmid were sent to Eurofins Genomics to be sequenced.

3.2.2 Protein Expression & Purification

3.2.2.1 Bacterial Transformation for Protein Expression

See Section 2.2.2.1.

3.2.2.2 Protein Expression

Expression of LOV-LPXTGGSG constructs was identical to Methods discussed in 2.2.2.2.

Expression of Sortase A was conducted in LB media (1% tryptone, .5% yeast extract, 1% NaCl), supplemented with 50 µg/mL kanamycin. 5mL overnight precultures were prepared and grown in a 37°C incubated shaker (New Brunswick Scientific). Precultures were used to inoculate a 20mL morning culture to an OD of 0.1. Cells were allowed to grow in a 37°C, 250 rpm shaker (New Brunswick Scientific). The culture was induced with 1mM IPTG at an OD of 0.7 – 1.0. Expression was allowed to continue for 4 hours at 37°C, 250 rpm, until harvest.

Expression of GGG-mRuby2, GGG-mCherry, GGG-ELP, & ELP-LPETG was done in TB media (2% tryptone, 2.4% yeast extract, .4% Glycerol, 10% TB Salts [0.17M KH₂PO₄, .72M K₂HPO₄]) supplemented with 50 µg/mL kanamycin. 2mL morning precultures were grown in 37°C shaker to inoculate 20mL (GGG-ELP, ELP-LPETG) and 25mL (GGG-mRuby2, GGG-mCherry), to an OD 0.1. The 20mL cultures were allowed to grow under leaky conditions (no IPTG induction) overnight at 37°C, 250 rpm. The 25mL cultures were induced with 200µM IPTG once an OD of 0.7 – 1.0 was reached. After induction, cultures grew overnight in a 20°C, 250 rpm incubated shaker (New Brunswick Scientific).

The expression conditions for Z-Ct-LPETG were similar to that of the oligoglycine substrates. Using LB media supplemented with 1.5% glycerol and 100 µg/mL ampicillin, morning precultures were grown in a 37°C, 250 rpm incubated shaker (New Brunswick Scientific) and used to inoculate a 25mL culture to an OD of 0.1. The culture was then induced with 200µM IPTG once it reached an OD of 0.7 –

1.0. The culture was then allowed to grown overnight in a 25°C (New Brunswick Scientific), 250rpm incubated shaker.

3.2.2.3 Protein Characterization

See Sections 2.2.2.3.

In addition, concentrations of ELP-containing peptides were further characterized due to their poor staining with Coomassie Blue. The nanodrop spectrophotometer (Thermo Fisher) was utilized to take the absorbance of the solution at 280nm. In combination of the theoretical estimated extinction coefficient and molecular weight, concentrations were obtained.

3.2.2.4 Protein Purification

LOV constructs and oligoglycine substrates containing a terminal His Tag were purified using the same methods as discussed in section 2.2.2.4.

Proteins tagged with ELP (GGG-ELP & ELP-LPETG) were purified via Inverse Transition Cycling (ITC). Elastin like polypeptide (ELP) purification is based on the principle that at elevated salt concentrations and temperatures, ELP tagged proteins become reversibly insoluble. With addition of ammonium sulfate to a final concentration of 0.5 M and heating the solution to 37°C, precipitate formed, and the samples were centrifuged. Supernatants were removed and the insoluble pellet was resolubilized in cold buffer. This cycle would be done two times to a sample, concentrating by a factor of 2 each time.

The thermally stable protein, Z-Ct-LPETG, was purified using heat. Samples of soluble lysate were heated to 70°C in a water bath for about 10 minutes. Samples

were then centrifuged at max speed for 10 minutes. Because the protein is thermally stable, it remains in the supernatant, and the pellet is discarded.

3.2.3 Functionality Experiments

3.2.3.1 Blue Light Independent Sortase Experiments

Sortase A reactions with the cloned oligoglycine substrates were conducted at a variety of described compositions. Reactions were prepared in 100 μ L volumes and incubated in a 37°C water bath. Negative controls were created by forgoing the addition of 6mM CaCl₂ to the reaction, as calcium is required for Sortase A transpeptidation. Reactions were quenched with the addition of SDS and samples were run on SDS-PAGE.

3.3 Results & Discussion

3.3.1 Generation of Constructs for Blue Light Induced Sortase A Transpeptidation

Three truncations of the LOV-LPXTGGSG construct were created using the cloning methods described in chapter 3.2.1. Confirmation of these constructs was done via DNA sequencing. An example of DNA sequencing results can be found in the Appendix.

In addition, two red fluorescent (RFP) oligoglycine substrates were generated, using the same methods. GGG-mCherry & GGG-mRuby2 were confirmed with restriction enzyme digest. Figure 16 shows the confirmation of these constructs with expected band sizes of about 0.725kb for the RFP's compared to the unaltered GGG-ELP insert size of 0.648kb.

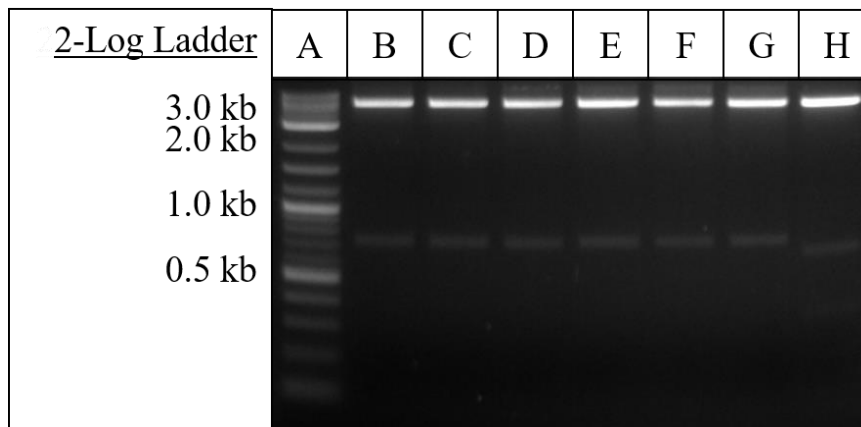


Figure 16: 2% agarose gel of digested GGG-mCherry, GGG-mRuby2, & GGG-ELP constructs. Constructs were digested with EcoRI & XhoI. Lane Profile: (A) 2 log ladder, (B-D) Ligated GGG-mCherry clones, (E-G) Ligated mRuby2 clones, (H) GGG-ELP. Positive clones are observed for all constructs screened.

3.3.2 Protein Expression & Purification

In addition to the constructs cloned, other constructs already available were necessary for the planned experiments. Sortase A was required to be expressed to facilitate the protein conjugation. Positive control constructs that contain the Sortase A recognition sequence that weren't hindered by the LOV protein, ELP-LPETG and Z-Ct-LPETG, were expressed. Additionally, a non-fluorescent oligoglycine tag, GGG-ELP, was produced. Expected sizes can be seen below in Table 2.

Table 2: Expected Size bands of Engineered Proteins

Construct	Expected Size
GGG-mCherry	28.6 kDa
GGG-mRuby2	28.6 kDa
GGG-ELP	18.4 kDa
ELP-LPETG	20.4 kDa
SrtA	20 kDa
LOV-X-LPXTGGSG	18.16 kDa

Initially all structures beside Z-Ct-LPETG were expressed and purified, as discussed in the methods. All proteins showed positive results with expected protein sizes showing up in each of the respective soluble lysates. As expected, Sortase A was highly expressed and did not require any further purification of the cellular lysate. The remaining constructs were each purified based on the methods defined, dependent on their respective purification tags (ELP or His6). Figure 17 below displays the expression and purification results.

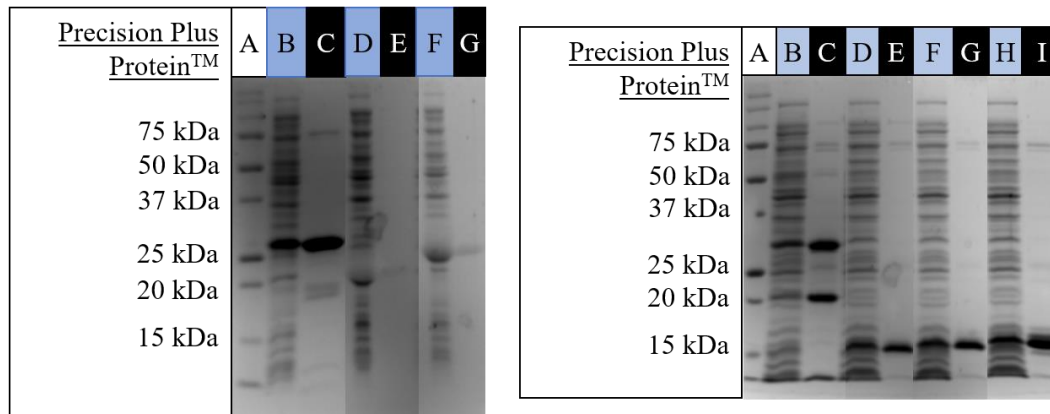


Figure 17: 12.5% SDS PAGE of expression & purification results. Color key: Blue (Soluble Lysate), Black (Elution/Purified). Lane Profile (Left): (A) Ladder, (B-C) GGG-mRuby2, (D-E) GGG-ELP, (F-G) ELP-LPETG. Lane Profile (Right): (A) Ladder, (B-C) GGG-mCherry, (D-E) LOV-537-LPITG, (F-G) LOV-540-LPETG, (H-I) LOV-541-LPETG.

Following analysis of the above gels, it was determined that the GGG-mCherry structure would be abandoned, as it was observed to have a greater fraction of truncated product compared to mRuby2. Because these two proteins fluoresce at approximately the same wavelength, mRuby2 was determined to be the more reliable construct of the two.^{23,24} The purified lanes of the ELP tagged protein show faint bands that, although at the expected size, look faint and diluted. This was determined to be because the ELP tag has a low percentage of basic residues which bind to Coomassie dye. Although not very visible, protein concentrations were confirmed via nanodrop.

Because the ELP-LPETG was especially hard to see on the gel, thermophilic Z-Ct-LPETG was chosen to be expressed as a more visual protein for initial Sortase A optimization. Purification was done with heat as stated in the methods. The results can be seen in the Appendix.

3.3.3 Optimization of Sortase Reaction

Using densitometry, fractions of purified samples were analyzed for their compositional percentage of the protein of interest. Before LOV2 variants could be tested, the optimal protein concentrations needed for an efficient Sortase A reaction was needed. The reversibility of the reaction prevents it from reaching a high extent of conversion. Initially reactions were set up with 10mM in Sortase A & ELP-LPETG & 30mM in oligoglycine substrate. The oligoglycine substrate was presented in excess in an attempt to drive the reaction forward. This did not prove true. In subsequent trials, reactant concentrations were doubled, and similar results showed. These figures can be seen in the Appendix.

Following analysis of Sortase A's reaction mechanism,¹³ concentrations were adjusted to limit the amount of Sortase A. This was done to attempt to drive the reaction toward transpeptidation rather than the hydrolysis pathway. The ratio used was 4:2:1 (GGG-X:ELP-LPETG:SrtA) with 15mM SrtA. Figure 18 below shows the results of this trial. Using densitometry, it was determined that about 10% of the GGG-mRuby2 reacted at these conditions. This value is not ideal enough for light dependent studies. A large range of values is desirable between light and dark states. Additional trials would have attempted to further limit the amount of Sortase. This has been reported to be done well when substrates have been equally present.²⁵

In addition, due to the poor staining of ELP in Coomassie blue and evidence of low GGG-mRuby2 reaction yields, it was determined a possible GGG-ELP + ELP-LPETG conjugate would be near impossible to see. This was the primary reason to decide to express the other recognition motif Z-Ct-LPETG.

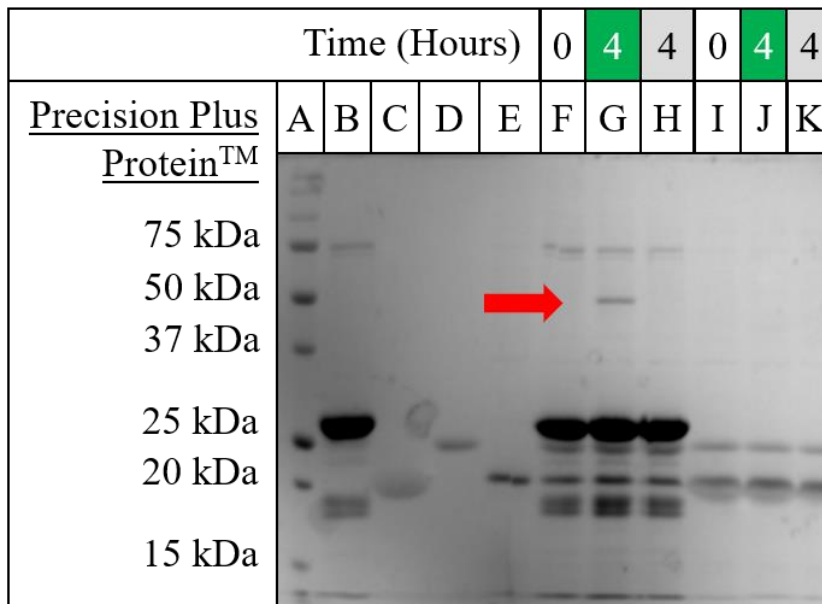


Figure 18: Light Independent Sortase A Transpeptidation. Color key: Green (4 hr Sortase A reaction) Grey (no calcium negative control). Lane profile: (A) Ladder, (B) GGG-mRuby2, (C) GGG-ELP, (D) ELP-LPETG, (E) Sortase A, (F-H) Sortase A + GGG-mRuby2 + ELP-LPETG, (I-K) Sortase A + GGG-ELP + ELP-LPETG. Transpeptidation band observed in lane G at about 49kDa.

3.4 Summary & Conclusions

In this chapter, the reversible Sortase A transpeptidation system was attempted to be controlled with blue light irradiation. By genetically fusing the Sortase A recognition sequence to the C-terminus of the LOV domain, it was hypothesized that the constructs would display similar dark-state peptide caging, as in the LOV-SpyTag constructs previously studied. Constructs were cloned, and resulting peptides were expressed and purified as expected. Light independent reactions were completed to begin the optimization of the Sortase A reaction. Due to time restrictions, light dependent experiments were not conducted.

Depending on the results on initial light dependent experiments, if a ligation band were only present in irradiated samples, further studies would be necessary. FRET experiments would be explored to test the reversibility of the system. Due to the fluorescence of the LOV domain, when conjugated to mRuby2 the construct should exhibit energy transfer from the LOV2 to the mRuby2 as the LOV2 is excited by blue light. The ratio of the mRuby2 to the LOV2 emission peaks can be tracked over time to study reaction kinetics. If the Sortase A reaction remained reversible the addition of nonfluorescent oligoglycine substrate (GGG-ELP) should result in a decrease of FRET response. Given this works we could use this system for light responsive hydrogels and real-time intracellular peptide tracking.

Chapter 4

CONCLUSIONS

4.1 Conclusions

The ability to control protein conjugations and the formation of protein complexes with light would serve as a powerful tool in biotechnology. Protein conjugation methods have been engineered for *in vitro* purposes, like the generation of microarrays, and for *in vivo* purposes, like the control gene expression. Being able to control protein conjugation with the tunability of light, would provide more versatile applications. Light is orthogonal to most biological functions and can be integrated as a simple temporal control switch to turn on and off biological functions in real time.

In the work presented, irreversible and reversible protein conjugation systems are investigated for their ability to be controlled by blue light. Using common cloning methods, a variety of constructs was created in order to incorporate key reaction peptide motifs into the J α helix of the AsLOV2 photoreceptor. These constructs were expressed, purified, and characterized using SDS PAGE.

When SpyTag was fused to the LOV2 domain, the irreversible SpyCatcher:SpyTag system displayed the ability to be dynamically caged. After 5 hours, LOV-537-SpyTag displayed 70% conjugation in blue light and 17% the dark. Genetic fusion of SpyCatcher to the LOV2 domain resulted in the inability to cage all together. Due to initial challenges in cloning and confirmation, light dependent experiments for the reversible Sortase A transpeptidation system were not completed. Light dependent experimentation will help determine the feasibility of this system. If

feasible, scientists could greatly increase their ability to control experimental variables.

REFERENCES

1. Types of Proteins.
<https://learn.genetics.utah.edu/content/basics/proteintypes/> (accessed May 1, 2019).
2. Lutz, Stefan, and Samantha M. Iamurri. "Protein engineering: past, present, and future." *Protein Engineering*. Humana Press, New York, NY, 2018. 1-12.
3. Ka-Hei Siu, Rebecca Chen, Qing Sun, Long Chen, Shen-Long Tsai, and **Wilfred Chen**, Synthetic Scaffolds for Pathway Enhancement, *Curr. Opin. Biotechnol.*, **36**, 98-106, 2015.
4. Murphy, John R. "Protein engineering and design for drug delivery." *Current opinion in structural biology* 6.4 (1996): 541-545.
5. Kim, Heejae, et al. "Bioengineering strategies to generate artificial protein complexes." *Biotechnology and bioengineering* 112.8 (2015): 1495-1505.
6. Jones, Susan, and Janet M. Thornton. "Principles of protein-protein interactions." *Proceedings of the National Academy of Sciences* 93.1 (1996): 13-20
7. Lotze, Jonathan, et al. "Peptide-tags for site-specific protein labelling in vitro and in vivo." *Molecular BioSystems* 12.6 (2016): 1731-1745.
8. Kim, Sanggil, et al. "Direct protein-protein conjugation by genetically introducing bioorthogonal functional groups into proteins." *Bioorganic & medicinal chemistry* 24.22 (2016): 5816-5822.
9. Hao, Ziyang, et al. "Introducing bioorthogonal functionalities into proteins in living cells." *Accounts of chemical research* 44.9 (2011): 742-751
10. Zakeri, Bijan, et al. "Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin." *Proceedings of the National Academy of Sciences* 109.12 (2012): E690-E697.
11. Li, Long, et al. "Structural analysis and optimization of the covalent association between SpyCatcher and a peptide Tag." *Journal of molecular biology* 426.2 (2014): 309-317.

12. Reddington, Samuel C., and Mark Howarth. "Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher." *Current opinion in chemical biology* 29 (2015): 94-99.
13. Clancy, Kathleen W., Jeffrey A. Melvin, and Dewey G. McCafferty. "Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition." *Peptide Science* 94.4 (2010): 385-396.
14. Arkenberg, Matthew R., Dustin M. Moore, and Chien-Chi Lin. "Dynamic control of hydrogel crosslinking via sortase-mediated reversible transpeptidation." *Acta biomaterialia* 83 (2019): 83-95
15. Proft, Thomas. "Sortase-mediated protein ligation: an emerging biotechnology tool for protein modification and immobilisation." *Biotechnology letters* 32.1 (2010): 1.
16. Mao, Hongyuan. "A self-cleavable sortase fusion for one-step purification of free recombinant proteins." *Protein expression and purification* 37.1 (2004): 253-263.
17. Möglich, Andreas, and Keith Moffat. "Engineered photoreceptors as novel optogenetic tools." *Photochemical & photobiological sciences* 9.10 (2010): 1286-1300.
18. Szymański, Wiktor, et al. "Reversible photocontrol of biological systems by the incorporation of molecular photoswitches." *Chemical Reviews* 113.8 (2013): 6114-6178.
19. Tischer, Doug, and Orion D. Weiner. "Illuminating cell signalling with optogenetic tools." *Nature reviews Molecular cell biology* 15.8 (2014): 551.
20. Zayner, Josiah P., Chloe Antoniou, and Tobin R. Sosnick. "The amino-terminal helix modulates light-activated conformational changes in AsLOV2." *Journal of molecular biology* 419.1-2 (2012): 61-74.
21. Lungu, Oana I., et al. "Designing photoswitchable peptides using the AsLOV2 domain." *Chemistry & biology* 19.4 (2012): 507-517.
22. Guntas, Gurkan, et al. "Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins." *Proceedings of the National Academy of Sciences* 112.1 (2015): 112-117.

23. Shaner, Nathan C., et al. "Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein." *Nature biotechnology* 22.12 (2004): 1567.
24. Yo Lam, Amy J., et al. "Improving FRET dynamic range with bright green and red fluorescent proteins." *Nature methods* 9.10 (2012): 1005.
25. Levary, David A., et al. "Protein-protein fusion catalyzed by sortase A." *PloS one* 6.4 (2011): e18342.

Appendix A

DETAILED PROCEDURE AND RESULTS FOR GENERATION OF BLISS SYSTEM CONSTRUCTS

As described in the methods section, primers were designed to amplify the LOV2 protein from the original iLID construct, received from Addgene, and to genetically fuse the SpyTag onto the C-terminus. The SalI restriction enzyme cut site was also incorporated into its genetic code. Table 3 below summarizes the primers used.

Table 3: Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-SpyTag construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow), SalI (blue), HindIII (green). The DNA to code for the SpyTag peptide is also highlighted in pink.

Primer Name	Primer Sequence (5' to 3')	Melt T
Forward Primer	CTT GTG GGA TCC GGG GAG TTT CTG G	57.6 °C
Reverse Primer 537	CAG CTT GTC GAC CAT TAC TAT GTG TGC GGC TGT CTT TTT GAT CAG CAT GAC	56.7 °C
Reverse Primer 540	GCA TGA GTC GAC CAT TAC TAT GTG TGC AAT CTG AAA GGC TGT CTT TTT GAT C	54.0 °C
Reverse Primer 541	CTT CTA GTC GAC CAT TAC TAT GTG TGC AGC AAT CTG AAA GGC TGT CTT TTT G	56.4 °C
Forward Adapter	TCG A CG CCT ACA AGC CGA CGA AGT AAA	-
Reverse Adapter	AGC TT T TAC TTC GTC GGC TTG TAG GCG	-

A PCR cycle was completed, as described in the methods. To analyze the results of the PCR, 2% agarose gels were run to confirm the expected size of the PCR inserts. All PCR reactions could be confirmed successfully beside the LOV537-Tag construct. An additional PCR reaction was conducted again on the LOV537-Tag construct and it was then confirmed in the same fashion as before. After analysis of the controls, it seems that the shared forward primer may have been subjected to plasmid contamination. Figure 19 below shows an example of a PCR confirmation for LOV-Tag 537.

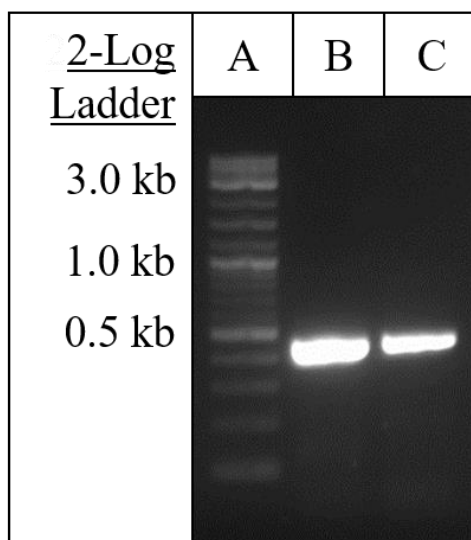


Figure 19: 2% agarose gel of LOV-Tag 537 PCR confirmation. Lane Profile: (A) 2-Log ladder, (B) PCR Sample, (C) PCR control, contaminated by plasmid

Following confirmation of a successful PCR, restriction enzyme digests were completed on both the LOV-X-Tag inserts as well as the iLID ligation vector. The LOV-X-Tag constructs were digested at BamHI & Sali sites and the vector was

digested at BamHI & HindIII sites. The iLID ligation vector digested product was run on 2/3% gel to separate the vector from the insert that was removed. The vector band was imaged and cut out of the agarose gel. The DNA in the slice of agarose was then extracted with the Zymoclean gel extraction kit as stated in the methods.

After phosphorylation of the adapters, a 3-piece ligation was then completed with the designed adapters, and digested vector and inserts containing the genetic sequences for the POIs. The ligation products were then transformed into NEB-5alpha cell line and plated on appropriate antibiotic selective plates. After growth was seen on the plate, colonies were picked and digested with BamHI & Sall to confirm the insert had been properly incorporated. A band representing an insert of about .450kb would confirm the cloning was successful. This can be seen below in Figure 20. The ligation and transformation processes were repeated for the other two constructs once more. This time a positive clone was picked for the 540 truncations. To obtain the 537 constructs, a 2-piece ligation was set up using the LOV540-SpyTag successful clone as a vector to clone into. All three constructs were then confirmed with DNA sequencing.

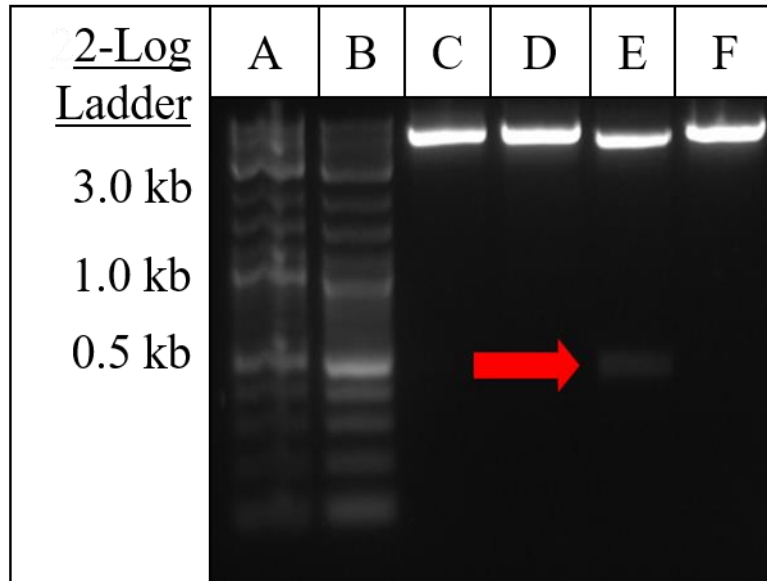


Figure 20: 2% Agarose gel of digested LOV-X-SpyTag Constructs. Lane Profile: (A-B) 2 Log Ladder, (C) LOV-537-SpyTag, (D) LOV-540-SpyTag, (E-F) LOV-541-SpyTag. It is observed that the only positive clone in this screening is LOV-541-SpyTag in lane 5.

In order to clone the SpyCatcher to the C-terminus of the AsLOV2 protein, with extra amino acids resulting in a restriction site, Gibson assembly was utilized. Primers were designed so that the LOV2 domain and SpyCatcher sequences had overlapping ends for Gibson assembly. Table 4 below summarizes the necessary primes.

Table 4: Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-SpyCatcher construct. Regions homologous to the template plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow) & HindIII (green).

Primer Name	Primer Sequence (5' to 3')	Melt T
SpyCatcher – Forward	GAT AGT GCT ACC CAT ATT AAA TTC TCA AAA CGT G	57.6 °C
SpyCatcher-Reverse	GTC CGT AAG CTT TTA GCC ATT TAC AGT AAC CTG ACC TTG C	58.7 °C
LOV2-Forward	CTT GTG GGA TCC GGG GAG TTT CTG G	57.6 °C
LOV2-Reverse-537	TTA ATA TGG GTA GCA CTA TCGG CTG TCT TTT TGA TCA GCA TGA C	55.7 °C
LOV2-Reverse-540	TTA ATA TGG GTA GCA CTA TCAA TCT GAA AGG CTG TCT TTT TGA TC	52.8 °C
LOV2-Reverse-541	TTA ATA TGG GTA GCA CTA TCA GCA ATC TGA AAG GCTG TCT TTT TG	54.4 °C

PCR was conducted as stated in the methods and 2% agarose gels were ran to observe the results. The two DNA fragments were then allowed to combine via Gibson assembly. The product of this reaction was then amplified via PCR where a HindIII restriction enzyme site was also incorporated. An agarose gel of this amplification can be seen in Figure 21 below with a band at .69kb.

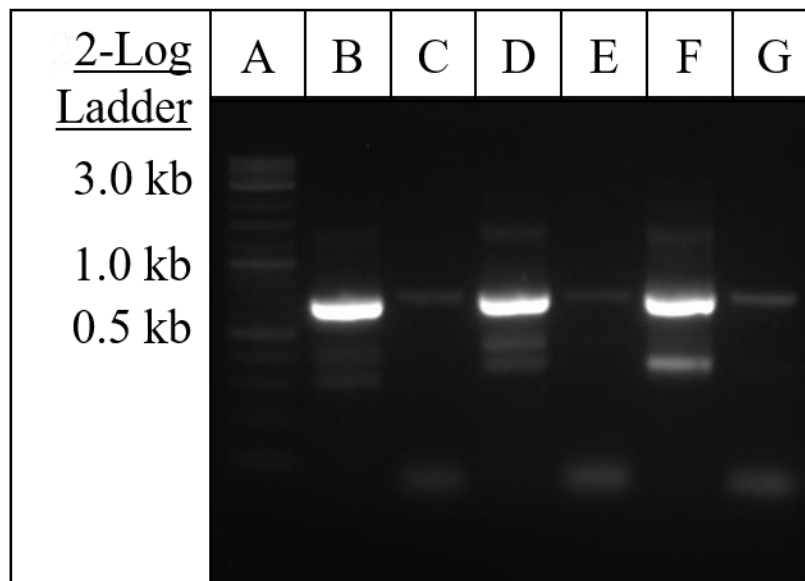


Figure 21: 2% agarose gel of PCR amplification of products of Gibson Assembly. Lane Profile: (A) 2-Log Ladder, (B) LOV-SC-537 Gibson Insert, (C) LOV-SC-537, (- control, no template), (D) LOV-SC-540 Gibson Insert, (E) LOV-SC-540, (- control, no template), (F) LOV-SC-541 Gibson Insert, (G) LOV-SC-541, (- control, no template).

A 2-piece ligation was then able to be performed with the amplified Gibson product and the iLID vector backbone. Both DNA sequences were cut with appropriate enzymes and ligated together. Ligated plasmids were then transformed into NEB5a and plated on appropriate antibiotic selective plates. Colonies were picked and plasmids were extracted using Zyppy plasmid extraction kit as stated in the methods. Plasmids were cut with HindIII & BamHI. Positive clones were determined by having the correct size band, which would be the expected size of the Gibson product (about .68kb). As seen in the Figure 22 below, positive clones were picked for the 537 and 540 truncations. Further screening of clones of the 541-truncation prevailed and positive clones were found for this construct as well.

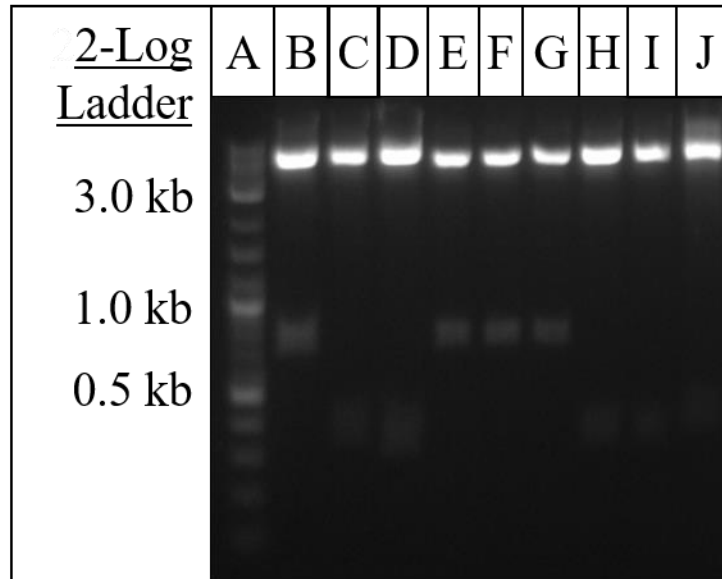


Figure 22: 2% Agarose gel of digested LOV-X-SpyCatcher Constructs. Lane Profile: (A) 2 Log Ladder, (B-D) LOV-537-SpyCatcher A-C, (E-G) LOV-540-SpyCatcher A-C, (H-J) LOV-541-SpyCatcher A-C. Positive clones are observed in LOV-537-SpyCatcher A & LOV-540 A, B, & C.

Appendix B

DETAILED PROTEIN EXPRESSION & CHARACTERIZATION RESULTS FOR BLISS SYSTEM

Protein expression of the iLID construct in pLysS cell line was induced with 333 μ M IPTG and let to grow overnight. Cells were harvested in the morning and components were obtained by the breaking of the cell membrane via sonication. POI's were extracted by sonication as stated in the Methods and concentrations were determined by a Bradford Assay, also stated in the methods. His-tag column purification was conducted, and samples were collected from each step. These samples were analyzed for their size via SDS-PAGE gels electrophoresis and ran on a 12.5% SDS gel.

Similar results were seen in expression of the engineered constructs as the unaltered iLID construct. Spy-X-Tag conjugates should be about 18.9 kDa and Spy-X-Catcher conjugates should be about 26.5 kDa. Additional images not included in Chapter 2 are seen in Figure 23 below. High levels of the POI can be seen in the soluble lysate and the elution bands of each construct.

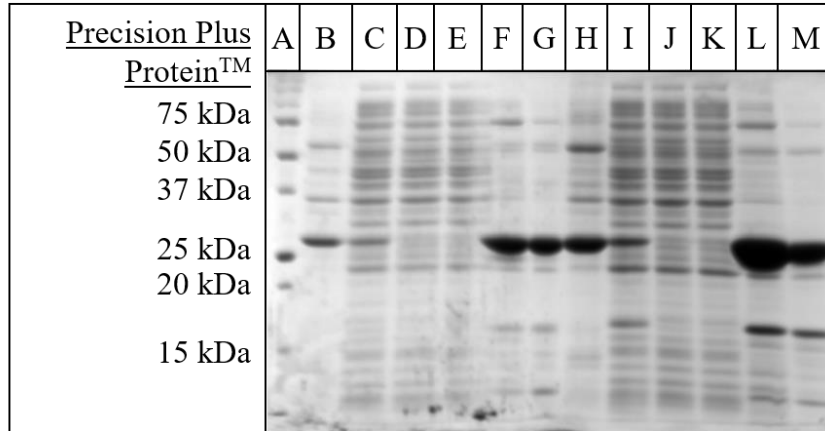


Figure 23: SDS PAGE of LOV-X-SpyCatcher purification process. Lane profile: (A) Precision Plus Ladder, (B-F) LOV-537-SpyCatcher (Insoluble Lysate, Soluble Lysate, Flow Through 1, Flow Through 2, Elution 1, Elution 2), (H-N) LOV-540-SpyCatcher (Insoluble Lysate, Soluble Lysate, Flow Through 1, Flow Through 2, Elution 1, Elution 2).

Once all constructs were confirmed, blue light experimentation was begun. Of the eluted samples, densitometry was used to determine the percentage of POI in each sample. These values were used to calculate samples compositions and were considered to determine ligation yield. Following initial results, only the LOV-X-SpyTag constructs were chosen for further experimentation. Reactions were conducted for a longer time and time points were taken more frequently. Figures X & X below show the results from the LOV-540-SpyTag & LOV-541-SpyTag constructs.

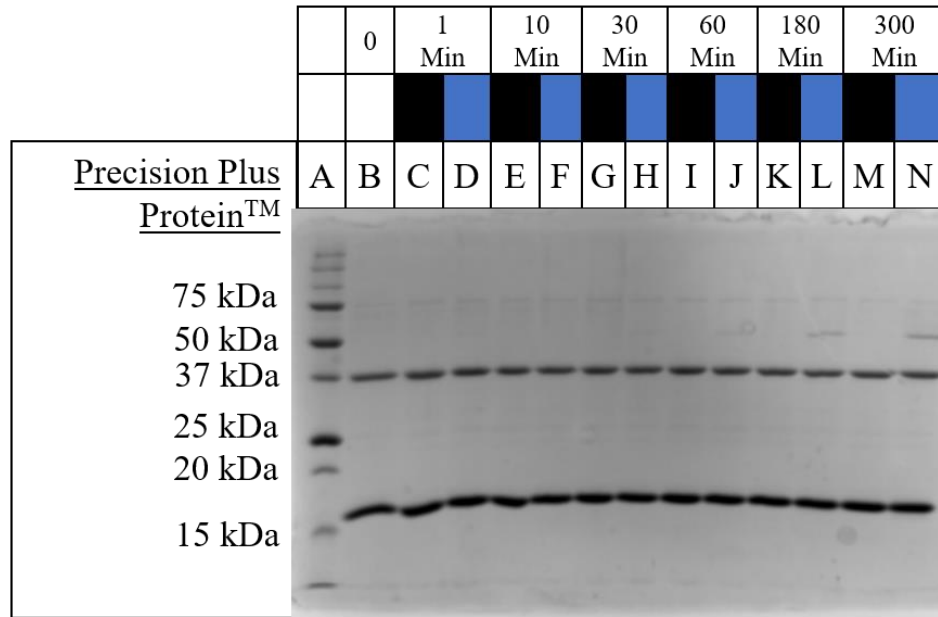


Figure 24: 12.5% SDS PAGE of LOV-540-SpyTag + SpyCatcher-ELP Experiments. Blue lanes show samples irradiated in blue light and black lanes show samples hidden from light.

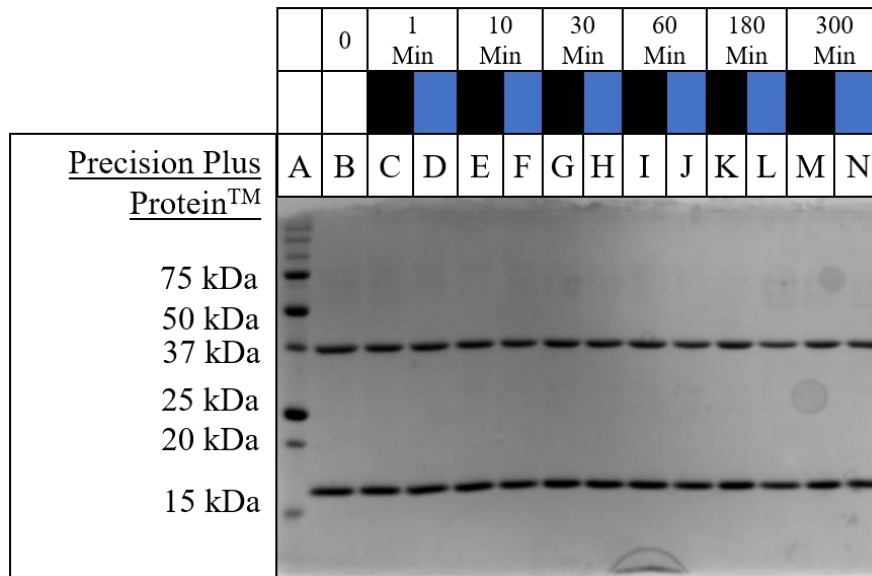


Figure 25: 12.5% SDS PAGE of LOV-541-SpyTag + ELP-SpyCatcher Experiment. Blue lanes show samples irradiated in blue light and black lanes show samples hidden from light.

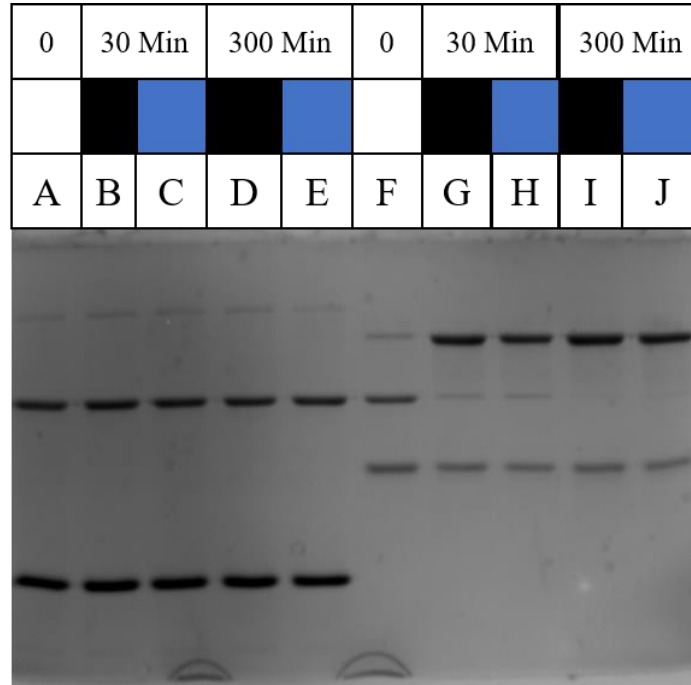


Figure 26: LOV-X-SpyTag Blue Light Experimental controls. Lane Profile: No ladder shown, (A-E) Negative control reacting SpyCatcher + iLID, (F-J) Positive control reacting SpyTag + SpyCatcher.

Appendix C

DETAILED PROCEDURE AND RESULTS FOR GENERATION OF LOV- CONSTRUCTS FOR SORTASE TRANSPEPTIDATION

Primers were designed to amplify the DNA for the LOV construct with a Sortase A recognition peptide sequence (LPXTGGSG) appended into three different locations of the $J\alpha$ helix. The constructs were cloned to append the recognition site at the 540 & 541 amino acid site where the third amino acid in the recognition sequence is Glutamic Acid (E) and at the 537 amino acid site where the third amino acid is Isoleucine (I). The most common amino acid seen in literature for X is glutamic acid, but at the 539th residue of the LOV domain is a highly conserved isoleucine. HindIII sites were also incorporated with PCR design. The original iLID structure was used as both a PCR template and a ligation vector backbone.

Table 5: Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-LPXTGGSG construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow) & HindIII (green). The DNA to code for the LPXTGGSG (red) peptide and the STOP (black) codon is also highlighted.

Primer Name	Primer Sequence (5' to 3')	Melt T
Forward Primer	CTT GTG GGA TCC GGG GAG TTT CTG G	57.6 °C
Reverse Primer 537	CA TGA AAG CTT CTA GCC AGA CCC GCC GGT AAT AGG AAG GGC TGT CTT TTT GAT C	44.6 °C
Reverse Primer 540	CA TGA AAG CTT CTA GCC AGA CCC GCC GGT CTC AGG AAG AAT CTG AAA GGC TGT CTT TTT G	51.7 °C
Reverse Primer 541	CTT CTA AAG CTT CTA CCC AGA TCC ACC TGT TTC CGG TAA AGC AAT CTG AAA GGC TGT C	52.8 °C

PCR reactions were conducted as stated in the Methods and the resulting products were run on a 2% agarose gel. Figures 27 & 28 below show the results of the PCR. Expected sizes of about .46kb are observed.

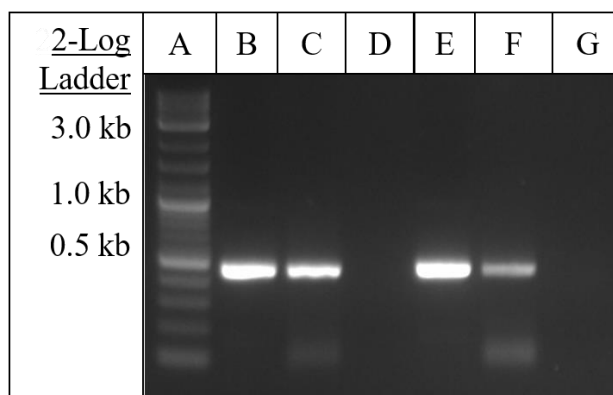


Figure 27: 2% Agarose gel for PCR results for amplification of inserts for LOV-540-LPETG & LOV-541-LPETG. Lane Profile: (A) 2 Log Ladder, (B-D) LOV-540-LPETG [sample, no plasmid control, no primer control], (E-G) LOV-541-LPETG [sample, no plasmid control, no primer control].

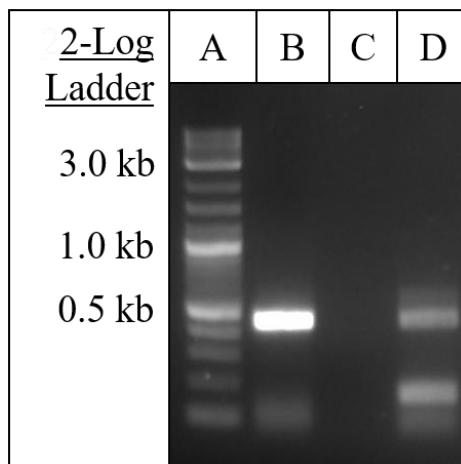


Figure 28: 2% Agarose gel for PCR results for amplification of inserts for LOV-537-LPITG Lane Profile: (A) 2 Log Ladder, (B) sample, (C) no primer control, (D) no plasmid control.

Following confirmation of a successful PCR, appropriate enzyme digests were performed overnight, and constructs were ligated and transformed. Clones were confirmed by DNA sequencing. An example DNA sequencing success is seen in Figure 29. Initially, it was thought the positive clones were achieved for the 541 and 540 truncations. However, after more careful analysis of the sequencing results, it was determined that the LOV-540 clone was not a positive one. The procedure was repeated, and a positive clone was confirmed. In addition, the cloning of the 537 constructs took longer than expected. It was also separately confirmed via DNA sequencing. An example of DNA sequencing results can be seen below. Positive clones were then transformed into production cell lines for protein expression.

```

Alignment of Sequence_1: [Sequence Window #2] with Sequence_2: [LOV2_LPXTGGSG_541.txt.xdna]
Similarity : 847/5162 (16.41 %)

Seq_1 1      ATGAGAGGATCGCATCACCATCACCATCACGGATCCGGGGAGTTTCTGGCAACCACACTG 60
          |||
Seq_2 1      ATGAGAGGATCGCATCACCATCACCATCACGgatcCGGGGAGTTTCTGGCAACCACACTG 60

Seq_1 61      GAACGGATCGAGAAAAATTCGTGATTACTGATCCGAGACTGCCTGACAAACCAATCATT 120
          |||
Seq_2 61      GAACGGATCGAGAAAAATTCGTGATTACTGATCCGAGACTGCCTGACAAACCAATCATT 120

Seq_1 121     TTTGCGAGCGATTTCCTTCTGCAGCTGACAGAATATTCTCGGGAAGAGATCCTGGGGCGC 180
          |||
Seq_2 121     TTTGCGAGCGATTTCCTTCTGCAGCTGACAGAATATTCTCGGGAAGAGATCCTGGGGCGC 180

Seq_1 181     AATTGCCGTTTTTCTGCAGGGACCCGAGACAGACCCGTGCCACTGTTCCGAAAATCAGAGAT 240
          |||
Seq_2 181     AATTGCCGTTTTTCTGCAGGGACCCGAGACAGACCCGTGCCACTGTTCCGAAAATCAGAGAT 240

Seq_1 241     GCTATTGACAACCAGACTGAAGTGACCGTTCAGCTGATCAATTATACCAAGAGCGGCAAG 300
          |||
Seq_2 241     GCTATTGACAACCAGACTGAAGTGACCGTTCAGCTGATCAATTATACCAAGAGCGGCAAG 300

Seq_1 301     AAGTTCCTGGAACGTGTTCCACCTGCAGCCGATGCGCGATTATAAGGGCGACGTCCAGTAC 360
          |||
Seq_2 301     AAGTTCCTGGAACGTGTTCCACCTGCAGCCGATGCGCGATTATAAGGGCGACGTCCAGTAC 360

Seq_1 361     TTCATTGGCGTGCAGCTGGATGGCACCGAACGTCTTCATGGCGCCGCTGAGCGTGAGGCG 420
          |||
Seq_2 361     TTCATTGGCGTGCAGCTGGATGGCACCGAACGTCTTCATGGCGCCGCTGAGCGTGAGGCG 420

Seq_1 421     GTCATGCTGATCAAAAAGACAGCCTTTCAGATTGCTTTACCGGAAACAGGTGGATCTGGG 480
          |||
Seq_2 421     GTCATGCTGATCAAAAAGACAGCCTTTCAGATTGCTTTACCGGAAACAGGTGGATCTGGG 480

```

Figure 29: DNA Sequencing Results for LOV-541-LPETGGSG. Alignment of sequences used via Serial Cloner software. The top sequence represents the sequenced DNA and the bottom is the theoretical construct. Green region represents the N-terminal Histidine Tag and the Purple represents the LPETGGSG motif.

Primers were also designed to incorporate Venus, a yellow fluorescent protein, to the N-terminus of the LOV protein. The purpose of this was to use its fluorescence in a FRET experiment to determine the reversibility of the system. Due to difficulty with cloning, time constraints, and the fact that the LOV protein does display some fluorescence, the cloning of these constructs was abandoned.

Primers were also designed to create oligoglycine tags. Primers amplified the portion of DNA that coded for a fluorescent protein (mRuby2, mCherry, GFP) in

different clones already available from group members. The primers designed to amplify the red fluorescent proteins (mRuby2, mCherry) were also designed to incorporate the EcoRI & XhoI restriction cut sites. The PCR inserts were cloned into a vector that would provide the protein product a N-terminal oligoglycine peptide and a C-terminal His Tag for purification purposes.

Table 6: Summary of PCR Primers & Ligation Adapters used in Cloning of LOV-LPXTGGSG construct. Regions homologous to iLID plasmid are shown in bold. The melting temperature of homologous region are reported. Key enzyme restriction sites are highlighted in the following manner: BamHI (yellow). HindIII (green), SacI (royal) , EcoR1 (navy), NdeI (purple), & XhoI(grey).

Primer Name	Primer Sequence (5' to 3')	Melt T
RFP FWD	GAA TGC GAA TTC GGA TCC ATG GTG TCC AAA GG	55.7 °C
RFP RVS	GAA TGC CTC GAG GAG CTC CTT ATA CAA TTC ATC C	50.7 °C
GFP FWD	GAA TGC CAT ATG GCT AGC GTG AGC	53.1 °C
GFP RVS	GAA TGC CTC GAG GAC AAG AGA CTC	51.4 °C

PCRs were conducted as stated in the methods. Successful reactions were confirmed via agarose gel electrophoresis. Results can be found in Figure 30 below.

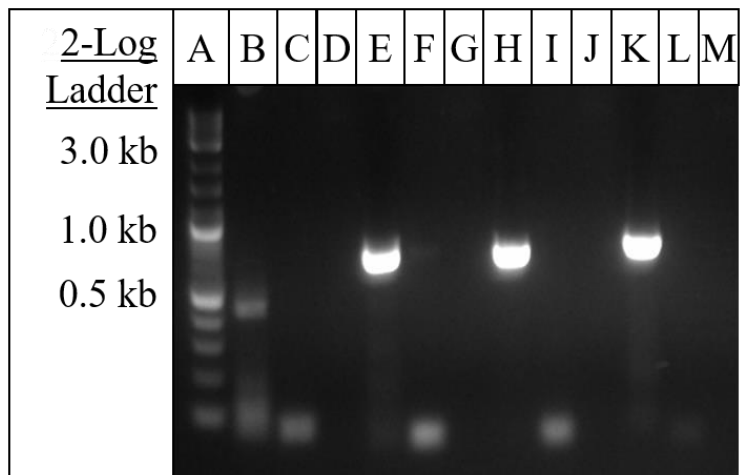


Figure 30: 2% Agarose gel for PCR results. After the ladder all lanes go in the following order: sample, no plasmid control, no primers control Lane Profile: (A) 2 Log Ladder, (B-D) LOV-537-LPETG, (E-G) GGG-mCherry, (H-J) GGG-mRuby2, & (K-M)GGG-GFP. Expected sizes are confirmed for all constructs.

Following confirmation of a successful PCR, appropriate enzyme digests were performed overnight. Following digestion, constructs were ligated and transformed into competent cells. GGG-mRuby2 & GGG-mCherry constructs were later confirmed with restriction enzyme digests, by observing the size of the inserted DNA. Unfortunately, there was difficulty in cloning the GGG-GFP structure, so this construct was determined unnecessary and abandoned. The results of the screening digest that was used to confirm the GGG-mRuby2 & GGG-mCherry clones is shown in the main body. Positive clones were then transformed in production cell lines for protein expression.

Appendix D

ADDITIONAL RESULTS FOR EXPRESSION AND PURIFICATION OF LOV- CONSTRUCUTS FOR SORTASE TRANSPEPTIDATION

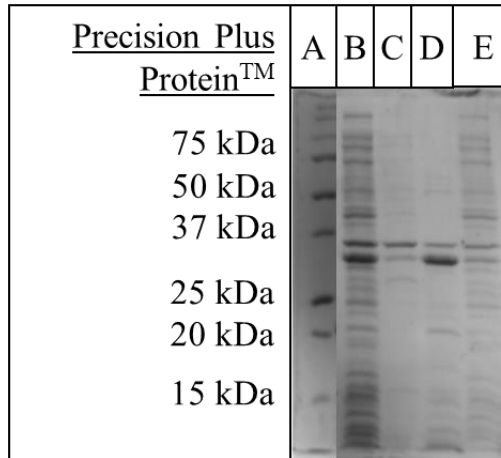


Figure 31: Expression and purification of Z-Ct-LPETG. Lane profile: (A) Ladder, (2) Soluble Lysate, (3) Insoluble Lysate, (4) Purified, (5) Resolubilized Pellet. Protein can be seen at about 29.7 kDa.

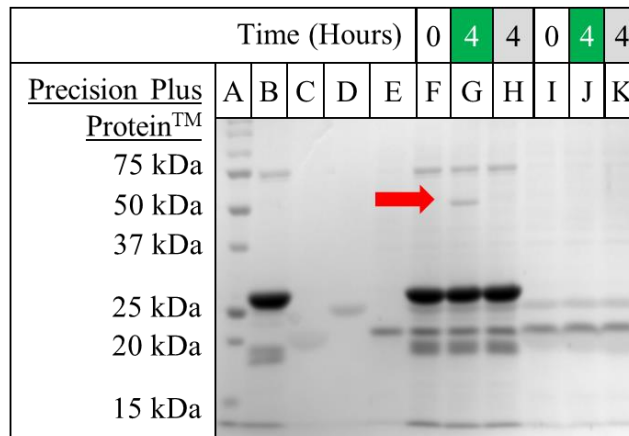


Figure 32: Sortase A reaction at 60mM GGG-X substrate and 20mM ELP-LPETG & Sortase A. Lane profile: (A) Ladder, (B) GGG-mRuby2, (C) GGG-ELP, (D) ELP-LPETG, (E) Sortase A, (F-H) Sortase A + GGG-mRuby2 + ELP-LPETG, (I-K) Sortase A + GGG-ELP + ELP-LPETG. Transpeptidation band observed in lane G at about 49kDa.