THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CHEMICAL ENGINEERING

DEVELOPMENT AND COMPARISON OF POTENTIAL HIV-TRANSMISSION PREVENTION PROTEINS NATIVE TO FRESHWATER CYANOBACTERIA

RYAN J. JONES
SPRING 2018

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Chemical Engineering
with honors in Chemical Engineering

Reviewed and approved* by the following:

Wayne Curtis
Professor of Chemical Engineering
Thesis Supervisor

Scott Milner
Joyce Chair and Professor of Chemical Engineering
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

As cases of Human Immunodeficiency Virus (HIV) continue to rise amongst adolescents and young adults in sub-Saharan Africa, the need for additional support becomes more urgent. Particularly in developing nations, access to anti-retrovirals is highly dependent on socioeconomic and psychological barriers that are not easily breached. Alternatively, there exists a lectin protein, cyanovirin-n (CVN), that prevents the transmission of HIV through inhibition of the gp120 and CD4+ receptor complex. This work sought to develop a platform to clone, express, and purify cyanovirin for end use in a topical anti-viral. CVN was successfully cloned into a cyanobacteria transformation vector, but could not transform *Synechococcus* 7002 due to apparent toxicity. In an effort to alleviate this toxicity, a cloning strategy was developed to fuse the CVN gene to an aggregation tag to alter its folding properties and reduce cellular toxicity upon expression. In tandem, a strategy was developed to isolate *Microcystis aeruginosa* from algal blooms in freshwater bodies near Lake Wallenpaupack in Northeast Pennsylvania. This cyanobacteria produces microvirin (MVN), a similar protein to CVN exhibiting less *in vivo* toxicity. Attempts to amplify the MVN gene from these water samples were unsuccessful; however, to facilitate future success of this project, a procedure to clone genes from cyanobacteria culture has been developed. Additional progress towards cloning MVN has involved isolation of *Microcystis aeruginosa* by serial dilution plating and non-photosynthetic contamination testing. Successful fusion of these anti-HIV transmission proteins to aggregation tags would provide a cost-effective and scalable platform for production of antiviral microbicides in *Synechococcus*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I cannot express enough gratitude to Dr. Wayne Curtis for his unwavering commitment to his undergraduate researchers. He has taught me everything I know about how *real* research is conducted and the dedication (and scrappiness) required to solve real-world problems. I am grateful for all that I have learned from him, and I will carry that knowledge with me throughout my professional career. I would also like to thank my Honors Adviser, Dr. Scott Milner, for reading this thesis and offering his insight. Furthermore, I must extend my appreciation to Dr. Barry O'Keefe from the National Cancer Institute for his generous gift of cyanovirin constructs and protein, without which this project would not have been possible.

I am wholeheartedly indebted to my mentor and friend Tina Shing Li Lai for teaching me everything I know about molecular cloning, protein expression, and plant biology. Her knowledge and patience have been infinitely valuable to me throughout my undergraduate career. I would also like to thank Erik Curtis for laying the groundwork for this project by developing the technique for cyanobacteria transformations and culture maintenance and Dr. Donald Bryant for generously gifting our lab with WT *Synechococcus* 7002.

The following colleagues have made my experience with undergraduate research an enjoyable one, and I am so thankful to them for their collaboration, advice, and friendship: Erica Lennox, Grant Gill, Nate Hamaker, Samwel Kariuki, Katie Legenski, Bennett McKinley, Luke Nugent, Morgan Shires, Noah Willis, and Krishnan Sreenivas.

Lastly, I cannot thank my family and friends enough for their continued love and support; in particular, my girlfriend Allie Abdalla who has been my greatest source of inspiration and encouraged me to succeed every step of the way.

Chapter 1

Introduction

The Human Immunodeficiency Virus (HIV) is a chronic, fatal virus that infects healthy human T cells, forcing the cell to reproduce the virus to further proliferate the infection. Loss of infected cells slowly deteriorates the immune system and the HIV develops into Acquired Immune Deficiency Syndrome (AIDS), making it nearly impossible for an individual to fight off diseases[1]. More alarming than the effects of HIV are the lack of current, feasible treatment options for the disease, turning an easily transmitted virus into a full-blown epidemic.

In 2016, 36.7 million individuals were affected by HIV worldwide; 2.1 million of these cases were in children under the age of 15. In the same year over 1.8 million diagnoses came to light, equating to approximately 5,000 individuals per day[2]. The majority of HIV cases arise in low to mid income countries in particular, with sub-Saharan Africa showing the greatest effects. Housing only 12 percent of the world population, sub-Saharan African countries account for roughly 71 percent of the global HIV burden, being transmitted primarily through heterosexual sex encounters and the vertical transmission from mother to child during pregnancy or childbirth[3,4].

A growing number of HIV cases have been treated in recent years using antiretroviral treatments (ART); approximately 18.2 million infected individuals were using this treatment by June 2016[2]. Though ART seems to be an increasingly effective measure for preventing the transmission of HIV, the costs and executions of antiretroviral treatments severely limits utility to the majority of infected individuals. Using sub-Saharan Africa as a primary example, ART would run approximately $1,700 per patient per year in these regions of the world, whereas the annual income for someone in this region is approximately $2,041[5, 6]. Financial burden aside, ART's are

typically inaccessible due to social and psychological barriers common in the third world[3]. This begs the question, how can this flawed system be improved?

Cyanovirin is a surface-binding lectin protein capable of binding to an HIV-infected cell and blocking contact-mediated transmission, acting similar to a shield or envelope, but can this application be practically implemented[7]? Is it comparable and/or more beneficial than the ART currently available to those infected by HIV? Why cyanovirin?

Cyanovirin is a protein from a cyanobacteria (blue-green algae) where its role appears to facilitate binding and colony formation. This protein has the potential for incorporation into a self-applied topical cream, providing a promising alternative to the current, limited, and expensive treatments available for the control of HIV[8]. It is somewhat ironic that this potentially beneficial protein is a component of Harmful Algal Blooms (HAB's). HAB's are currently seen by the Environmental Protection Agency as a threat to human ecosystems, aquatic ecosystems, and the economy (Figure 1)[9,10]. Cyanobacteria that produce lectins form colonies; combined with gas vesicle floatation to the surface, these cyanobacteria can form dense concentrations. In addition to causing acute and chronic illnesses through both direct and indirect contact, these HAB's proliferate drastically in the summer months, negatively impacting the global carbon cycle, killing wildlife and greenery, polluting local water reservoirs, and disrupting natural ecosystem functions[11]. These dangerous water sources cause a number of economic issues, increasing the need for water treatment facilities and decreasing the value of properties and recreational venues[12]. Unfortunately, despite the prevalence of natural lectins in the environment, separation of biomolecules from natural mixtures is extremely difficult, inefficient, and expensive. Therefore, the strategy developed in this thesis is to genetically engineer a greatly simplified separation

approach that will be placed into an alternative cyanobacteria that does not have toxins but retains the advantage of autotrophic (photosynthetic) production.

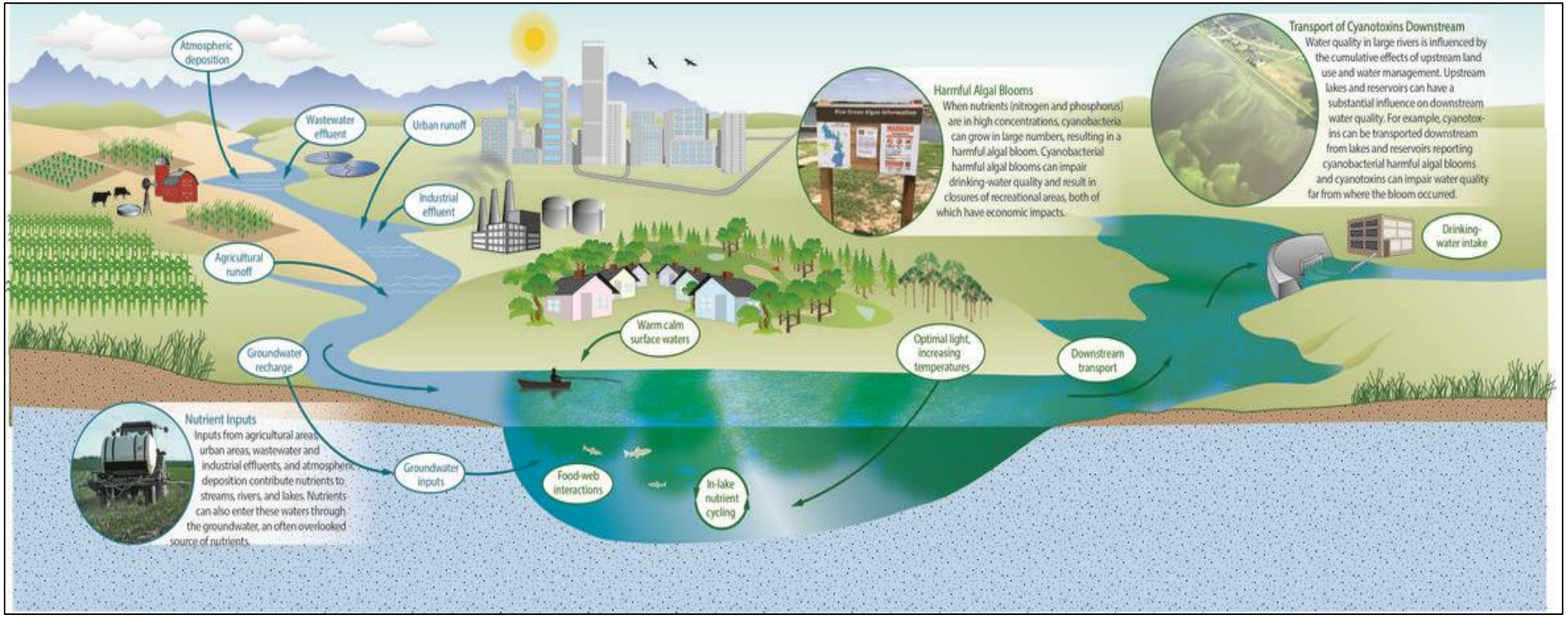

Figure 1. Environmental Impacts of Harmful Cyanobacterial Algal Blooms (HAB's)[13]

Chapter 2

Background Information

2.1 The Anti-HIV Protein(s)

The HIV virus spreads primarily through direct contact with the infection and through the exchange of fluids typically during unprotected sex or intravenous drug use. After exposure, the virus can cross the vaginal epithelium through infection of the epithelial cells, transcytosis or transmigration via donor cells through the epithelial membrane, or breaches in the membrane from lesions (Figure 2). Once inside the epithelium, the virus infects other cells when gp120 (glycosylated protein 120) on the viral surface envelope binds to receptors on healthy cells. The most common target receptors for HIV infection are CD4+ and DC-SIGN both of which play significant roles in antigen detection via notification of other immune cells, particularly C8 cells which are responsible for killing infectious cells[14, 15].

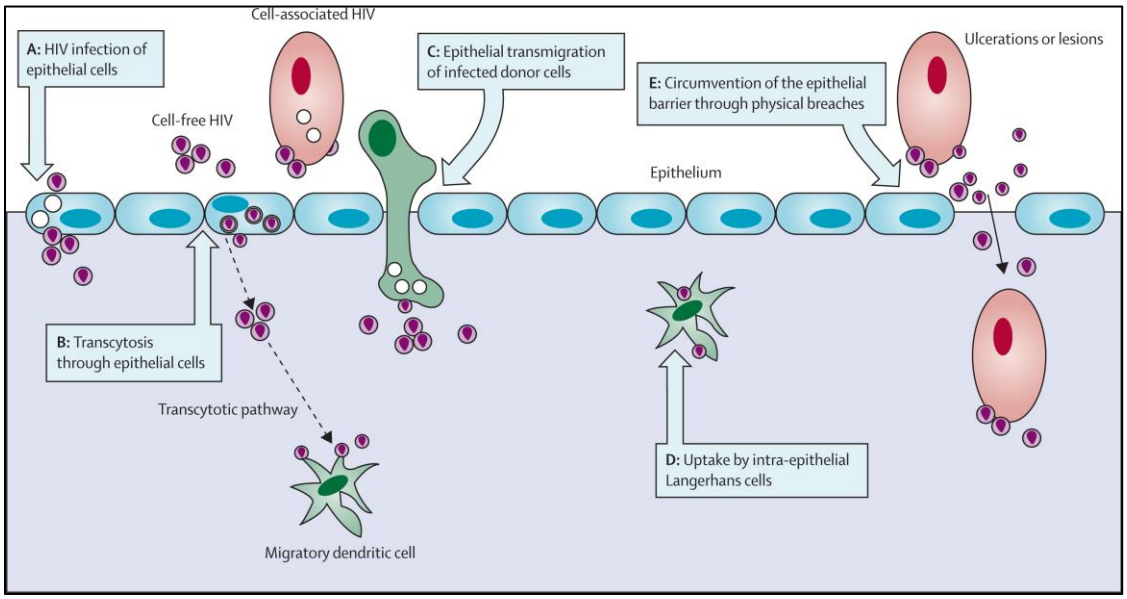

Figure 2. Mechanisms of HIV Breach into the Vaginal Epithelium[16]

There exists a handful of lectin (sugar-binding) proteins that exhibit anti-HIV transmission capabilities by binding to the gp120 surface protein of the HIV virus and preventing it from contacting the CD4+ receptor (Figure 3)[8].

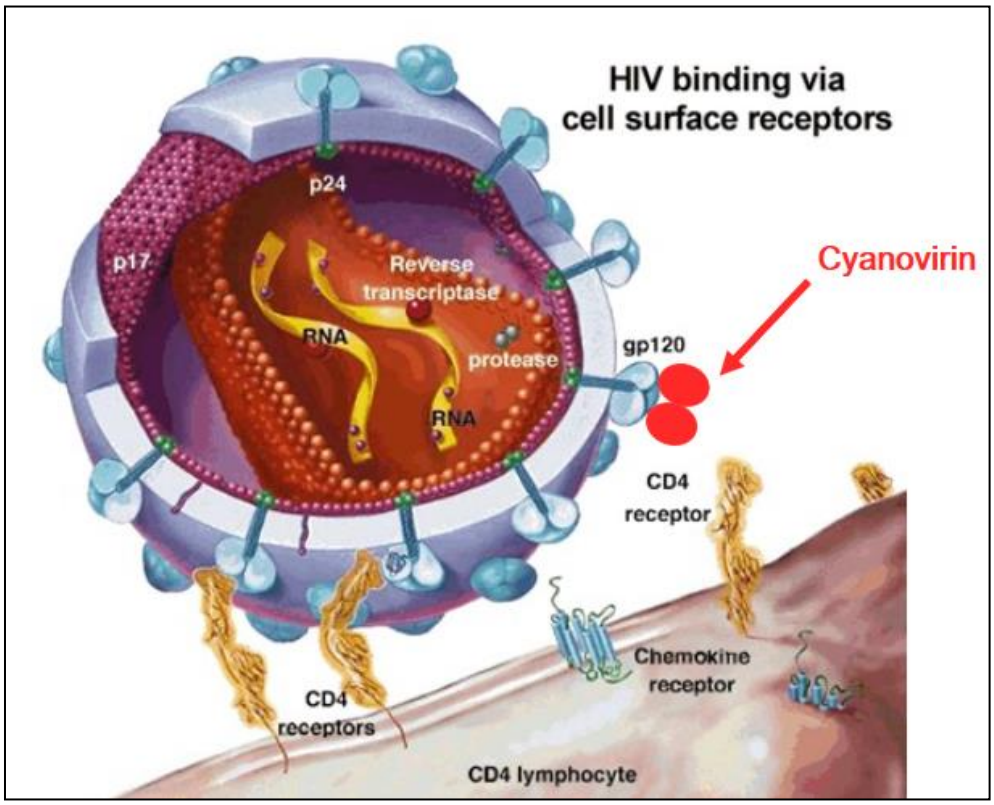

Figure 3. Cyanovirin Disruption of the CD4 Receptor via gp120[17]

Three of the main proteins studied for application in antiviral microbicides that exhibit this mechanism of action are cyanovirin-n (CVN), native to the cyanobacterium *Nostoc ellipsosporum*, microvirin (MVN), native to the cyanobacterium *Microcystis aeruginosa*, and griffithsin, native to the tobacco-relative plant *Nicotiana benthamiana*[18]. Although they cannot be used to generate a "cure" for the virus, they have practical applications (e.g. in topical creams) to reduce the spread of the virus in undeveloped areas where population density is high and access to anti-retroviral treatment is low or up-and-coming. These creams could be produced and distributed with relatively low costs compared to anti-retroviral treatment and would likely not require a certified physician

to administer. However, given the project focus of expressing these anti-HIV proteins in cyanobacteria, griffithsin was not pursued as a potential microbicide candidate despite its efficacy and safety profile[19].

Although cyanovirin has exhibited reasonable efficacy in treating SIV-infected rhesus macaques (*Macaca mulatta*), there is a notable safety issue of cytotoxicity[20]. The desirable protein to pursue in this case is microvirin because, despite having 33% homology to the cyanovirin protein, it has shown drastically less cytotoxicity while producing similar treatment results in preliminary studies[21, 22].

2.2 Disadvantages of *E. coli* as an Expression Platform for Recombinant Proteins

Despite the versatility of *E. coli* as an expression system, there are several disadvantages that negatively impact scalability and cost effectiveness. Among these are inclusion body formation, inducible promoters, and the requirement of separate strains for plasmid proliferation and protein expression.

During T7 overexpression of recombinant proteins, *E. coli* will often sequester the foreign protein in vesicles called inclusion bodies that alter the size, shape, and configuration of the protein, making it difficult to purify the protein without further processing (processing that is typically not conducive to scaling) (Figure 4)[23].

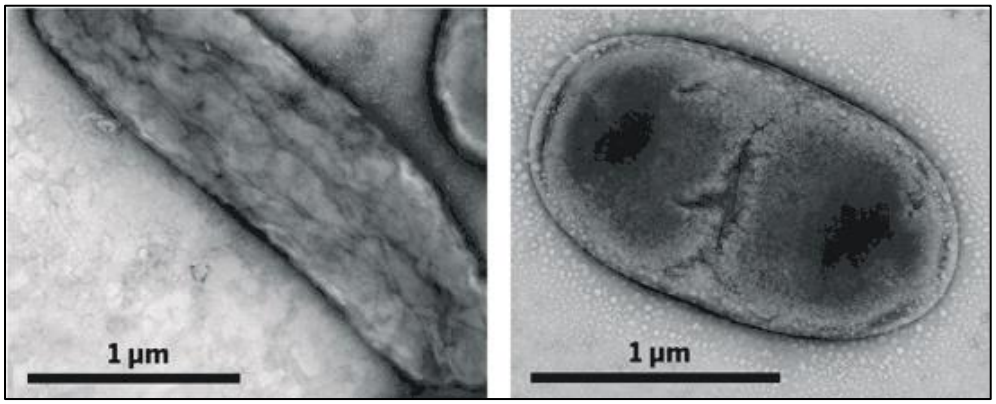

Figure 4. Inclusion Body Formation (Right) *in E. coli* due to Overexpression of Recombinant Protein[24]

Expression in *E. coli* via the T7 promoter also requires activation of the lac operon to facilitate transcription via the RNA polymerase. This activation process requires use of an inducing agent; Isopropyl β-D-1-thiogalactopyranoside (IPTG) is typically used but it is very costly, especially for large-scale cultures.

Additionally, different strains of *E. coli* are required for different applications. TOP10 cells are used for generic cloning work and plasmid proliferation, while protein expression requires BL21 cells that contain T7 polymerase. This requires twice the amount of transformations and agar plates per construct to confirm successful protein expression.

Protein expression via cyanobacteria requires a slightly more complicated scale-up due to dependence on carbon dioxide and light for adequate growth, but the added perks of advantaging the CPC constitutive promoter and eliminating inclusion body formation justify the strategy[25].

Chapter 3

Materials and Methods

3.1 Organisms

All molecular cloning was accomplished using TOP10 *E. coli* as the host organism for transformation and cryopreservation. The marine cyanobacterium *Synechococcus* sp. PCC 7002 (WT provided by Dr. Donald Bryant) was used as the host for all cyanobacteria transformations.

Additionally, in an effort to obtain WT *Microcystis aeruginosa*, samples were harvested September 2017 from three freshwater bodies in Northeastern Pennsylvania: Lake Wallenpaupack, Porter's Lake, and Decker's Pond. The motivation behind taking samples from multiple sources was ultimately to compare the variations in the internal region of the microvirin gene (between the functional carbohydrate-binding ends) as a function of the source environment.

3.2 Plasmids

The cyanovirin gene was amplified from pET26B-CVN, supplied by Dr. Barry O'Keefe. This form of the CVN gene was a synthetic construct optimized for expression in *E. coli*. The intein-ELP complex was amplified from pET/ELP-I-CAT provided by Dr. David Wood, Professor of Chemical Engineering at The Ohio State University. The backbone source for cloning work for cyanobacteria homologous recombination was pAQ1-Ex-YFP, constructed in-house by Erik Curtis[26]. Plasmid maps and sequences for these genes and constructs can be seen in Appendices B and D, respectively.

3.3 Oligonucleotides

A table of all relevant oligonucleotides can be found in Table 1 in Appendix C. They are categorized by function including traditional cloning, fusion PCR, sequencing, and organism confirmation. This table also details the restriction site, base pairs for successful cleavage, and the target gene sequence.

3.4 Molecular Cloning Methods

3.4.1 Restriction Digest Cloning

All genes were cloned into the corresponding expression vectors via restriction digest cloning, the general procedure for which is outlined below:

1. Assuming the gene sequences are known, 5'→ 3' forward and reverse oligonucleotides were designed to amplify the first and last ~10 base pairs of the gene of interest. Basic 5'→3' primer anatomy is characterized by 3-6 base pairs for cleavage close to the ends followed by the desired enzyme restriction site (determined by the location of the multiple cloning site (MCS) in the vector) followed by a start or stop codon or base pairs to ensure the gene is in frame if a start codon is present in the restriction site) followed by the gene fragment. Careful attention is required when designing the reverse primer to ensure the reverse complement of the sequence of interest is used.
2. The genes were then amplified via Polymerase Chain Reaction (PCR) in a BioRad T100 thermalcycler. PCR reactions include the gene source (typically genomic or

plasmid DNA), the forward and reverse oligos, deoxyribonucleotide triphosphates (dNTP's) which are the ATGC building blocks of DNA strands, a DNA polymerase (Taq is the most common for simple amplifications), a salt-based DNA polymerase buffer, and nuclease-free water. PCR amplification has three main steps: 1.) Denaturation at 95-98°C to break the hydrogen bonds between the DNA strands, 2.) Annealing at 50-65°C (temperature dependent on the melting temperature of the oligos and the polymerase used) to bind the primers to the gene of interest, and 3.) Elongation at 72°C to allow the polymerase to stitch together the nucleotides and "proofread" the DNA strand. This process continues for 30-35 cycles with an initial denaturation and final elongation step.

3. PCR amplification was confirmed via DNA gel electrophoresis on agarose (0.7-1.0 wt%) gels. Successful amplicons were viewed under Ultraviolet light, and either the gel fragments were excised using a scalpel and purified by gel extraction (Qiagen Gel Extraction Kit) or the bulk PCR product was purified (Omega E.Z.N.A. Cycle Pure Kit).
4. Amplicons and their corresponding expression vectors were then double-digested with restriction enzymes to yield compatible "sticky ends" on each end of the amplicon and the plasmid. Restriction double-digest reactions include the DNA template (either amplicon or plasmid), the two restriction enzymes, restriction digest buffer (compatible with both restriction enzymes), and nuclease-free water. The reaction conditions vary based on the enzyme pair but most double-digests proceed at 37°C for 3 hours followed by enzyme inactivation at 65-80°C for 20 minutes to prevent random cutting (i.e. star (*) activity). In some cases, sequential

digestions are necessary if each enzyme requires a different buffer or reaction temperature. If plasmid DNA for use as reaction template is only available in TOP10 *E. coli* cryostock form, it can be generated using the Qiagen Mini-prep Kit.

5. Digestion was confirmed via DNA gel electrophoresis on agarose (0.7-1.0%) gels. Digested fragments were viewed under Ultraviolet light, and the gel bands were excised using a scalpel, and purified by gel extraction to remove unused template, oligos, and unwanted insert fragments. Alternatively, the vector can be treated with alkaline shrimp phosphatase (rSAP) post-digestion to remove the phosphate groups on the ends of the 5'→3' overhangs, and the entire PCR product can be purified without gel extraction (Omega E.Z.N.A. Cycle Pure Kit). In this project, both approaches were attempted.
6. Insert and backbone fragments were then combined (in a ratio between 1:1 and 9:1 by weight depending on the size of the pieces) in the presence of T4 DNA ligase, T4 DNA ligase buffer, and nuclease-free water. All ligation reactions were conducted overnight in an insulated 4°C ice-water slurry, after which time the ligase was heat inactivated at 65°C for 20 minutes.
7. Ligation was confirmed via electroshock or heat shock transformation of the ligation products into TOP10 *E. coli.* Upon shock, transformants were revived with nutrient-rich Super Optimal Catabolite-repressive (SOC) media and plated on Luria Broth (LB) media under appropriate selection pressure and allowed to grow at 37°C for 16-18 hours. Subsequently, formed colonies were purified on fresh LB plates to eliminate satellite colonies which do not contain the plasmid of interest. Satellite colonies form when the main colony releases beta-lactamase (BLA) into the

surrounding media, degrading the antibiotic resistance (particularly common with Ampicillin resistance).

8. Purified colonies were then screened by colony PCR which is essentially the same as traditional PCR with the exception that the template is the colony, and the initial denaturation time at 98ºC is increased to 10 minutes to ensure the cell walls are lysed allowing the oligos to access the DNA. (note this only works for *E. coli* template and failed to work for cyanobacteria and wild samples). Additionally, bracket primers are typically designed for use in colony PCR. These oligos amplify the gene of interest but also the region of plasmid flanking the gene to ensure that the gene is actually cloned into the plasmid.
9. Again, PCR amplicons were confirmed via DNA gel electrophoresis on agarose (0.7-1.0%) gels. Colonies that yielded amplification were subsequently grown up in liquid LB media under selection pressure, plasmid DNA was extracted, and sent for sequencing to ensure there were no mutations in the gene, promoter, or the surrounding backbone during the cloning process.
10. Sequence-verified clones were then cryopreserved in TOP10 *E. coli* or stored as plasmid DNA for subsequent applications.

Using this general procedure, CVN was cloned into the pAQ1 expression system with His-tags (denoted His(+)) and without His-tags (denoted His(-)). Although the process is similar for both constructs, there are slight deviations.

For the His(+) construct, the forward primer used to amplify CVN was designed with a thrombin protease cleavage site between the restriction site (NdeI) and the gene sequence. In the final construct, the thrombin site falls in between the His-tag and the gene of interest so the His-

tag can be cleaved enzymatically post-translation. For the His(-) construct, the forward primer was designed with FatI as the restriction site instead of NcoI (the restriction site present in the backbone). This substitution is acceptable because FatI will not splice the gene and NcoI and FatI have compatible 5'→3' sticky ends when digested and desirable because the FatI (CATG) restriction site ends with a start codon while NcoI (CCATGG) would require insertion of two extra base pairs to amplify the gene of interest in frame. The reverse primer (utilizing the BamHI restriction site) and the primer set annealing temperature (63^{o}C) is the same for both constructs.

These amplicons were then double digested with the corresponding enzymes. However, FatI and BamHI do not have compatible digestion buffers or reaction temperatures. To circumvent this issue, the His(-) insert was sequentially digested: first with BamHI-HF at 37^{o}C for three hours in NEBuffer 2.1 followed by heat inactivation at 65^{o}C to prevent enzymatic star(*) activity, then with FatI at 55^{o}C for an additional three hours followed by heat inactivation at 80^{o}C to prevent enzymatic star(*) activity. To avoid the need for changing buffers, high-fidelity BamHI was used and the incubation time was increased to three hours. The His(+) amplicons were double digested with NdeI and BamHI at 37^{o}C in CutSmart Buffer for three hours. To generate backbone for each construct, pAQ1-Ex-YFP was double digested in the same way as each of its corresponding inserts with the exception that FatI, which is not a single cutter in pAQ1, was replaced with NcoI. This vector is a particularly useful source of backbone DNA because in the instance where it either fails to completely digest or re-ligates on itself post-digestion, these colonies can be easily identified as false positives by their fluorescence under blue light. After gel purification of the digestion fragments, an insert to vector ratio of 7:1 was used for both constructs to maximize the probability of successful ligation because the CVN gene is relatively small in comparison to the backbone (306 bp compared to 5.5 kb).

All ligation reactions were transformed into TOP10 *E. coli* via electroporation, revived with SOC medium, and plated on LB+Amp100 (notation for LB media supplemented with Ampicillin antibiotic to a final concentration of 100 µg/ml). After colony purification, each colony was screened via PCR using a gene-specific forward primer and a backbone-specific reverse primer (T7 Reverse) – referred to colloquially in the lab as "bracket primers" to avoid false positives that result from using internal fragment primers only. Positive colonies were sequenced to screen for gene mutations prior to cryopreservation.

3.4.2 Fusion PCR

Fusion PCR is the process of fusing pieces of DNA together through Polymerase Chain Reaction. This technique is particularly useful when the gene of interest needs to be amplified from a eukaryotic organism but the expression system is prokaryotic (e.g. *E. coli*). The active portions of genes in eukaryotic systems (i.e. exons) are typically split by one or more inactive portions (i.e. introns) that can be differentiated by eukaryotic expression systems. Prokaryotes, however, cannot make this differentiation and will express the entire gene "as is." Thus, for expression in *E. coli,* introns must be spliced out, and fusion PCR is the most common technique utilized[27].

The easiest way to accomplish this is to first consider the entire gene (introns and all), and design forward and reverse primers with the appropriate restriction sites to yield compatible sticky ends within the MCS of the desired expression vector. Then design primers to amplify each exon individually with a 5'→3', 10-12 bp overhang that complements the end of the exon preceding it (forward fusion primer) or the beginning of the exon succeeding it (reverse fusion primer) (see Figure 5). Each exon can then be amplified individually and purified by gel extraction. These

exons are then used as the template for the final PCR reaction using the outer primers designed for the entire gene, and the polymerase can stitch each exon together advantaging the overlap between adjacent exons. Typically, high fidelity (HF) polymerases (e.g. Phusion or Q5) are used for the final PCR reaction because they have "proofreading capabilities" and can limit the number of Single Nucleotide Polymorphisms (SNP's) in the final fusion product.

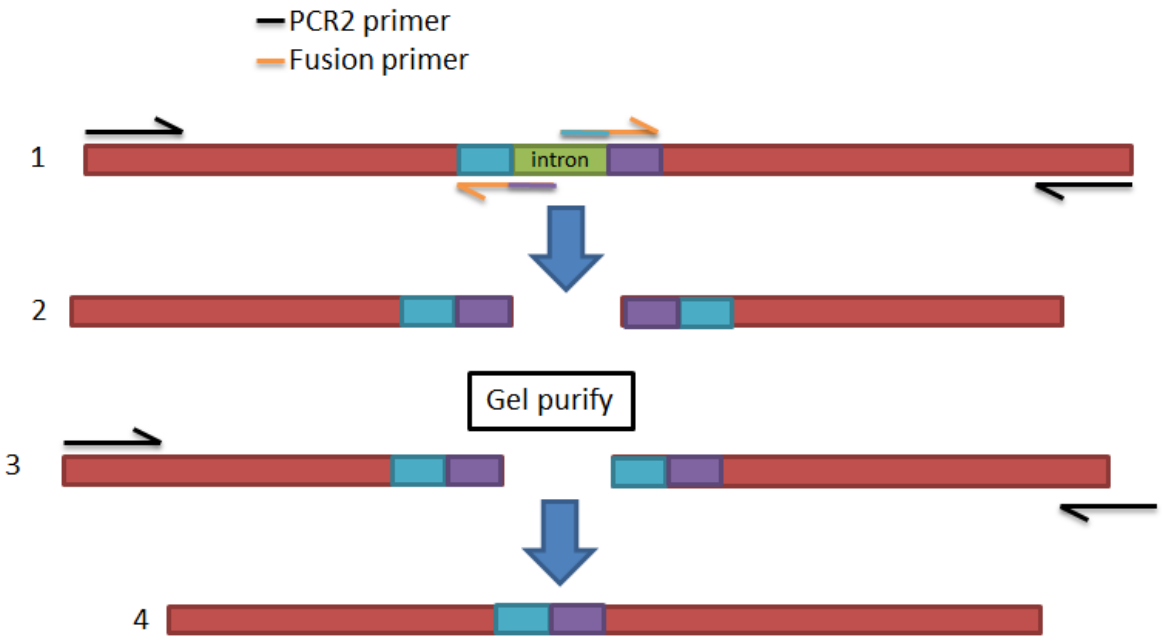

Figure 5. Schematic of Fusion PCR

Fusion PCR was explored in a particularly novel way in this project. During the process of transforming the cyanovirin-containing expression vector, it was hypothesized that the expression of the CVN protein was killing the host organism (discussed in detail in Chapter 4). To alleviate this issue and provide a sustainable platform for large-scale purification of the CVN protein, primers were designed to amplify the cyanovirin gene from its host plasmid, pET26B-CVN, and the elastin-like peptide and intein (ELP+INT) tag from its host plasmid, pET/ELP-I-CAT, and fuse them together, treating them like two exons. The inclusion of this tag in the expression system will

minimize cytotoxicity in the host organism possibly caused by the carbohydrate binding domain aspects of this protein that could interact broadly within the cell.

3.5 Protein Expression in Cyanobacteria

Cyanobacteria are viable hosts for protein expression because they are "naturally competent," essentially meaning they are capable of accepting foreign DNA from the surrounding environment[28]. The expression vector utilized for transformation into *Synechococcus* sp. PCC 7002 was pAQ1-Ex-YFP, constructed by Erik Curtis, a recent alum of CurtisLab who adapted this protein expression technique to our lab. This construct is a derivative of pAQ1, a high copy number plasmid native to *Syn.* sp 7002. The plasmid contains a constitutive CPC (c-phycocyanin) promoter that expresses genes encoded in the MCS in cyanobacteria and kanamycin and spectinomycin resistant genes for plasmid proliferation and expression under selection pressure in *E. coli* and cyanobacteria, respectively. This entire region is flanked by sequences that are homologous to the native pAQ1 plasmid found in WT cyanobacteria, allowing the host organism to accept the linearized plasmid into the chromosome[26, 28].

Transformation of constructs into cyanobacteria is accomplished by linearizing the plasmid either by digestion with one or both of the single-cutting restriction enzymes neighboring the flanks or via PCR amplification of the promoter region in between the flanks. This linearized DNA is then combined with ~3 mL of exponentially growing *Syn*. 7002 without antibiotic selection pressure to avoid killing the WT host. The inoculum is then allowed to grow for 1-2 days under bright light and CO_2 supplementation and subsequently plated on A+ medium under appropriate selection pressure (spectinomycin 50 in this instance). The plates are first stored under

low light to avoid photo-bleaching and slowly transitioned to bright light and CO_2 supplementation until single colonies appear (typically 1-2 weeks from harvest). These colonies can then be screened via PCR using bracket primers to confirm successful homologous recombination into the host organism[26,28].

3.6 Protein Purification Techniques

With the current technology available, proteins expressed in CurtisLab have typically been cloned with an N-terminus polyhistidine tag and purified via nickel affinity chromatography resin[29,30]. Noting the labor-intensive procedure, expensive materials, and the difficulty of scalability, the transition to other purification platforms has become increasingly desirable.

Among the alternative protein purification methods being considered is solubility tag purification. A method has been developed such that the gene of interest can be cloned into an expression system with an N-terminus aggregation tag composed of an Elastin-like-Peptide (ELP) solubility tag followed by an intein (INT) spacer. Upon expression of the entire protein (ELP+INT+recombinant protein), the ELP tag can be self-aggregated at low temperatures, and the pH of the lysate can be altered to facilitate self-cleavage of the ELP tag via the INT spacer. Subsequent self-cleavage of the INT spacer can yield purified target protein for further application[31]. This method allows for scale-up of recombinant protein purification without chromatography columns and resins.

Incorporation of this protein purification technique into a cyanobacteria expression platform requires manipulation of the expression vectors because the aggregation tag was developed as a construct driven by a T7 leaky promoter. To avoid expression in *E. coli*, primers

were designed to amplify the ELP+INT tag and fuse it to the anti-HIV genes (CVN and MVN). Subsequently, the fusion product can then be cloned into the pAQ1 expression system and transformed into *Synechococcus*, advantaging the strength and versatility of the CPC promoter and the scalability of the ELP+INT purification system.

3.7 Procurement and Revival of Wild (Potential) *M. aeruginosa* Samples

In late summer of 2017, the principal investigator on this project, Dr. Wayne Curtis, travelled to Northeast Pennsylvania in search of algal blooms that might contain *M. aeruginosa*. He procured water samples from three nearby bodies including Lake Wallenpaupack, Porter's Lake, and Decker's Pond, which are all freshwater bodies but each of which has a varying degree of human influence (Figure 6). *M. aeruginosa* is carried to the surface by gas vesicles in the presence of sunlight; therefore blooms were located by searching in the late afternoon and following the direction of the wind into nearby alcoves. Samples were stored in conical tubes and plastic jars and transported back to State College for processing on September 26, 2017. Unfortunately, these samples were not transferred to fresh media until three weeks later, which likely stunted proliferation of *M. aeruginosa*.

Figure 6. Bloom of Algae Blown Towards Leeward Lakeshore of Lake Wallenpaupack as Potential Source of *Microcystis aeruginosa*

Approximately 2.5 mL of each wild sample was sub-cultured into separate 22.5 mL aliquots of Wayne's Freshwater Algae Media for Chlorella (WFAMC) and HEPES-modified BG-11 medium to compare the impact of media selection on growth and maintenance. The subcultures were stored in a Percival Model E41L2 at 25ºC and 2% carbon dioxide supplementation. Each sample was sub-cultured every seven days for two months into fresh media to maintain exponentially growing cultures to obtain monocultures. In early December, these samples were moved out of a carbon dioxide environment to slow growth and were sub-cultured every three weeks into fresh media to maintain the stock.

3.8 Maintenance and Preservation of Cyanobacteria

During regular use, the transformation host organism, *Synechococcus* sp. PCC 7002, was maintained in liquid A+ medium under low light, shaking at 200 RPM. For short-term storage, liquid culture was subbed on A+ medium slants in pre-sterilized 15 mL Falcon tubes and stored under low light. For long-term storage, cyanobacteria liquid culture can be cryogenically preserved according to the procedure developed by Brand[32].

3.9 Achieving Axenic WT Microcystis strains

The WT lake samples were maintained in liquid HEPES-modified BG-11 for about a month to ensure the organisms were consistently growing well. Approximately 50 µL of each sample was then spread on HEPES-modified BG-11 plates and allowed to grow for one week under bright light in a carbon dioxide-supplemented environment. Significant overgrowth was noticed after this first week so each sample was dilution-streaked on fresh HEPES-modified BG-

11 plates. After another week of growth under the same conditions, at least three organisms were distinguishable on each plate: a fibrous fungus, white bacterial colonies, and blue-green algae colonies. About 5-6 isolated, green colonies were chosen from each plate (totaling 34 colonies) and individually subbed into 2 mL of fresh HEPES-modified BG-11. These colonies were allowed to grow under the same conditions and harvested during their exponential growth phase (approximately 72 hours later). Each colony was dilution streaked onto a fresh HEPES-modified BG-11 plate and allowed to grow under the same conditions for 1-2 weeks. To confirm whether or not the cultures were axenic, single colonies from each of these plates were inoculated in 3 mL of fresh HEPES-modified BG-11 and allowed to grow for 1-2 weeks under low light, shaking at 200 RPM. These cultures were spread onto R2A plates, allowed to grow for 1-2 weeks under bright light and CO_2 supplementation, and examined for bacterial contamination. Single colonies from contaminated plates were re-inoculated in fresh HEPES-modified BG-11 and again spread on R2A plates where lack of bacterial growth would suggest axenic. Cultures were then visualized under a Motic 2.0 Digital Microscope to assess whether or not they were monocultures.

Chapter 4

Results and Discussion

4.1 Assembly of pAQ1-CVN

Summary

This section describes the successful cloning of the cyanovirin (CVN) gene into pAQ1, a homologous recombination vector for cyanobacteria transformation.

Details

PCR amplification of the cyanovirin gene from pET26B-CVN using standard Taq polymerase was successful as indicated by Figure 7. Due to the thrombin site addition in the His(+) amplicons, these bands (lanes 1-3) are slightly larger than the His(-) amplicons (lanes 4-6). To minimize the loss of the CVN amplicons due to the inefficiencies of gel extraction, the bulk PCR products were purified and used as template for enzymatic digestion as described in Chapter 3.

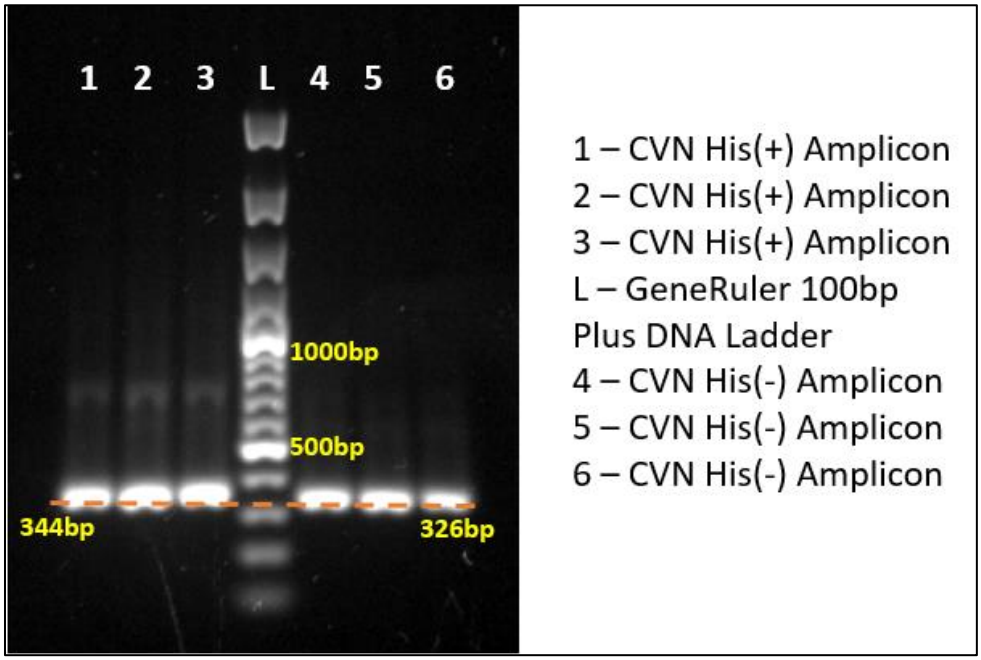

Figure 7. PCR Amplification of CVN from pET26B-CVN

The pAQ1-YFP vector was also double digested with corresponding enzymes that yield compatible sticky ends with each desired insert (also described in Chapter 3). However, electroshock transformation of the subsequent ligations reactions yielded no colonies when plated on LB+Amp100. The absence of false positives (i.e. re-ligated backbone) was suggestive of transformation failure. In order to eliminate the other possibility of failed ligation, the ligation reactions were repeated and run on an agarose gel. Typically, ligation reactions are run in small volumes (~10 µL) and produce small amounts of the desired construct. In order to alleviate this issue, the ligation reactions were scaled up ten-fold to be able to visualize the DNA. The corresponding gel can be seen below in Figure 8.

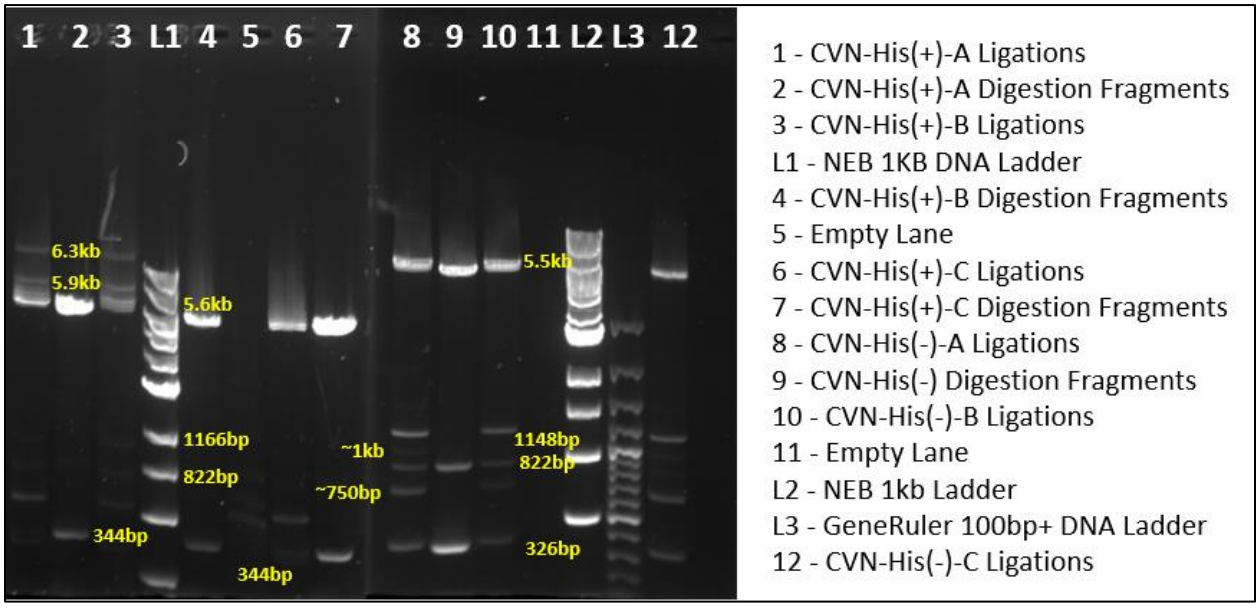

Figure 8. pAQ1-CVN Ligation Confirmation

To note, the letters (A, B, and C) merely indicate technical replicates of the same exact procedure (i.e. inserts that were amplified in the same way but were individually purified and processed). Lanes 2, 4, 7, and 9 show digested pAQ1 vector and CVN insert fragments (again, slight differences in size are indicative of the amplicons with and without the thrombin site). Smears in lanes 1 and 3 (attempted His (+) ligation) are suggestive of successful ligation because

of the likely presence of both linearized and circular plasmid DNA. A closer look seems to show ~3 bands: the top band (~6.3 kb) is suggestive of non-digested pAQ1-YFP vector, the middle band is suggestive of successful ligation (~5.9 kb), and the bottom band is suggestive of digested, but un-ligated pAQ1 vector (~5.6 kb). Lane 6 (also attempted His(+) ligation) did not appear to be obviously successful, but there is always the possibility that ligated product is in concentrations too small to visualize on an agarose gel. Lower down on the gel, digested, un-ligated amplicons with thrombin sites are visible around 350 bp. Comparing lanes 1, 3, and 6 versus 2, 4, and 7, the predominance of this band correlates to a lower likelihood of successful ligation. The attempted His(-) ligations (lanes 8, 10, and 12) show large bands nearly identical to that of the digestion fragments (lane 9), but no visible ligation product is apparent. The most striking band on this gel is the band at ~1150 bp (much more significant in the His(-) ligation reactions). Interestingly, this band indicates ligation between the inserts (CVN and YFP). The vector digestion reactions were treated with Alkaline Shrimp Phosphatase (rSAP) to dephosphorylate the backbone sticky ends and prevent pAQ1-YFP re-ligation. However, given that the digested YFP fragments had not been removed prior to ligation, the cyanovirin gene was able to ligate to YFP because they shared complementary sticky ends. This ligation is also more energetically favorable given their proximity in size compared to the backbone. This 1150 bp band is visible on the left (His(+) amplicon) gel as well, but significantly fainter. The left gel also shows significantly less YFP (~800 bp) altogether which corresponds to the successful pAQ1-CVN His(+) ligations noted in lanes 1 and 3. It is worth noting that it is uncertain what the bands at ~750 bp and ~1000 bp represent (they are much more prominent on the right gel than the left). It was considered that they could be the result of digested pET26B-CVN template leftover from the PCR reaction but this is unlikely because NcoI does not normally cut anywhere in this plasmid.

Overall, this gel is suggestive that utilization of the E.Z.N.A Cycle Pure kit to purify PCR and digestion reactions simply results in too many potential, more energetically favorable side reactions/ligations. However, if this kit is preferred, it is worth considering a subsequent PCR after purification to further increase amplicon concentration and dilute other contaminants. The CVN PCR amplifications and subsequent digestions were repeated utilizing gel extraction as the preferred DNA isolation technique despite the notably lower yields. To reduce contamination and increase the final DNA concentration in this step, excess agarose was stripped from the excised fragments. Ligation and electroshock transformation were re-attempted in the same way as discussed in Chapter 3.

As a positive transformation control, 1 µL of pAQ1-Ex-YFP pure plasmid was transformed into the TOP10 electrocompetent *E. coli* and plated on LB+Amp100. Transformation yielded 25 pAQ1-CVN His(+) colonies and 59 pAQ1-CVN His(-) colonies. Of these 84 total colonies, 15 pAQ1-CVN His(+) colonies and 25 pAQ1-CVN His(-) colonies were streaked onto fresh LB+Amp100 plates to eliminate contamination from the background plasmid on the original ligation plate. More His(-) colonies than His(+) colonies were selected as a precaution given that this ligation had likely failed previously. If no positive colonies were found, more could be colony purified from the original transformation plate. The positive control plate yielded hundreds of colonies; a few were selected for colony purification. One of the resultant plates (viewed under blue light) is pictured below as confirmation of successful transformation (Figure 9).

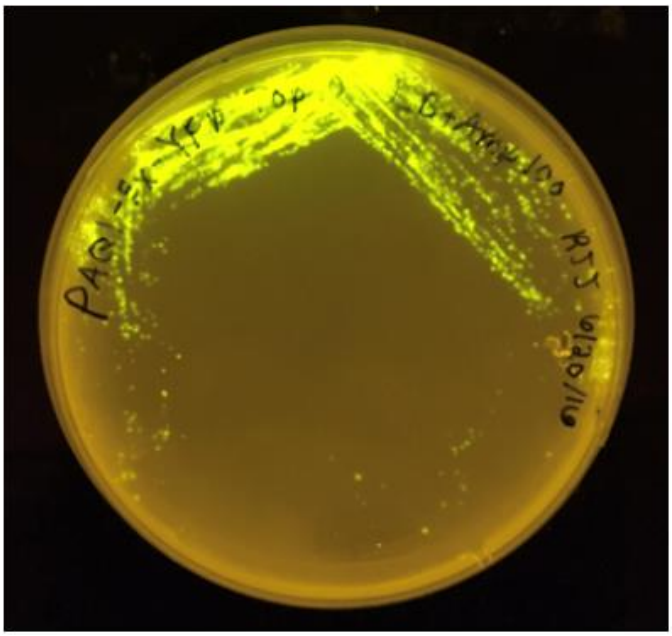

Figure 9. pAQ1-YFP Electroshock Transformation Positive Control

Each pAQ1-CVN colony was subsequently used as template for colony PCR using the corresponding gene-specific forward primers and backbone-specific reverse primers to confirm successful ligation into the vector and to confirm that the CPC promoter region was not mutated in the cloning process. The colony PCR results are pictured below in Figure 10.

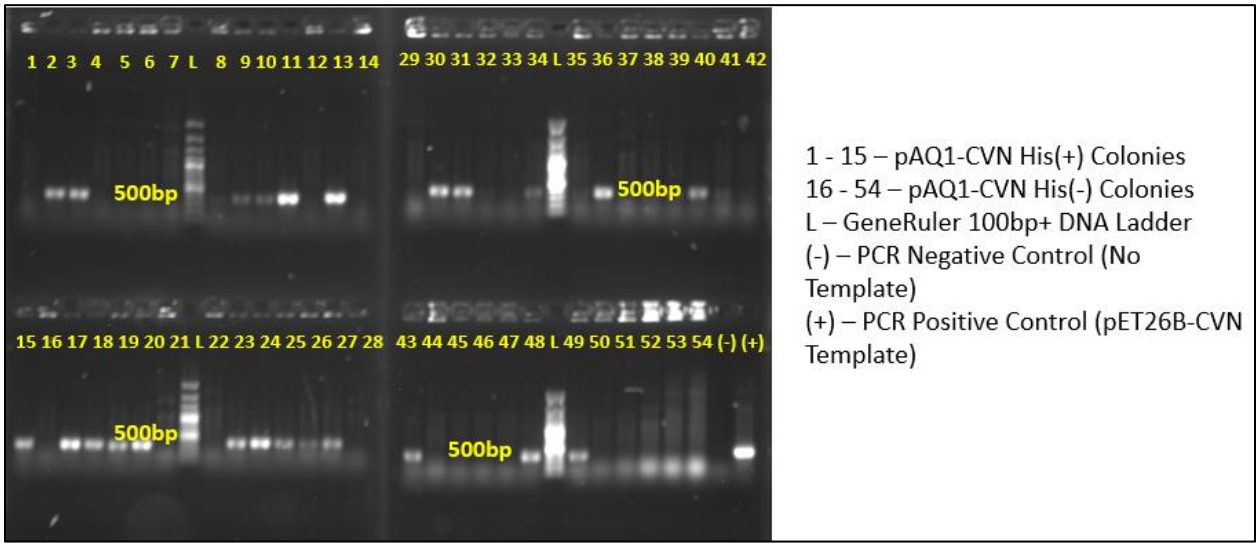

Figure 10. pAQ1-CVN Colony PCR Confirmation

Despite poor ladder separation in the right gel, there was successful amplification in lanes 2, 3, 9, 10, 11, 13, 15, 17, 18, 19, 20, 23, 24, 25, 26, 27, 30, 31, 34, 36, 40, 43, 48, and 49. Of these positive colonies, three of the most promising colonies for each construct were chosen for plasmid proliferation and sent to the Penn State sequencing center to verify the genes were in frame and not mutated.

Of the six colonies that were sent for sequencing, the promoter and gene regions aligned perfectly for four of them (pAQ1-CVN His(+) colonies 15 and 22 and pAQ1-CVN His(-) colonies 4 and 32). Utilizing the Benchling sequence alignment tool, it was determined that pAQ1-CVN His(-) colony 1 was missing six base pairs at the beginning of the CVN gene, including the FatI restriction site (Figure 11). It is likely not a coincidence that this deletion occurs right at the cut site, suggesting that the 5' sticky ends did not successfully ligate. However, linearized plasmid is typically degraded by nucleases in *E. coli*. Although this topic is not widely studied, there is the possibility that the linearized DNA was taken up by the cells via homologous recombination if any region of the plasmid aligned with the *E. coli* genome, but there is no way to confirm this. Nonetheless, this plasmid and cryostock were discarded as a mutant.

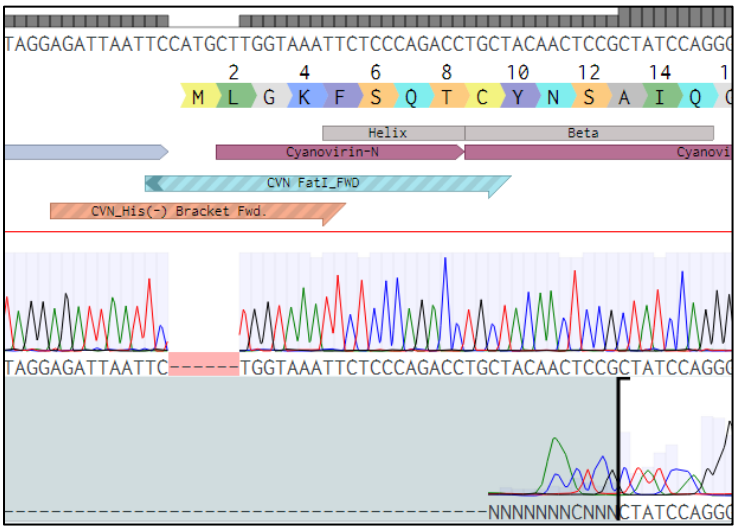

Figure 11. pAQ1-CVN His(-) Colony 1 Mutation

The other mutation was exhibited by pAQ1-CVN His(+) colony 24. However, in this instance, the mutation is silent at the 100th codon in the cyanovirin coding sequence. The adenine base pair in the template target sequence was replaced with a guanine, both of which translate to glutamic acid in this codon (as pictured in Figure 12). No further action was required, and this plasmid was acceptable to use for downstream applications.

Figure 12. pAQ1-CVN His(+) Colony 24 Silent Mutation

4.2 Transformation of pAQ1-CVN

Summary

This section describes the unsuccessful attempts to transform *Synechococcus* 7002 with pAQ1-CVN due to apparent toxicity. Fusion of this gene to an aggregation tag is proposed as a possible solution for toxicity mitigation.

Details

Sequence-verified pAQ1-CVN His(+) and His(-) plasmids were generated from *E. coli* cryostocks and used as templates for independent digestion reactions. The plasmids were linearized using SphI which cuts directly after the spectinomycin-adjacent pAQ1 flank. Alternatively, NsiI, which cuts directly before the CPC promoter-adjacent pAQ1 flank, can be used as the restriction enzyme. Either one effectively linearizes the DNA along a flank that can perform homologous recombination into the chromosome of WT *Synechococcus*.

This linearized DNA was inoculated in 3 mL of exponentially-growing WT *Syn.* 7002 and allowed to grow for 1-2 days under 2% carbon dioxide supplementation (See Appendix F for the cyanobacteria bubbler rig setup). As a positive control, pAQ1-Ex-YFP was linearized and inoculated in the same way. As a negative control, WT *Syn* 7002 with no DNA additives was observed under the same growth conditions. The cyanobacteria was harvested and plated according to the procedure discussed in Chapter 3.

The positive and negative controls similarly exhibited reasonable growth (OD_{730} ~0.5-1) after two days while the His(+) and His(-) construct inocula exhibited little to no growth (insignificant change in OD_{730}). As expected, the plated His(+) and His(-) inocula did not yield any colonies. The positive control plate (pictured in Figure 13) shows successful transformation, indicated by the cyanobacteria glowing a bright yellow-green under blue light.

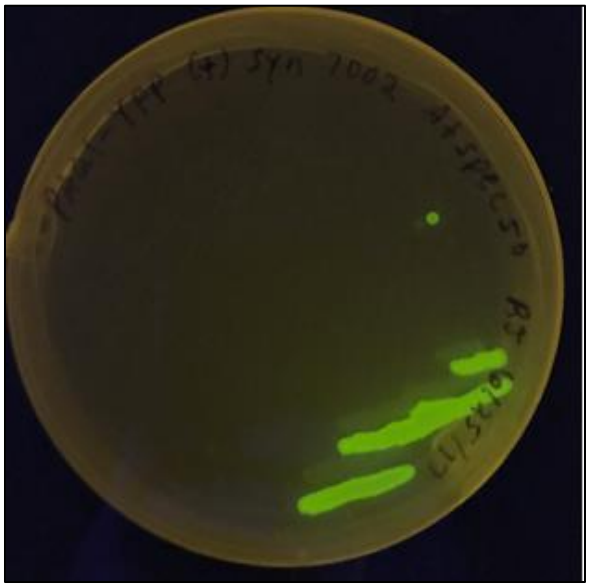

Figure 13. pAQ1-YFP Syn. 7002 Transformation Positive Control

This transformation was repeated under the same conditions and yielded a similar result (no growth for the desired CVN constructs). This is suggestive that as cyanovirin is expressed by the constitutive CPC promoter, it releases toxins into the cytoplasm, preventing cellular replication. Therefore, the organism dies shortly after inoculation. Further replicates would be required to confirm this observation. Assuming this theory is plausible, in order to develop a large-scale purification platform for cyanovirin protein, this cytotoxic property must be subdued. Around the time of this discovery, other CurtisLab protein purification projects were centering around the application of an Elastin-Like Peptide (ELP) and Intein (INT) aggregation tag which precipitates the entire protein complex at low temperature and self-cleaves under controlled changes in pH, leaving behind the desired recombinant protein. This sparked the innovative idea to fuse the cyanovirin protein to this ELP+INT tag that would ultimately alter the size, shape, and folding nature of the protein (hopefully reducing the concentration of toxins emitted into the cytoplasm) while also providing a sustainable and scalable platform for protein purification.

4.3 CVN/ELP Fusion Cloning

Summary

In this section, the attempt to create the self-cleaving solubility tag fusion to CVN is described. Although this was not successful, a strategy to achieve this by PCR outside the repeat region of ELP is presented and fully discussed in Chapter 5.

Details

The cloning strategy to produce the desired construct (pAQ1-INT+ELP+CVN) requires the same pAQ1 backbone containing the CPC promoter. However, the insert needs to be pieced together from two different source plasmids. From Table 1 in Appendix C, primers CVN_INT_FUS_Fwd. and CVN_BamHI_Rev. were used to amplify the CVN gene from the pET26B-CVN vector with an ~10 bp, N-terminus overlap with the end of the Intein sequence. Similarly, primers ELP_NcoI_Fwd. and CVN_INT_FUS_Rev. were used to amplify the ELP+INT gene combination from the pET/ELP-I-CAT vector with an ~10 bp, C-terminus overlap with the beginning of the CVN gene. A third PCR can be performed using primers ELP_NcoI_Fwd. and CVN_BamHI_Rev. to fuse the two pieces together, advantaging the adjacent overlap between genes. This ELP+INT+CVN insert can then be digested and ligated into the pAQ1 backbone to yield the desired construct (All target constructs can be seen in Appendix B).

A first attempt at amplifying the individual pieces is shown in Figure 14.

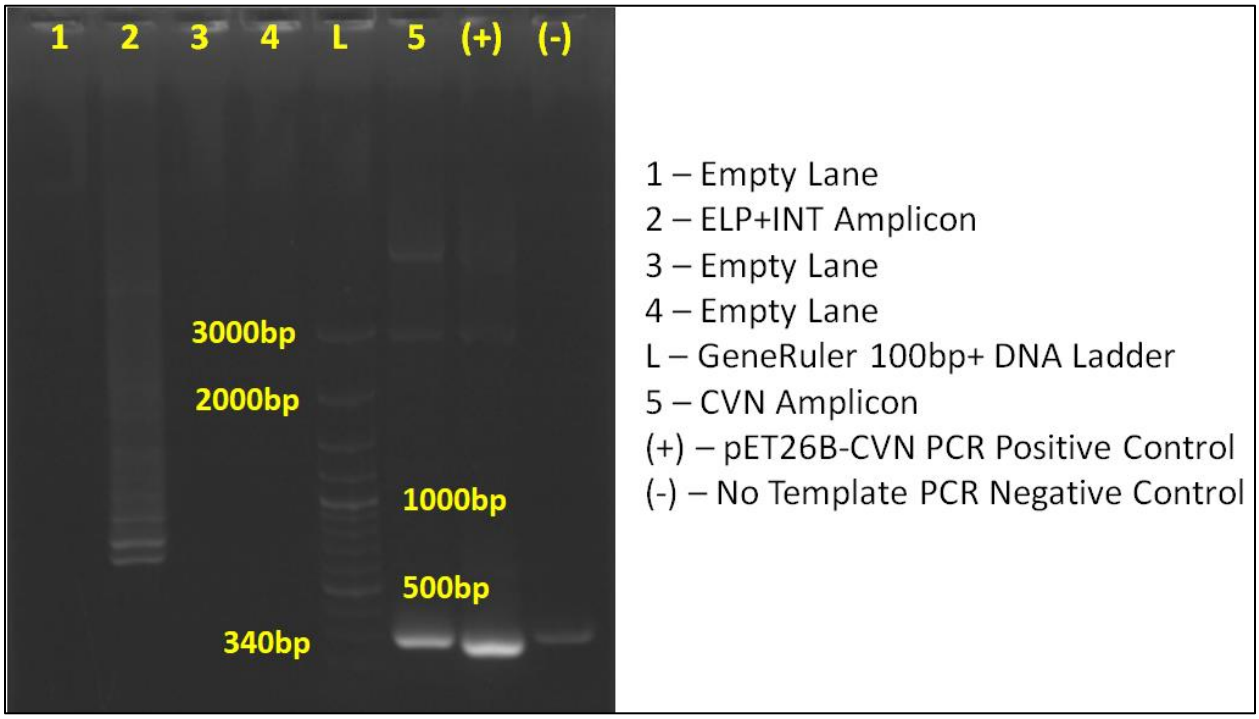

Figure 14. ELP+INT and CVN PCR Amplification Attempt

Amplification of the CVN gene (lane 5) succeeded without issue while the ELP+INT amplification (desired size = 2.2 kb; lane 2) is a smear with decipherable bands at ~700, 800, and 900 bp. This is likely due to the repetitive nature of the ELP DNA sequence. Although there are ~10 unique base pairs at the beginning of the ELP sequence in the host plasmid, the last five base pairs of the ELP_NcoI_Fwd. primer overlap with 32 other regions in the ELP gene. To circumvent this issue, a novel forward primer (denoted ELP_NcoI_Fwd._Revised) was designed such that the restriction site (in this case NcoI) was near the 3' end of the oligo. The first 23 base pairs of the primer align with the region of the pET/ELP-I-CAT backbone eight base pairs removed from the ELP gene. The next six base pairs are the restriction site (NcoI), and the last two are essential for ensuring the entire gene is translated in frame (in this case "GA" was inserted to add a dormant glycine amino acid to the N-terminus of the protein upon translation). When this insert is digested, the base pairs upstream of the cut site will be cleaved to leave the desired ELP+INT+CVN insert.

A first attempt at this amplification yielded no band whatsoever. Presumably, any number of reasons could have resulted in this failure (e.g. primer dimers, sub-quality DNA template, non-specific binding, etc.), but the methodology is sound. At this time, the cloning is still being troubleshot.

All other development of this project was then focused on isolating the microvirin gene from the Northeastern Pennsylvania lake samples. When the MVN gene is isolated, the fusion cloning discussed in this chapter may be worthwhile to pursue given that both CVN and MVN similarly emit cytotoxins upon expression.

4.4 MVN Microscopic Imaging

Summary

In this section, there are a lot of observations of wild "algae" samples aimed at isolating *M. aeruginosa* monocultures and eventually axenic cultures for cloning of MVN. No monocultures have been achieved yet, but it is suggestable that maintenance on various media types significantly alters the population of the culture, decreasing the chance of isolating *M. aeruginosa*.

Details

During the process of maintaining the wild lake samples, a couple of physical observations were made that helped confirm or deny the possible presence of *M. aeruginosa*. After storage under bright light for several weeks, the dense Lake Wallenpaupack jug sample showed translucent, blue-green algal blooms on the surface of the water. The jug was then moved into complete darkness for a week, and the surface of the water was noticeable different. The floating

blooms had now become clustered masses that either sunk to the bottom or just below the surface (Figure 15).

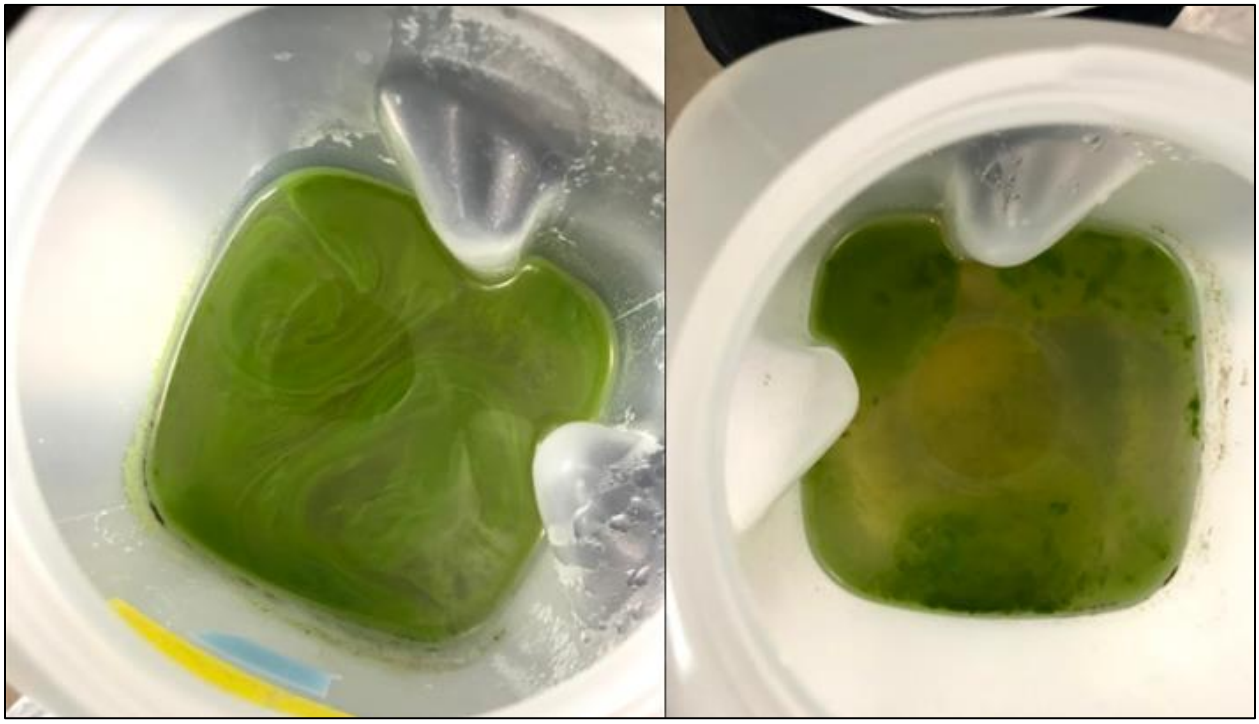

Figure 15. Lake Wallenpaupack Jug Sample Before (Left) and After (Right) Exposure to Darkness

This observation aligns with the presence of *Microcystis aeruginosa* given that this organism is carried to the surface by gas vesicles at peak sunlight hours and sinks when the sun sets. However, this is not conclusive proof that the desired organism is present because surely there are other organisms who exhibit similar behavior.

In contrast, one of the Decker's Pond plates exhibited significant overgrowth and bubbling when streaked on HEPES-modified BG-11 medium and stored under carbon dioxide supplementation (Figure 16). It is likely that exposure to carbon dioxide resulted in photorespiration whereby the organism fixes the oxygen from the light source rather than the carbon dioxide and actually releases already fixed carbon dioxide into the environment, thus

resulting in air bubbles. Photorespiration is equally common in eukaryotic microalgae and cyanobacteria, but this observation differentiates this culture from the others.

Figure 16. Decker's Pond Eukaryotic Samples of HEPES-modified BG-11 Medium

To obtain a more comprehensive understanding of the composition of the lake samples, each of the nine lake samples and the dense Wallenpaupack jug sample were visualized under a Motic 2.0 digital microscope at a total magnification of 400X. A hemocytometer was used to estimate the diameter of the cells in solution. The juxtaposition of these pictures is shown in Figure 17.

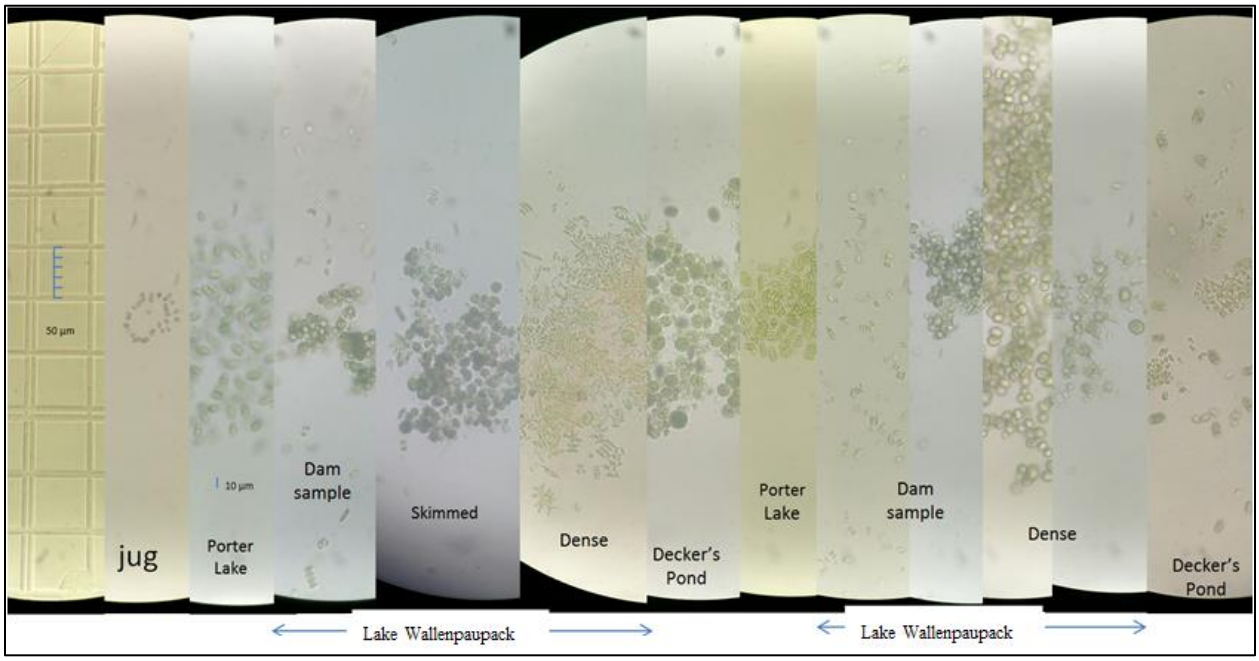

Figure 17. Motic Digital Microscope Images of WT Freshwater Samples (400X)

Most of the above cultures have multiple, distinguishable organisms present, particularly the Dense Wallenpaupack and Decker's pond samples. Interestingly enough, the first Decker's pond sample (corresponding to the plate pictured in Figure 16) is clearly eukaryotic as anticipated, but the second Decker's pond sample is almost entirely small colony algae (possibly containing *M. aeruginosa*). The only difference between these two cultures is that the second sample was cultured only on HEPES-modified BG-11, while the first sample started off on WFAMC for three subcultures. These first few weeks on WFAMC may have been selecting for eukaryotic microalgae rather than *Microcystis*, and by the time the switch the HEPES-modified BG-11 was made, it was too late to recover. This suggests that media choice may play a large role in obtaining monocultures and axenic cultures of *Microcystis aeruginosa*. Both of the Dense Wallenpaupack samples show an incredible diversity of phytoplankton (not shocking given the overwhelming amount of activity in this region of the lake). The skimmed sample looks like a mix of the Dense Paupack and Decker's pond samples, showing evidence of both large, eukaryotic microalgae and cyanobacteria.

The jug sample invariably looks like the Dense Wallenpaupack samples because they are essentially from the same population (only difference is the jug sample has been continuously proliferating in lake water. Both of the Lake Wallenpaupack ("dam") samples exhibit similar composition with a notable mix of microalgae stacked in quads and clusters that appear to be colony cyanobacteria. Perhaps the most promising sample in the lot is the first Porter's Lake sample with dispersed, circular, colony algae. Oddly enough, the second Porter's Lake sample looks very dissimilar to the first with a diversity of plankton and quad stacks, suggesting that the effects of media selection are more complex than originally expected.

To summarize, the only samples that do not look promising for isolation of *M. aeruginosa* are Decker's Pond 1 and Porter's Lake 2. All other cultures exhibit clusters of small colony algae with an approximate cell diameter of 5 um (consistent with literature values of 3-5 μm for *M. aeruginosa*)[33]. These other cultures should continue to be maintained on HEPES-modified BG-11 to select for cyanobacteria. The imminent goal on this front is to get a cyanobacteria monoculture first and then focus on isolating axenic strains via dilution streaks on R2A (noting that many algae cultures are monocultures but not axenic). Then, once a colony PCR technique has proved successful in cyanobacteria, these samples should be used to attempt to isolate the microvirin gene for cloning into the pAQ1 expression system.

4.5 Troubleshooting Cyanobacteria Colony PCR

Summary

This section attempts to demonstrate isolation of *M. aeruginosa* monoculture (or even presence in mixed culture) through PCR. Several methods for PCR amplification direct from

culture failed, including the *Syn.* 7002 positive control; thus prompting further development of the technique. Ultimately, genomic DNA extraction prior to PCR was adopted as the strategy moving forward.

Details

The difficulties of working with WT samples are numerous but among the most relevant are selecting for the appropriate organism (assuming it is even present), confirming and isolating that organism, and cloning from that organism. There is ample documentation and congruency on conducting colony PCR from *E. coli*, but the techniques used for colony PCR from cyanobacteria are numerous and result-variable depending on the species and strain.[34] Several techniques to amplify the MVN gene from a handful of dilution streaked, lake sample colonies were attempted to no avail.

As simply an adaptation of *E. coli* colony PCR, the cyanobacteria colony was tapped with a pipette tip, used as template for a PCR reaction with MVN-specific primers, and denatured for 30 minutes at 98°C prior to the first PCR cycle.

A similar method was attempted whereby the colony was tapped and diluted in 10 µL of milli-Q water and pre-boiled for 30 minutes at 98°C. Subsequently, 1 µL of this diluted colony was then used as template for PCR with MVN-specific primers.

Both of the resulting gels turned up empty. It was then hypothesized that the cell membrane of the cyanobacteria was not properly lysing due to inherent complexities not present in *E. coli*, and thus there is no DNA in solution for the primers to anneal to.[35] To test this, the above approaches were attempted again using both MVN-specific and degenerate 16s primers.[36] As a positive control, *Syn.* 7002 was treated identically to the MVN colonies and used as template for a PCR using degenerate 16s primers. Given that the *Syn.* 7002 in the lab was an axenic, prokaryotic

organism, the 16s rDNA gene should amplify assuming the cells are lysing adequately. This technique also yielded no amplification, suggesting that stronger lysing methodologies might be needed.

Exponentially-growing MVN and *Syn.* 7002 culture was combined with a formulated PCR detergent (1% v/v Triton X-100, 5mM DTT, 500mM Tris-HCl) in a ratio of nine parts culture to one part detergent. The cells were well mixed and irradiated in a tabletop microwave at 700W for 80 seconds.[37] The disrupted cells were then used as template for PCR using degenerate 16s primers and MVN-specific primers. Again, no amplification was achieved.

A handful of troubleshooting measures was discussed including primer redesign, sonication of the culture using a mini horn tip, and genomic DNA prep prior to PCR. Re-evaluation of the MVN-specific primers proved unhelpful because there were no stable secondary structures in the PCR reaction temperature range, and sonication seemed extreme and could possibly shear the DNA if not executed carefully, leaving the next viable option as genomic DNA extraction. Phenol-chloroform DNA extraction from Syn. 7002 was performed using TRIzol reagent to yield ~3 µg of genomic DNA (gDNA). This gDNA (~150 ng) was then used as the template for PCR using degenerate 16s rDNA primers.

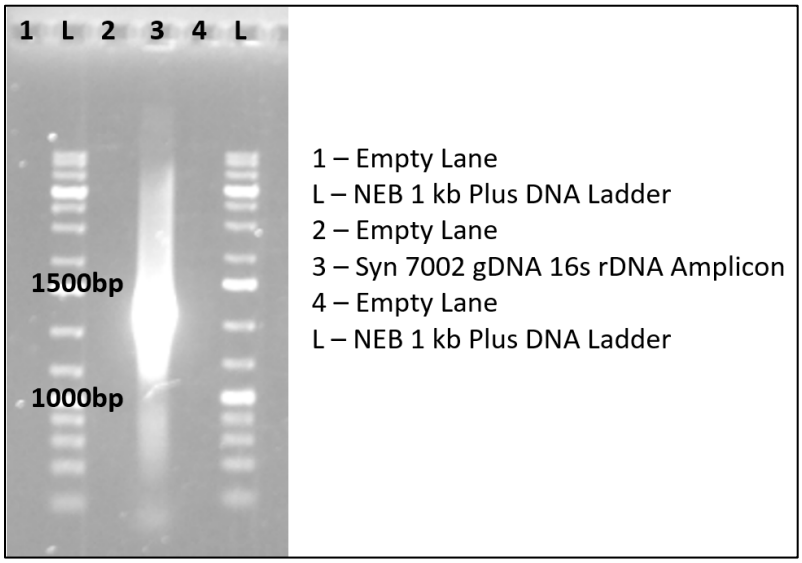

Figure 18. *Syn.* 7002 gDNA 16s rDNA Successful Amplification

This PCR was successful, showing a bright smeared band (expected band size of 1,522 bp) in lane 3 indicative of 16s rDNA presence (Figure 18). This confirms that attempts at cyanobacteria colony PCR were merely not breaching the cell membranes, leaving no DNA in solution for amplification. Despite countless failed attempts to clone MVN from the lake samples, this result presents a new path forward for determining if *Microcystis aeruginosa* is present in any of the freshwater samples.

Chapter 5

Future Work

5.1 Summary of Current Project Status

5.1.1 Cyanovirin and Fusion to the Solubility Tag

At this point, the cyanovirin gene has been successfully cloned, both with and without His-tags, into the pAQ1 expression system, but transformation of these vectors into *Syn.* 7002 has been repeatedly unsuccessful. It is hypothesized that the expression of cyanovirin is inherently toxic to the cyanobacteria, making this approach to protein production impossible. To alleviate this issue, a cloning strategy has been developed to fuse the cyanovirin gene to an aggregation tag comprised of an Elastin-Like Peptide (ELP) and an intein (INT). This tag will increase the size of the cyanovirin protein by more than eight-fold, altering its folding properties, hopefully reducing the toxicity of CVN (and MVN), while also providing a low-cost protein purification platform. As it stands now, this cloning has failed. A current theory explaining why amplification of the ELP+INT tag failed is that since the ELP_NcoI_Fwd._Revised forward primer has a unique structure to accommodate the repeat in the ELP sequence, six out of twenty-seven bases of the forward primer do not match up with the pET/ELP-I-CAT plasmid template in order to generate the NcoI restriction site (discrepancies indicated by lowercase bases in Figure 19 below). It was not expected that these contradictions would prevent successful primer binding given that the other 80% of the primer matched the template plasmid.

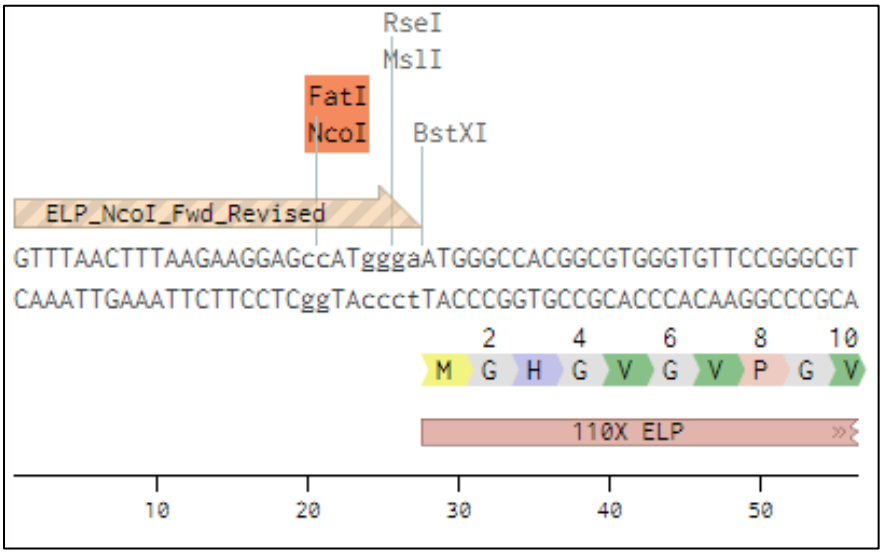

Figure 19. Discrepancy Between ELP_NcoI_Fwd._Revised Forward Primer and pET/ELP-I-CAT Plasmid Template

Assuming that the issue is primer binding, a few options are available (all three options utilize the same reverse primer "E-I-C_FUS_Rev."):

1. Utilizing the current primers (ELP_NcoI_Fwd._Revised and E-I-C_FUS_Rev), a touchdown PCR (TD-PCR) should be performed to test the impact of annealing temperature on this specific amplification. The suggested annealing temperature for this primer pair from the NEB Tm calculator is 54°C based on the primers' melting temperatures. However, these annealing temperatures are only rough approximations because primer Tm can be heavily dependent on template concentration, primer concentration, and buffer composition[38]. This issue is further complicated in this case because the binding region of the primer does not perfectly match the template. Thus, a PCR should be attempted where the annealing step operates on a temperature gradient of 58°C to 48°C (or possibly lower) where the annealing temperature only changes slightly between cycles. Primer specificity decreases with lower annealing temperatures because the primers have significantly

less kinetic energy and can locate an adequate binding site more easily. This may yield multiple amplicons, but the desired fragment can be excised using gel extraction.

2. Another option (albeit a very simple option) would be to add more bases to either end of the oligo matching the template plasmid. Given that the 3' end is where the mismatching occurs, it would make sense to start here and perhaps add another seven nucleotides or so. This will essentially "bury" the mismatching region in the middle of the oligo and the specificity of the ends should increase the likelihood of a binding event.
3. The last obvious option would be to re-design the forward primer traditionally (NcoI restriction site near the 5' end preceded by extra base pairs for cleavage) such that the binding region matches the template DNA perfectly, specifically the region immediately preceding the ELP repeat. In this instance, the primer would not have any mismatches and should bind without any issues. This is (by far) the least desirable option, however, because when the protein is expressed *in vivo*, there will be several random amino acids on the N-terminus. The hope would have to be that because the ELP protein is so massive (579 AA) the few extra amino acids will not impact solubility and folding during the purification step. Even this option is limited, however, because there are stop codons very close to the beginning of the ELP gene in pET/ELP-I-CAT. Therefore, if this is the only option remaining, a reasonable forward primer to use would be: atataCCATGGgaaggagatatacatatgggccacggcgtg. How far the primer extends into the ELP region is flexible, but this is the general anatomy. This primer has been

denoted “ELP_NcoI_Fwd_EGDIH” in Table 1 in Appendix C, indicating the five amino acids that separate the cut site from the ELP sequence (“EGDIH”).

One of the above approaches should yield successful amplification of the ELP+INT tag. This tag can then fused to the cyanovirin gene by amplifying the CVN gene from pET26B-CVN using primers E-I-C_FUS_Fwd. and CVN_BamHI_Rev. Both fragments (ELP+INT and CVN) should be purified via gel extraction and then used as template for a third PCR reaction using primers ELP_NcoI_Fwd._(insert notation for correction version) and CVN_BamHI_Rev. Given the large size difference between these two fragments, it would make sense to treat it like a ligation reaction and use seven-fold or so more CVN fragment mass then that of ELP+INT to increase the likelihood of fusion. This fusion product should subsequently be gel purified, double-digested with NcoI and BamHI, and ligated into a TOPO vector so that the gene is in an accessible plasmid for further applications and fusion PCR will not need to be repeated. At this point, the insert should be re-amplified, re-digested, ligated into pAQ1 (digested to remove YFP), and transformed into TOP10 *E. coli* for long term storage. This final plasmid should then be digested with either SphI or NsiI (or both), inoculated in exponentially growing WT *Syn.* sp. PCC 7002, grown in the bubbler rig for 1-2 days and plated on A+ Spectinomycin 50 plates until colonies appear. At which point, these colonies should be inoculated in liquid A+ and cryopreserved according to the procedure by Brand[32].

5.1.2 Isolation of *M. aeruginosa* and the MVN Gene

On the other front, attempts to clone the microvirin gene directly from the freshwater lake samples have been unsuccessful despite promising physical observations and microscopic

imaging. Attempted PCR of the 16s rDNA directly from an axenic culture of *Syn.* 7002 was unsuccessful, while genomic DNA extraction prior to PCR yielded the desired amplification. This genomic DNA extraction procedure should now be used on the isolated colonies from the lake samples and the resulting DNA can be used as template for PCR using the degenerate 16s primers and the MVN-specific primers. If these PCR's are repeatedly unsuccessful, then it is suggestable that *Microcystis aeruginosa* is not present in the lake samples and alternative sources are required to obtain this organism. If the MVN gene is successfully amplified, then it can be cloned into a TOPO cloning vector for subsequent cloning applications including fusion to the ELP+INT tag and transformation into *Syn.* 7002 (following the same logic as discussed above in terms of CVN). Part of the motivation behind identification and sequencing of the MVN gene from different sources was to assess natural variation in the non-domain region of the protein (not to mention a proof of concept for large-scale algal bloom extraction from freshwater lakes), but simply isolating one variation of the gene should now be the imminent goal. Alternative future strategies could be to abandon isolation of the WT organism and either simply purchase *Microcystis aeruginosa* from the UTEX algae bank (see Appendix E for Bold 3N medium recipe) for subsequent genomic DNA extraction and PCR or synthesize the MVN gene for insertion into the intein-solubility tag construct.

It is noteworthy that the wild lake samples need continuous maintenance over the course of the project or until a reliable MVN source is determined. There are currently nine liquid samples being stored at room temperature under moderate light for maintenance purposes. Of these nine samples, four are labeled "take two," indicative of proliferation from the lake sample using HEPES-modified BG-11 medium, while the original samples started in WFAMC. For anyone following up this work directly, the Lake Wallenpaupack samples are divided into three categories

labeled as follows: “Dense Paupack” (indicating a highly concentrated sample of algal bloom taken from Lake Wallenpaupack), “L. Wallenpaupack” (indicating a sample taken from nearby where the culture was not as dense), and “Top Skim” (indicating a sample skimmed from the top of the “Dense Paupack” jug where excessive floating organisms were noted). Additionally, there is a stack of ten plates in the 4°C refrigerator that were dilution streaked from the original lake samples. Several colonies from these plates were numbered and dilution streaked on two to three rounds of HEPES-modified BG-11 medium plates (See Figures 20 and 21 to match up the colony number with the source sample).

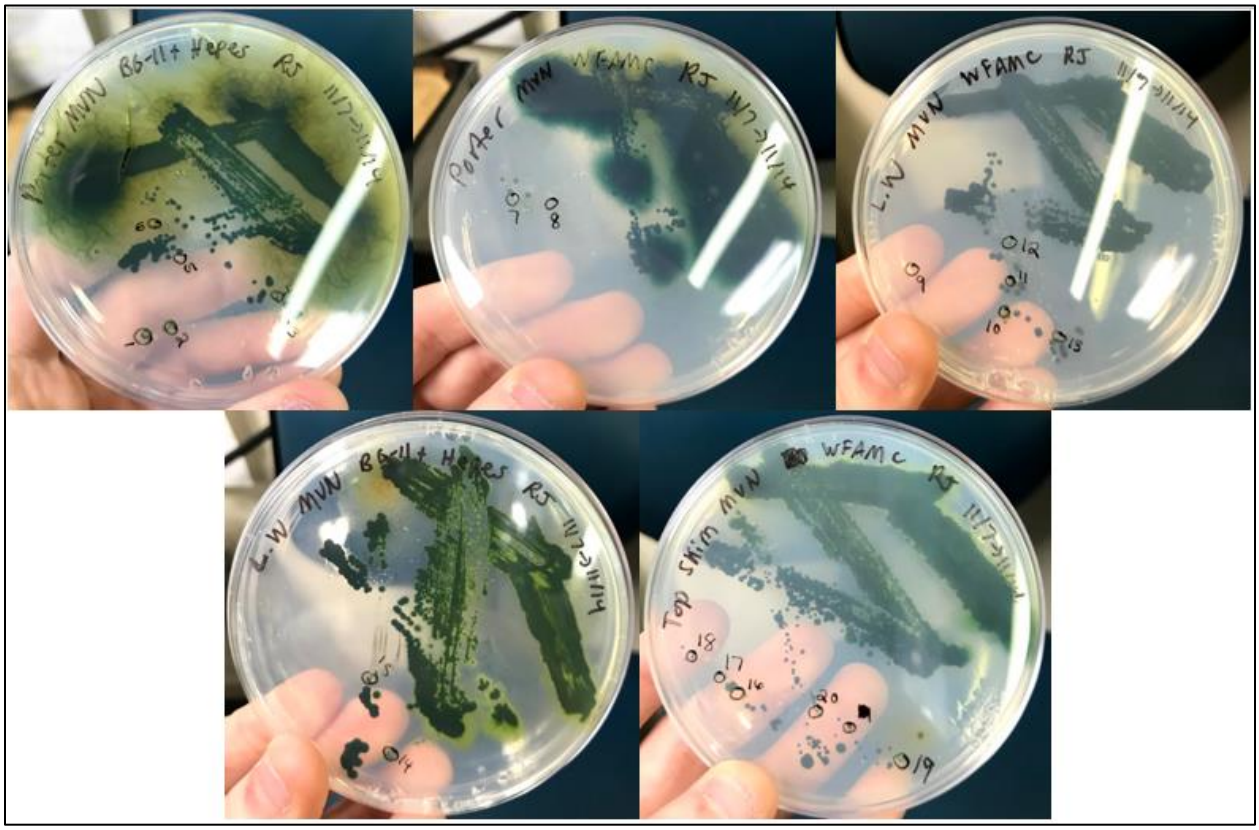

Figure 20. Wild Lake Water Samples (Colonies 1-20)

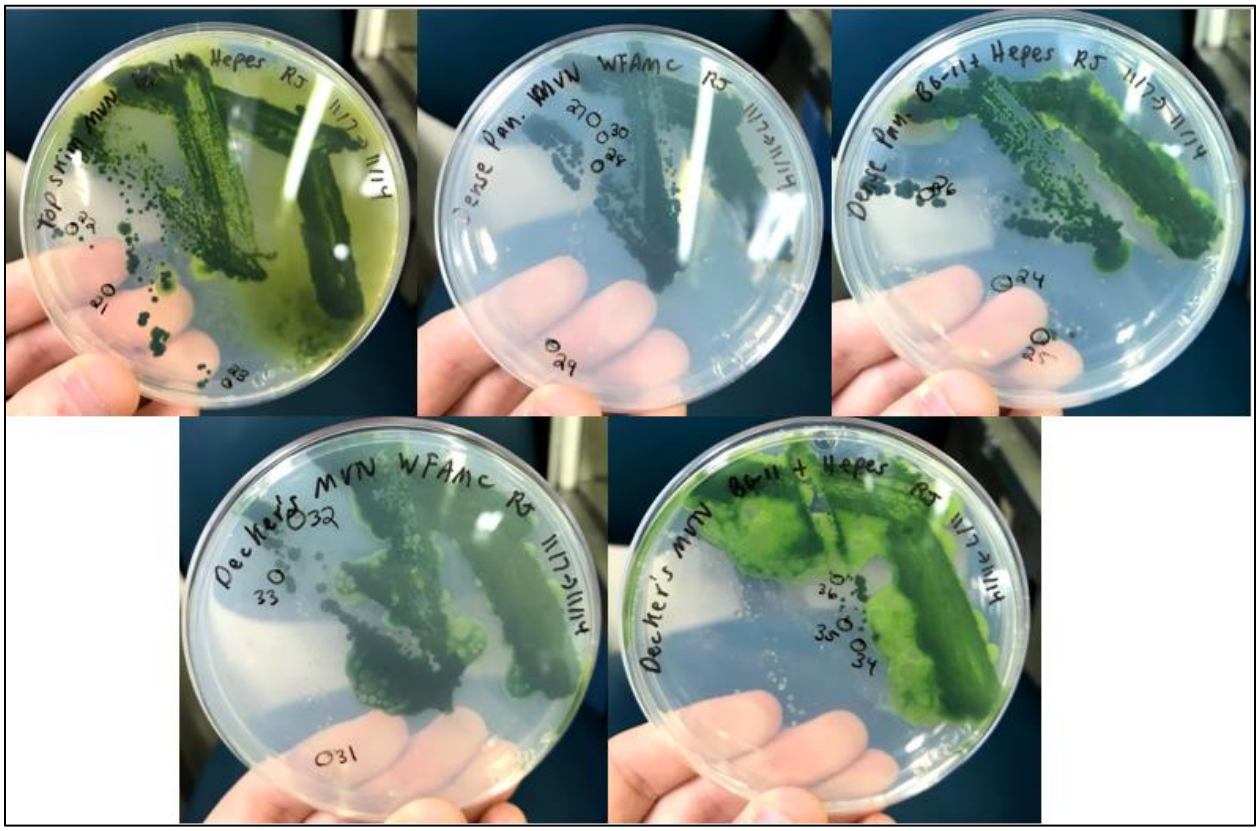

Figure 21. Wild Lake Water Samples (Colonies 21-36)

Several stacks of plates of the individual colonies were dilution streaked and also stored in the 4°C refrigerator. Select colonies were also streaked on R2A plates for bacterial contamination testing, and these plates can be found in the same refrigerator. These plates can be used for generation of axenic strains of *M. aeruginosa* if it is reasonable to assume the organism is present. When these plates are no longer viable (arguably whenever the organism will no longer proliferate in fresh media or the cells look visibly dead), colony isolation would need to be started from scratch from the original wild samples (thus the importance to continue their maintenance). None of the above mentioned cultures or colonies were cryogenically preserved because they were not confirmed monocultures.

5.2 Large-Scale Purification by Self-Cleaving Aggregation Tag

When either the CVN or MVN gene is successfully fused to the ELP+INT tag and cloned into the pAQ1 expression system, steps can be made toward optimizing the purification process. The basic premise of the aggregation tag purification is that the ELP tag will greatly increase the solubility of the protein at low temperatures (allowing it to be separated from insoluble proteins and cellular debris) and the INT tag will cleave itself in an acidic environment, leaving behind the purified recombinant protein (Figure 22).

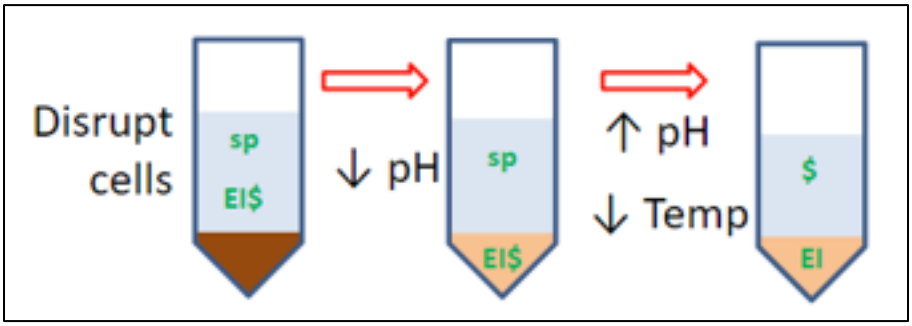

Figure 22. Purification of Recombinant Protein via Aggregation Tag

At this stage, there is the additional complication that the aggregation tag purification method developed by Wu et. al. is optimized for extraction of recombinant proteins from an *E. coli* host, utilizing terrific broth in lieu of an induction agent. Given that the CPC promoter expresses genes constitutively, no additional media formulation or induction agent is necessary, greatly simplifying the transition on that front. However, because the size, shape, and structure of *E. coli* cells differ greatly from that of *Synechococcus* (confirmed from various failed attempts at cyanobacterial colony PCR), it is likely that lysis of the harvested cells and processing of the cell

lysate will require tweaking of the protocol. The exact changes will have to be experimentally determined, but some reasonable starting points are as follows[31].

- Utilizing lysis buffers of varying pH and lysozyme concentration
- Inclusion of protease inhibitors
- Varying the sonication pulse rate and/or time exposed to sonic waves
- Altering the salt content and volume used of cleaving buffer
- Varying the temperature and time of the cleavage reaction

It would be advisable to become proficient in the purification procedure using *E. coli* prior to attempting it with cyanobacterial hosts. The lab currently has cryostocks of pET/ELP-I-CAT and pET/ELP-OsPLA (courtesy of Katie Legenski's and Bennett McKinley's work in CurtisLab), both of which are candidates for a proof of concept run[39].

Additionally, large-scale purification of the anti-HIV proteins will also require scale-up of the cyanobacterial growth and transformation rig. It would be considerable to use a large translucent culture bath with a sparger for carbon dioxide supplementation. This has not yet been attempted and will require additional development.

5.3 HIV Ketchup Microbicide

A long term goal of this project is to express the anti-HIV proteins in transgenic tomatoes grown hydroponically to produce an antiviral topical microbicide in areas with sub-optimal soil conditions. Proof of concept exists for the expression of cyanobacteria-native genes in transgenic tomatoes, but significant legwork would be required to get this off the ground, including obtaining an appropriate plasmid, designing fruit-specific promoters, developing and refining a tomato

transformation procedure via *Agrobacterium tumefaciens*, equipping the lab and greenhouse for automated, hydroponic growth conditions, and developing and scaling a protein purification procedure[40]. In turn, this component of the project would correlate closely to another project in CurtisLab funded by the Bill and Melinda Gates Foundation that is geared toward improving the agricultural yields of staple cash crops in sub-Saharan Africa through treatment with somatic-embryogenesis (SE) enhancing proteins (see Appendix H). These projects coupled would strengthen the quality of life in these developing areas by providing sustainable food sources and access to health-care aids that they can grow in their own backyards.

Appendix A

List of Abbreviations

BG-11 – Blue-Green 11 Medium

CAT – Chloramphenicol Acetyl Transferase

CVN – Cyanovirin-n

ELP – Elastin-like Peptide

HAB – Harmful Algal Bloom

INT – Intein

MVN – Microvirin

PCR – Polymerase Chain Reaction

SYN – *Synechococcus*

WFAMC – Wayne's Freshwater Algae Medium (Chlorella-modified)

WT – Wild Type

Appendix B

Plasmid Maps

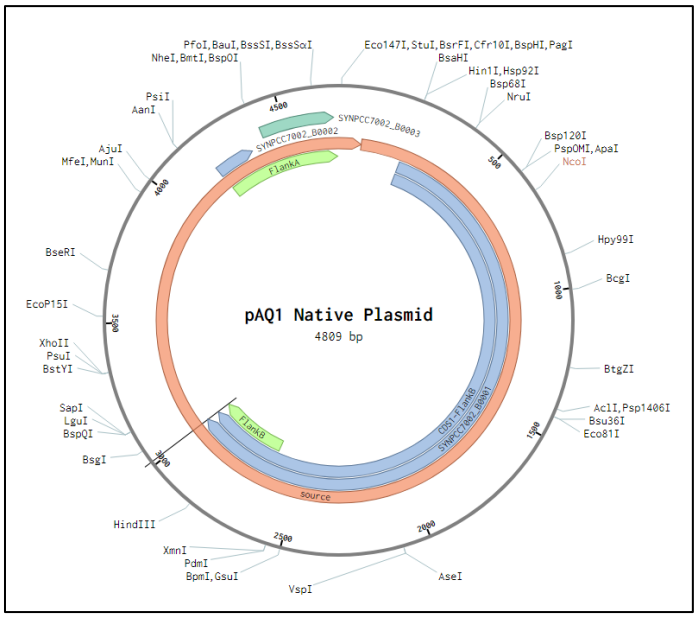

Figure 23. pAQ1 Plasmid Native to *Syn.* sp. PCC 7002

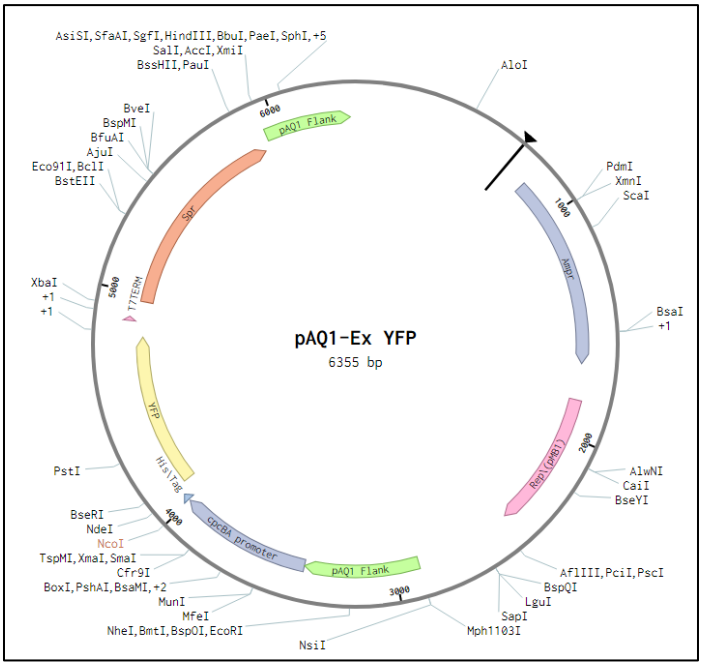

Figure 24. pAQ1-YFP Plasmid Genetically Engineered with pAQ1 Flanks

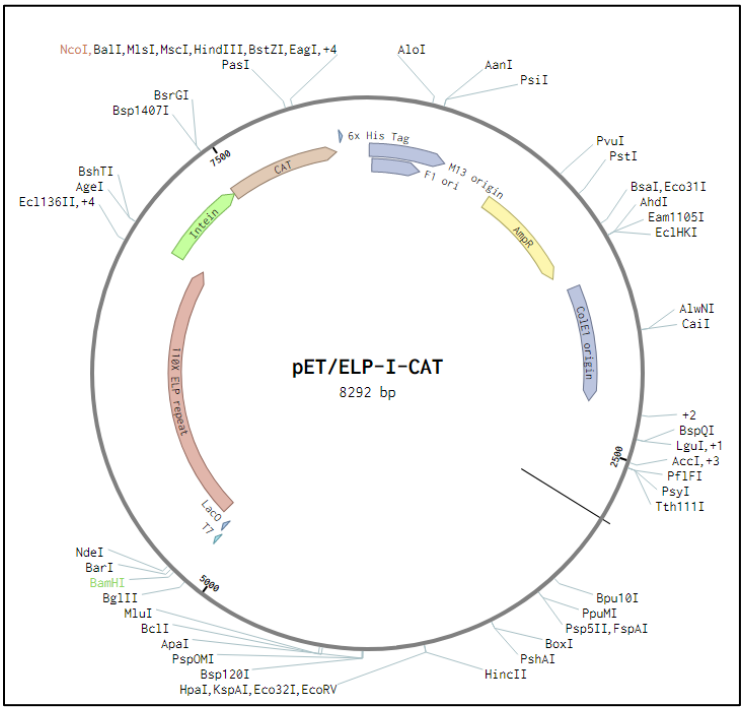

Figure 25. pET/ELP-I-CAT Plasmid

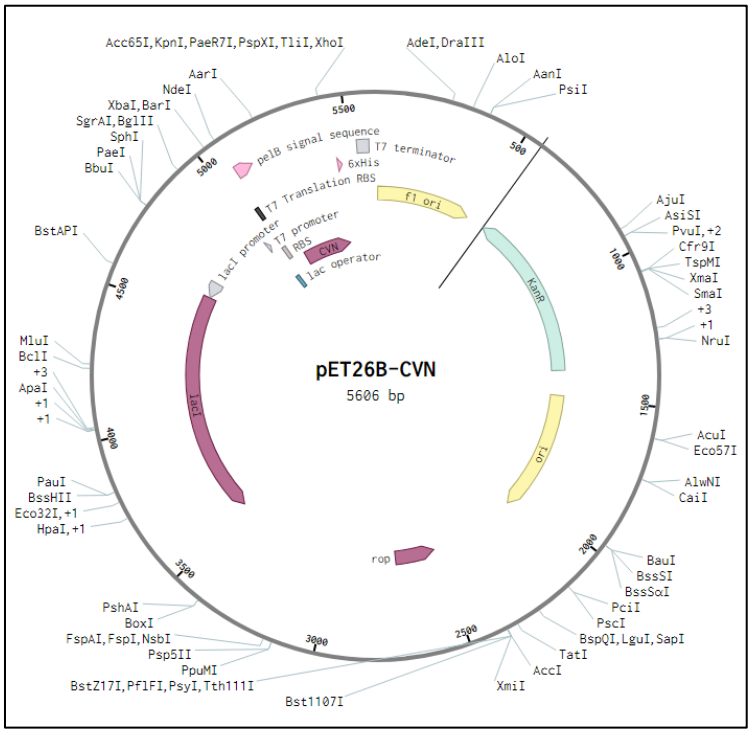

Figure 26. pET26B-CVN Plasmid

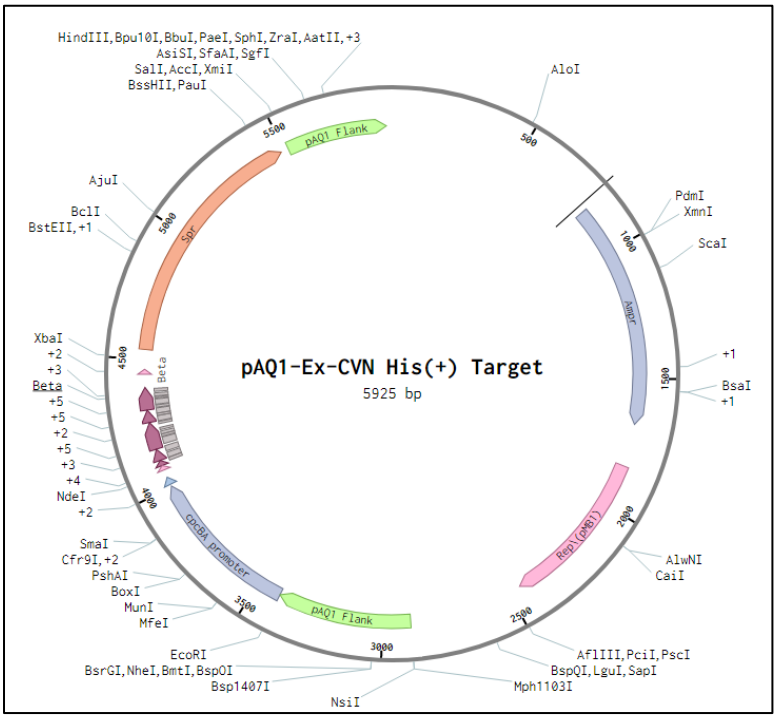

Figure 27. pAQ1-Ex-CVN His(+) Target Plasmid

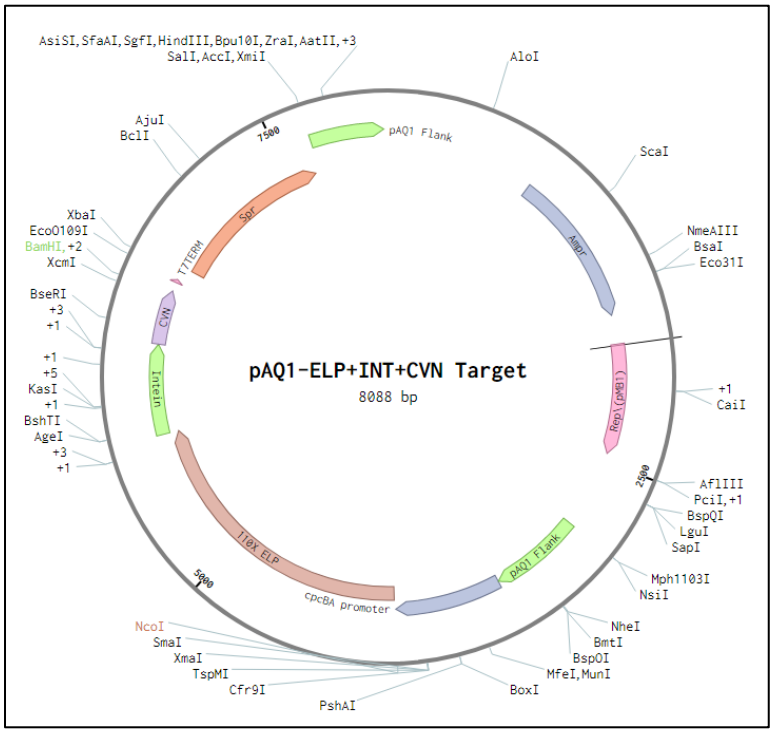

Figure 28. pAQ1-ELP+INT+CVN Target Plasmid

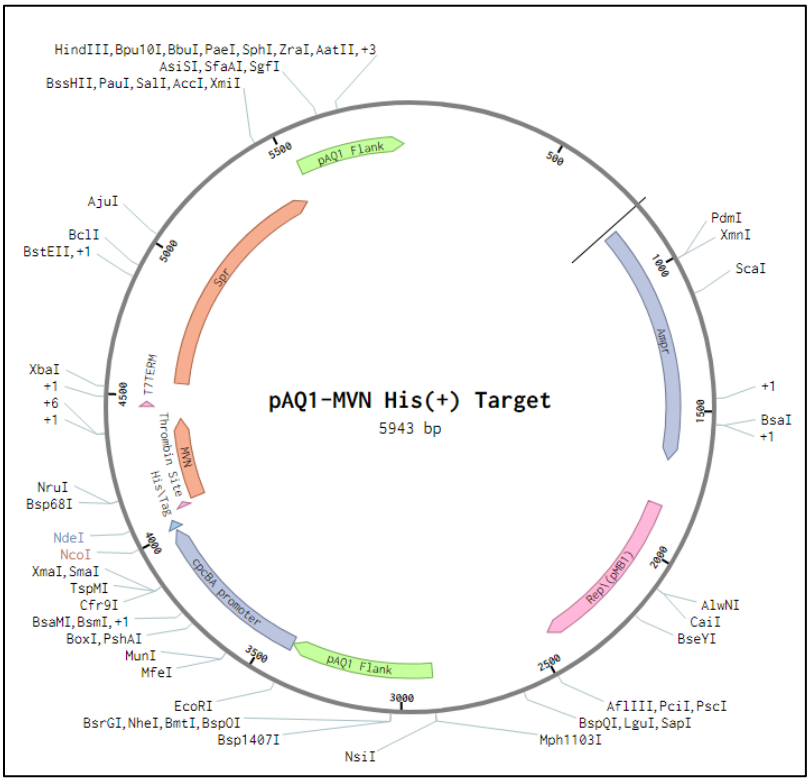

Figure 29. pAQ1-MVN His(+) Target Plasmid

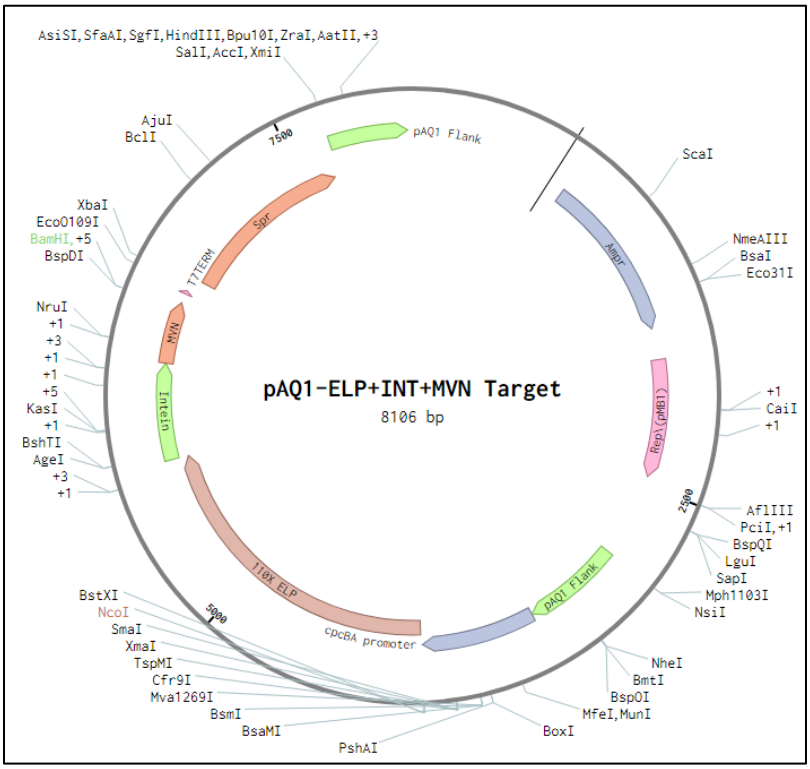

Figure 30. pAQ1-ELP+INT+MVN Target Plasmid

Appendix C

Primers

Table 1. Oligonucleotide Sequences

Primer Name	Sequence (restriction site in all caps)	Use
CVN_NdeI_Fwd.	aggaCATATGctggtgccgcgcggcagccttggtaaattctcccagacctgct	Cloning/ Sequencing
CVN_FatI_Fwd.	taggtcCATGcttggtaaattctcccagacctgct	Cloning/ Sequencing
CVN_BamHI_Rev.	catgGGATCCttattcgtatttcagggtaccgtcgatgttagc	Cloning
E_I_C_FUS_Fwd.	aaggggttgttgtacacaaccttggtaaattc	Fusion
E_I_C_FUS_Rev.	ctgggagaatttaccaaggttgtgtacaacaac	Fusion
ELP_NcoI_Fwd.	atcactCCATGGgaatgggccacggcgt	Cloning
ELP_NcoI_Fwd. (Revised)	gtttaactttaagaaggagCCATGGga	Cloning
ELP_NcoI_Fwd._ EGDIH	atataCCATGGgaaggagatatacatatgggccacggcgtg	Cloning
MVN_NdeI_Fwd.	aggaCATATGctggtgccgcgcggcagccctaatttttcgcacacttg	Cloning/ Sequencing
MVN_FatI_Fwd.	aggactCATGcctaatttttcgcacacttg	Cloning/ Sequencing
MVN_BamHI_Rev.	catgGGATCCctatccaatttccagttggc	Cloning
E_I_M_FUS_Fwd.	aaggggttgttgtacacaaccctaatttttcgcacacttg	Fusion
E_I_M_FUS_Rev.	caagtgtgcgaaaaattagggttgtgtacaacaac	Fusion
16s_Deg_Fwd.	gagtttgatcatggctcag	Organism Confirmation
16s_Deg_Rev.	ctacggctaccttgttacg	Organism Confirmation
T7_Univ_Rev.	gctagttattgctcagcgg	Sequencing

*All oligos were purchased from Integrated DNA Technologies (IDT)

Appendix D

Sequences

All of the following sequences can be found on the CurtisLab Benchling domain.

Cyanovirin (CVN) DNA Sequence

```
CTTGGTAAATTCTCCCAGACCTGCTACAACTCCGCTATCCAGGGTTCCGTTCTGACCTCCACCT
GCGAACGTACCAACGGTGGTTACAACACCTCCTCCATCGACCTGAACTCCGTTATCGAAAACGT
TGACGGTTCCCTGAAATGGCAGCCGTCCAACTTCATCGAAACCTGCCGTAACACCCAGCTGGCT
GGTTCCTCCGAACTGGCTGCTGAATGCAAAACCCGTGCTCAGCAGTTCGTTTCCACCAAAATCA
ACCTGGACGACCACATCGCTAACATCGACGGTACCCTGAAATACGAATAA
```

Microvirin (MVN) DNA Sequence

```
CCTAATTTTTCGCACACTTGTAGCAGTATCAACTACGACCCTGACAGCACAATCCTGAGTGCTG
AGTGCCAAGCTCGCGATGGAGAATGGCTCCCTACCGAACTGAGGCTTAGTGACCATATCGGTAA
TATAGATGGGGAATTGCAGTTCGGGGATCAAAACTTCCAAGAAACCTGCCAAGATTGTCACCTT
GAGTTCGGGGATGGAGAGCAATCCGTATGGTTGGTGTGTACTTGTCAAACAATGGATGGGGAAT
GGAAATCTACCCAAATACTGTTAGACAGTCAGATCGATAATAACGACAGCCAACTGGAAATTGG
ATAG
```

110X Elastin-Like Peptide (ELP) DNA Sequence

```
ATGGGCCACGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTG
TTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCC
GGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGT
GGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTG
GTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGT
TCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCG
GGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTG
GCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGG
TGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTT
CCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGG
GTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGT
TGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGT
GTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGC
CGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGG
TGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCA
GGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTG
TTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCC
GGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGT
GTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGG
```

GTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGT
GCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCG
GGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCG
CAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGG
TGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTT
CCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGG
GTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGG
GCTGGTGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATT
TCAGAATTC

Intein (INT) DNA Sequence

GCCCTCGCAGAGGGCACTCGGATCTTCGATCCGGTCACCGGTACAACGCATCGCATCGAGGATG
TTGTCGgTGGGCGCAAGCCTATTCATGTCGTGGCTGCTGCCAAGGACGGAACGCTGCATGCGCG
GCCCGTGGTGTCCTGGTTCGACCAGGGAACGCGGGATGTGATCGGGTTGCGGATCGCCGGTGGC
GCCATCcTGTGGGCGACACCCGATCACAAGGTGCTGACAGAGTACGGCTGGCGTGCCGCCGGGG
AACTCCGCAAGGGAGACAGGGTGGCGCAACCGCGACGCTTCGATGGATTCGGTGACAGTGCGCC
GATTCCGGCGCGCGTGCAGGCGCTCGCGGATGCCCTGGATGACAAATTCCTGCACGACATGCTG
GCGGAAGAACTCCGCTATTCCGTGATCCGAGAAGTGCTGCCAACGCGGCGGGCACGAACGTTCG
GCCTCGAGGTCGAGGAACTGCACACCCTCGTCGCCGAAGGGGTTGTTGTACACAAC

pAQ1-Ex-CVN His(+) Plasmid DNA Sequence

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTA
GCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGC
CCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGT
CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCA
AAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC
TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAAC
CCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAA
ATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTG
ATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA
TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAA
ACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG
ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCAC
TTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGT
CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC
CAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGG
GATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC
GTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACT
TACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTT

CTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTG
ATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTA
ACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGAT
CCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTT
GTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGAT
ACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC
TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG
TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGG
TCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGT
CGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA
TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACA
TGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
CCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGT
TTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGC
ACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAA
TTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATAC
TCAAGCTATGCATGGGGTTTTCTCGTGTTTAGGCAGCATCTTGATCCGCCTCATCATTGTTATA
GACTTCGAGCAAATCTCGTAAAACAAGCTCCCGGATATAGTCGCTGAAATTTAAGCCTTGCTCA
TTGGCTTTTTTAGTAGCCGCGCTGTACATGAGTGGGGGAAAACGAACCGCGCTCGCTGATTTTA
AATTCTTATTCGGTCTAGTCGTCATGGGATCGCCTAAGAAAGTCTCTATCATTTTACAGTATCC
AAAAGATTTGACACCCCCATTCATGGTGGTATTTTTTCTTTTCCTTTTCCCCATAGCACTGTGG
CTAGCAATAAAGCTATGGGCGATCCCTACATTTATTCTGTAGCACCAACGCTACAGCCCCTTAA
TTTCCTGTGGATAAACAGTGTGGAAATTGAGAAGAACACATGAGAATTTGTCCAGTTTTATTTT
GATGGTTATTTTTTGCGGTTGCTTTTTAAGGGAATTGTGCGTGTGGTTTCCAGTCCCCATCTGT
GCATAAGAGAAAGGAATTCGTTATAAAATAAACTTAACAAATCTATACCCACCTGTAGAGAAGA
GTCCCTGAATATCAAAATGGTGGGATAAAAAGCTCAAAAAGGAAAGTAGGCTGTGGTTCCCTAG
GCAACAGTCTTCCCTACCCCACTGGAAACTAAAAAAACGAGAAAAGTTCGCACCGAACATCAAT
TGCATAATTTTAGCCCTAAAACATAAGCTGAACGAAACTGGTTGTCTTCCCTTCCCAATCCAGG
ACAATCTGAGAATCCCCTGCAACATTACTTAACAAAAAAGCAGGAATAAAATTAACAAGATGTA
ACAGACATAAGTCCCATCACCGTTGTATAAAGTTAACTGTGGGATTGCAAAAGCATTCAAGCCT
AGGCGCTGAGCTGTTTGAGCATCCCGGTGGCCCTTGTCGCTGCCTCCGTGTTTCTCCCTGGATT
TATTTAGGTAATATCTCTCATAAATCCCCGGGTAGTTAACGAAAGTTAATGGAGATCAGTAACA
ATAACTCTAGGGTCATTACTTTGGACTCCCTCAGTTTATCCGGGGGAATTGTGTTTAAGAAAAT
CCCAACTCATAAAGTCAAGTAGGAGATTAATTCCATGGGCCATCATCATCATCATCATCATCAT
CATCACAGCAGCGGCCATATCGAAGGTCGTCATATGCTGGTGCCGCGCGGCAGCCTTGGTAAAT
TCTCCCAGACCTGCTACAACTCCGCTATCCAGGGTTCCGTTCTGACCTCCACCTGCGAACGTAC
CAACGGTGGTTACAACACCTCCTCCATCGACCTGAACTCCGTTATCGAAAACGTTGACGGTTCC
CTGAAATGGCAGCCGTCCAACTTCATCGAAACCTGCCGTAACACCCAGCTGGCTGGTTCCTCCG
AACTGGCTGCTGAATGCAAAACCCGTGCTCAGCAGTTCGTTTCCACCAAAATCAACCTGGACGA
CCACATCGCTAACATCGACGGTACCCTGAAATACGAATAAGGATCCGGCTGCTAACAAAGCCCG

AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCT
AAACGGGTCTTGACGGGTTTTTTGTCTAGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGT
TTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAA
GCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACT
ATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTAC
GGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCG
TAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCCC
TGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCG
TGGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAG
GTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACA
TAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTA
TTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGC
GAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAA
GGATGTCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAA
GCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAG
AATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAATGTCTAACAATTCG
CTCGACTCTAGGTCGACGCCTCCTGAATAAATCTATTTATACAGGGGTTGGACACGGCCCCTAA
TTTTGCTTGGTCACGCTGTAACCAATGAGCAAAGACCTTTTCGCGCTGATCGTTAGGGGCGATC
GCCTCATAGACCCGCTGACGGTTATCCCGCGATCCAAAGCGCCCCTTGTATTCCAATTTCCGGT
CAAGCTTGCCTAGGAGCGCCTGAGCAATTTCGAGCGGGCTCATTTTGTCGGTGATCGTTCTGCC
AAGAATCCACTTGATCGGCTTGGCATACCGCAAACAATTTTCCTTGAACCCTTCAAGGCTCGCC
CCGGTGAATGTCACCCCTGGTTTTAAGAACTGATGGATGTTGAGTAGCTCTAACAGGTGAATCT
TTGGTGAGAGGCATGCGACGTCGGGCCCAATTCGCCC

pAQ1-ELP+INT+CVN Plasmid Map Sequence (Target)

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTA
GCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGC
CCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGT
CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCA
AAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC
TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAAC
CCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAA
ATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTG
ATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA
TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAA
ACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG
ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCAC
TTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGT
CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGC

CAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGG
GATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC
GTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACT
TACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTT
CTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
GACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTG
ATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTA
ACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGAT
CCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTT
GTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGAT
ACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC
TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG
TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGG
TCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGT
CGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA
TGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACA
TGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
CCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGT
TTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGC
ACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAA
TTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATAC
TCAAGCTATGCATGGGGTTTTCTCGTGTTTAGGCAGCATCTTGATCCGCCTCATCATTGTTATA
GACTTCGAGCAAATCTCGTAAAACAAGCTCCCGGATATAGTCGCTGAAATTTAAGCCTTGCTCA
TTGGCTTTTTTAGTAGCCGCGCTGTACATGAGTGGGGGAAAACGAACCGCGCTCGCTGATTTTA
AATTCTTATTCGGTCTAGTCGTCATGGGATCGCCTAAGAAAGTCTCTATCATTTTACAGTATCC
AAAAGATTTGACACCCCCATTCATGGTGGTATTTTTTCTTTTCCTTTTCCCCATAGCACTGTGG
CTAGCAATAAAGCTATGGGCGATCCCTACATTTATTCTGTAGCACCAACGCTACAGCCCCTTAA
TTTCCTGTGGATAAACAGTGTGGAAATTGAGAAGAACACATGAGAATTTGTCCAGTTTTATTTT
GATGGTTATTTTTTGCGGTTGCTTTTTAAGGGAATTGTGCGTGTGGTTTCCAGTCCCCATCTGT
GCATAAGAGAAAGGAATTCGTTATAAAATAAACTTAACAAATCTATACCCACCTGTAGAGAAGA
GTCCCTGAATATCAAAATGGTGGGATAAAAAGCTCAAAAAGGAAAGTAGGCTGTGGTTCCCTAG
GCAACAGTCTTCCCTACCCCACTGGAAACTAAAAAAACGAGAAAAGTTCGCACCGAACATCAAT
TGCATAATTTTAGCCCTAAAACATAAGCTGAACGAAACTGGTTGTCTTCCCTTCCCAATCCAGG
ACAATCTGAGAATCCCCTGCAACATTACTTAACAAAAAAGCAGGAATAAAATTAACAAGATGTA
ACAGACATAAGTCCCATCACCGTTGTATAAAGTTAACTGTGGGATTGCAAAAGCATTCAAGCCT
AGGCGCTGAGCTGTTTGAGCATCCCGGTGGCCCTTGTCGCTGCCTCCGTGTTTCTCCCTGGATT
TATTTAGGTAATATCTCTCATAAATCCCCGGGTAGTTAACGAAAGTTAATGGAGATCAGTAACA
ATAACTCTAGGGTCATTACTTTGGACTCCCTCAGTTTATCCGGGGGAATTGTGTTTAAGAAAAT
CCCAACTCATAAAGTCAAGTAGGAGATTAATTCCATGGGAATGGGCCACGGCGTGGGTGTTCCG
GGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTG
TTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGG

TGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTT
CCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGG
GTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGG
CGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGT
GTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTC
CGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGG
TGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGC
GGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTG
TTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCC
GGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGT
GGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTG
GTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGT
TCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCG
GGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTG
GCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGG
TGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTT
CCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGG
GTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGT
TGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGTGGGT
GTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCAGGTGTTCCTGGTGTAGGTGTGC
CGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTGTTCCGGGTGCAGGCGTTCCGGG
TGGCGGTGTGCCGGGCGTGGGTGTTCCGGGCGTGGGTGTTCCGGGTGGCGGTGTGCCGGGCGCA
GGTGTTCCTGGTGTAGGTGTGCCGGGTGTTGGTGTGCCGGGTGTTGGTGTACCAGGTGGCGGTG
TTCCGGGTGCAGGCGTTCCGGGTGGCGGTGTGCCGGGCGGGCTGGTGAGCTCGAACAACAACAA
CAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTCAGAATTCGCCCTCGCAGAGGGC
ACTCGGATCTTCGATCCGGTCACCGGTACAACGCATCGCATCGAGGATGTTGTCGGTGGGCGCA
AGCCTATTCATGTCGTGGCTGCTGCCAAGGACGGAACGCTGCATGCGCGGCCCGTGGTGTCCTG
GTTCGACCAGGGAACGCGGGATGTGATCGGGTTGCGGATCGCCGGTGGCGCCATCCTGTGGGCG
ACACCCGATCACAAGGTGCTGACAGAGTACGGCTGGCGTGCCGCCGGGGAACTCCGCAAGGGAG
ACAGGGTGGCGCAACCGCGACGCTTCGATGGATTCGGTGACAGTGCGCCGATTCCGGCGCGCGT
GCAGGCGCTCGCGGATGCCCTGGATGACAAATTCCTGCACGACATGCTGGCGGAAGAACTCCGC
TATTCCGTGATCCGAGAAGTGCTGCCAACGCGGCGGGCACGAACGTTCGGCCTCGAGGTCGAGG
AACTGCACACCCTCGTCGCCGAAGGGGTTGTTGTACACAACCTTGGTAAATTCTCCCAGACCTG
CTACAACTCCGCTATCCAGGGTTCCGTTCTGACCTCCACCTGCGAACGTACCAACGGTGGTTAC
AACACCTCCTCCATCGACCTGAACTCCGTTATCGAAAACGTTGACGGTTCCCTGAAATGGCAGC
CGTCCAACTTCATCGAAACCTGCCGTAACACCCAGCTGGCTGGTTCCTCCGAACTGGCTGCTGA
ATGCAAAACCCGTGCTCAGCAGTTCGTTTCCACCAAAATCAACCTGGACGACCACATCGCTAAC
ATCGACGGTACCCTGAAATACGAATAAGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA
CGGGTTTTTTGTCTAGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGT
TATGACTGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCG
TGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAAC
AAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTT
GGCGTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGG
ATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGA
AACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCCCTGGAGAGAGCGAG

ATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAG
CTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCC
AGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTG
GTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAA
ATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCT
TACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCC
GACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTT
ATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCCACTA
CGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAATGTCTAACAATTCGCTCGACTCTAGGT
CGACGCCTCCTGAATAAATCTATTTATACAGGGGTTGGACACGGCCCCTAATTTTGCTTGGTCA
CGCTGTAACCAATGAGCAAAGACCTTTTCGCGCTGATCGTTAGGGGCGATCGCCTCATAGACCC
GCTGACGGTTATCCCGCGATCCAAAGCGCCCCTTGTATTCCAATTTCCGGTCAAGCTTGCCTAG
GAGCGCCTGAGCAATTTCGAGCGGGCTCATTTTGTCGGTGATCGTTCTGCCAAGAATCCACTTG
ATCGGCTTGGCATACCGCAAACAATTTTCCTTGAACCCTTCAAGGCTCGCCCCGGTGAATGTCA
CCCCTGGTTTTAAGAACTGATGGATGTTGAGTAGCTCTAACAGGTGAATCTTTGGTGAGAGGCA
TGCGACGTCGGGCCCAATTCGCCC

Appendix E

Media Recipes

Modified BG-11_0 with HEPES

For Culturing of Synechosistisis PCC 6803

	MW	[final]	[stock]	prep / L	/ 250 mL
HEPES		1.1 g/L	--	1.1 g	0.22 g
HEPES BG-11 Macro (10x)				100 mL	25 mL
$NaNO_3$		1.5 g/L	15 g/L		
K_2HPO_4		45.9 mg/L	0.459 g/L		
$CaCl_2 \bullet 2H_20$ (B)	147	47.6 mg/L	0.476/L		
$CaCl_2$ (anhydrous)	111	35.9 mg/L	0.359 g/L		
$MgSO_4 \bullet 7H_20$ (C)	246.5	75 mg/L	0.75 g/L (C)		
$MgSO_4$ (anhydrous)	120	36.5 mg/L	0.365 g/L		
Citric acid		6.5 mg/L	0.065 g/L		
Sodium carbonate		20 mg/L	0.2 g/L		
Ferric ammonium citrate		24 mg/L	0.24 g/L		
A5 + Co Trace (1000x)				1 mL	0.25 mL
$MnCl_2 \bullet 4H_20$		1.81 mg/L	0.181 g/100mL		
H_3BO_3 (boric acid)		2.86 mg/L	0.286 g/100mL		
$ZnSO_4 \bullet 7H_20$		0.222 mg/L	0.0222 g/100mL		
$Na_2MoO_4 \bullet 2H_20$ (sodium molybdate)		0.39 mg/L	0.039 g/100mL		
$CuSO_4 \bullet 5H_20$ (E)		0.079 mg/L	7.9 mg/100mL (D)		
$Co(NO_3)_2 \bullet 6H_20$		49.4 µg/L	4.94 mg/100m L		
Fe-EDTA•[$2H_20$] (F) (iron-EDTA)	403.1	40.3 mg/L	4.0 g/L	0.26 mL	0.06 mL
Fe(III)-EDTA (anhydrous)	367.1				
pH = 8.0 (record initial pH)					
Agar				12 g	3g

(A) (Sodium phosphate monobasic, dihydrate MW=156) If use $NaH_2PO_4 \bullet H_20$ (MW=138), stock contains 1.5 g/L; if using anhydrous NaH_2PO_4 (MW=120), stock contains 1.3 g/L.

(B) (dihydrate MW=147) If using anhydrous $CaCl_2$ (MW=111), stock contains 1.13 g/L

(C) (heptahydrate MW=246.5) If using anhydrous $MgSO_4$ (MW=120), stock contains 1.22 g/L

(D) (manganese II sulfate, monohydrate MW=169.0) If use anhydrous $MnSO_4$ (MW=151), stock contains 8.93 g/L.

(E) (pentahydrate MW=249.7) If using anhydrous $CuSO_4$ (MW=159.6), stock contains 0.016 g/L.

(F) (Ferrous EDTA: $FeNaC_{10}H_{12}N_2O_8$; anhydrous MW=367.1) The original Iron-chelator solution was created by adding 27.3 mg/L $FeSO_4 \bullet 7H_2O$ (FW=278) plus 37.3 mg/L Na_2-EDTA [although the degree of hydration of di-sodium EDTA is not specified on several venders (Sigma, Gibco) the dihydrate (FW=372.2) is the only available form, and this would give the same molarity (0.100 mM) for both Fe and EDTA. The iron-sodium-EDTA is now available, but apparently it has a variable degree of hydration (3 batches - all were different). Equivalent Fe and EDTA molarity can be calculated: Fe-EDTA$\bullet[2H_2O]$ (FW=403.1), 4.03 g/L; [$\bullet 2.5H_20$] (FW=412.1), 4.12 g/L; [$\bullet 3H_20$] (FW=421.1), 4.21 g/L

Adapted from: Don Bryant's Lab at The Pennsylvania State University

A+ Medium
For growth of *Synechococcus* sp. PCC 7002

If chemicals are available as anhydrous forms, hydrates or as di-sodium salts etc, then the grams per liter need to be re-calculated!

Powder or stock solution	1 Liter
NaCl (300 mM final conc.)	18g
KCl (8 mM final conc.)	0.6g
$NaNO_3$ (11.8 mM final conc.)	1.0g
$MgSO_4$ x 7 H_2O (20.3 mM final conc.)	5.0g
KH_2PO_4 (50 g/L)	1mL
$CaCl_2$ (37 g/L)	7.2 mL
Na_4-EDTA (tetra Na salt) (3 g/L)	10 mL
$FeCl_3$ x 6 H_2O (3.89 g/L liter 0.1 N HCl)	1 mL
Tris (100 g/L, pH 8.0)	10 mL
P1 Metals (1000x)	1 mL
Only for Agar plates: $Na_2S_2O_3$	2 g
Only for glycerol-containing A+:	80% glycerol
for 11 mM final conc.:	1 mL
for 27.5 mM final conc.:	2.5 mL
ddH2O	to 1 L

Autoclave and cool down to room temperature, then add
Vitamin B12 (4 mg/L) 1 mL
[final conc. in medium: 4 µg/L]

To make agar plates, add 14 g/L Bacto-Agar
Autoclave 40-50 min. The larger the volume the longer lakes the sterilization process!
Wait to cool down to 50-55 ºC, and then add Vitamin B12 stock solution and appropriate antibiotic(s).

Use glass pipettes and NOT Gilson pipettes!

Suggested final concentration of antibiotic (s) for A+ plates.

Kanamycin	100 µg/mL
Streptomycin/spectinomycin	50 µg/mL
Erythromycin	20 µg/mL
Chloramphenicol	10 µg/mL
Gentamycin	20 µg/mL

P1 Metals (to make 1 liter 1000x stock solution)

H_3BO_3 (554 mM)	34.26 g
$MnCl_2$ x 4 H_2O (21.8 mM)	4.32 g
$ZnCl_2$ (2.31 mM)	0.315 g
MoO_3 (85%)	0.03 g
$CuSO_4$ x 5 H_2O (12.0 µM)	0.003g
$CoCl_2$ x 6 H_2O (51.1 µM)	0.01215 g
ddH_2O	to 1 L

KH_2PO_4 stock

KH_2PO_4 → 0.367 M	50 g/L

$CaCl_2$ stock

$CaCl_2$ → 0.333 M	37 g/L

Na-EDTA stock

Na_4-EDTA → 7.89 mM	3 g/L

$FeCl_3$ stock

$FeCl_3$ x 6 H_2O → 14.4 mM	3.89 g/L

Dissolve $FeCl_3$ in 0.1 N HCl

Tris/HCl stock

Tris → 0.826 M	100 g/L

Adjust pH to 8.0 with conc. HCl

Modified Bold 3N Medium

For Culturing of Microcystis aeruginosa

	MW	[final]	[stock]	prep / L	/ 250 mL
Bold 3N Macro (10x)				100 mL	25 mL
$NaNO_3$		0.75 g/L	7.5 g/L		
KH_2PO_4		175 mg/L	0.175 g/L		
K_2HPO_4		74.8 mg/L	0.748 g/L		
$CaCl_2 \bullet 2H_20$ (A)	147	25.0 mg/L	0.250 g/L		
$CaCl_2$ (anhydrous)	111	18.9 mg/L	0.189 g/L		
$MgSO_4 \bullet 7H_20$ (B)	246.5	74 mg/L	0.740 g/L		
$MgSO_4$ (anhydrous)	120	36 mg/L	0.36 g/L		
NaCl		25 mg/L	0.025 g/L		
P-IV Metals (100x)				10 mL	2.5 mL
$Na_2EDTA \bullet 2H_20$ (C)		744 mg/L	7.44 g/100mL		
$FeCl_3 \bullet 6H_20$		97.3 mg/L	0.973 g/100mL		
$MnCl_2 \bullet 4H_20$		26.4 mg/L	0.264 g/100mL		
$ZnCl_2$		5.04 mg/L	50.4 mg/100mL		
$CoCl_2 \bullet 6H_20$		1.09 mg/L	10.9 mg/100mL		
$Na_2MoO_4 \bullet 2H_20$ (sodium molybdate)		0.39 mg/L	0.039 g/100mL		
Soilwater GR+ Medium (40x)		40.3 mg/L	4.0 g/L	40 mL	10 mL
Green House Soil(D)		1 tsp/L	1 tsp/200mL		
$CaCO_3$		1 mg/L	1 mg/200mL		
Autoclave					
Vitamin B12 (1000x)		4 µg/L	4 mg/L	1 mL	0.25 mL
Agar				15 g	3.75 g

(A) (dihydrate MW=147) If using anhydrous $CaCl_2$ (MW=111), stock contains 1.13 g/L

(B) (heptahydrate MW=246.5) If using anhydrous $MgSO_4$ (MW=120), stock contains 1.22 g/L

(C) Fully dissolve in stock before adding other components

(D) Purchased from UTEX. Prior to its use in soil-water media, treat soil in batches by placing it in a heat-resistant pan lined with aluminum foil, fill the soil to a so depth of ¼ inch, and bake at 150 °C for 2 hours. After it cools, cover the pan with aluminum foil and store in darkness at room temperature. Avoid excessive moisture during storage.

Adapted from: UTEX Culture Collection of Algae

Appendix F

Cyanobacteria Bubbler Rig Setup

The following growth apparatus for cyanobacterial transformation was originally designed by Erik Curtis, a former undergraduate in the lab who adapted the procedure from Dr. Donald Bryant's lab (Figure 31). The following equipment was used to setup and run the rig:

- 50 mL culture tubes
- Sponge plugs (to facilitate mass transfer of gas into the culture)
- Pasteur pipettes (to disperse the gas mixture into the liquid culture)
- Cotton (to act as an insulator between the sponge plug and the culture tube wall/Pasteur pipette and trap solid particulates)
- Silicon tubing and fittings
- Air filters (VWR PTFE 0.45 micron allowance) (for gas sterilization)
- Air pump
- Ring stand and culture tube rack
- Premixed gas manifold (for even distribution of carbon dioxide to multiple individual cultures)
- Percival (Model# E41L2) (for bright light and carbon dioxide supplementation)

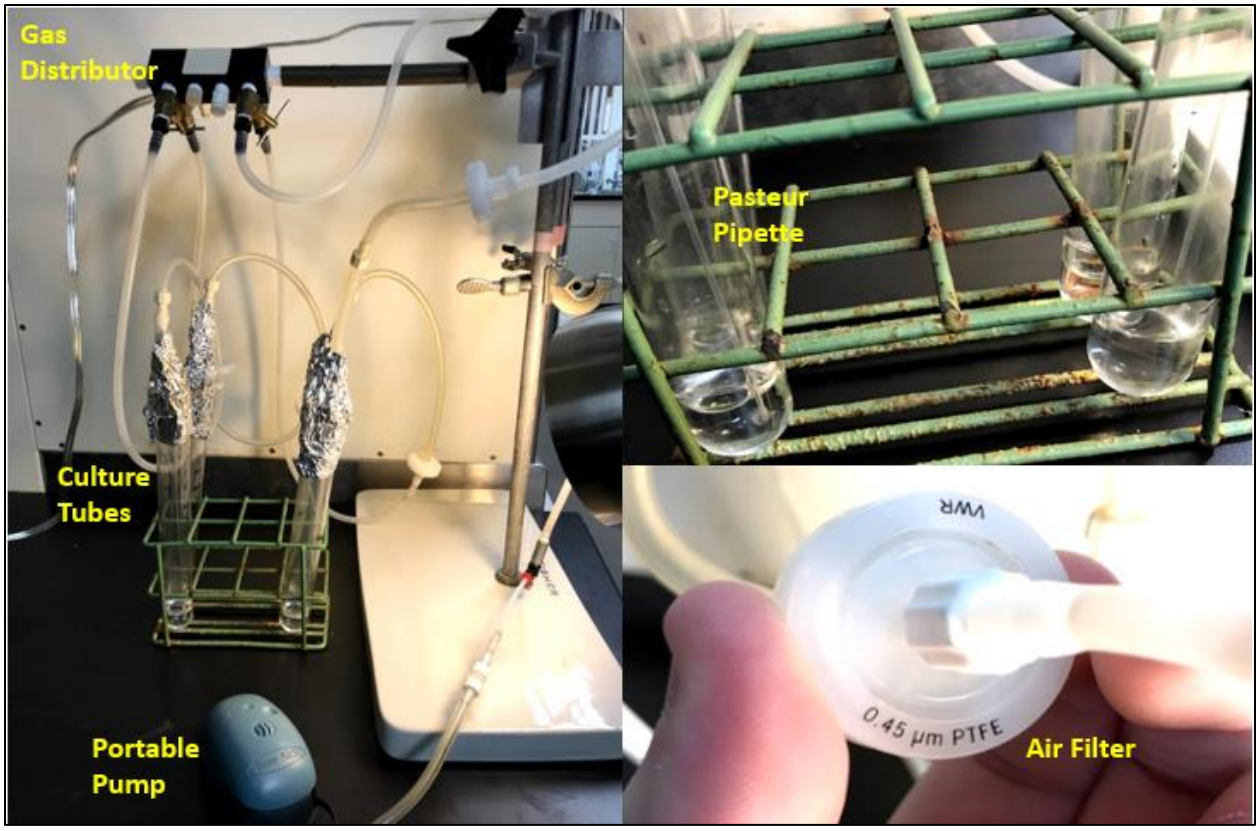

Figure 31. Setup of the Cyanobacteria Bubbler Rig for Transformation into *Syn.* 7002

Appendix G

Cloning Resources

The following resources have proved invaluable for learning and refining the techniques of molecular cloning:

- Sambrook Molecular Cloning Volumes 1-3 (Essentially "the Bible" of molecular cloning and always a good resource for troubleshooting)
- Benchling Research Tools (Public, collaborative research engine for sequence management, virtual PCR's and digestions, construct assembly wizard, and sequence alignments)
- NEB PCR Protocols (PCR reaction recipes for various polymerases and types of DNA templates)
- NEB Tm Calculator (PCR reaction annealing temperature calculator for a given primer pair, primer concentration, and DNA polymerase)
- IDT Oligo Analyzer (Tool for analyzing primer hairpins, melting temperatures, and other secondary structures)
- NEB Compatible Cohesive Ends (Resource for niche cloning applications where restriction site generation or re-cleavable ligation is key)
- NEB Cleavage Close to the Ends of DNA Fragments (Resource for designing primers containing 5' restriction sites for optimal digestion)
- NEB Ligation Calculator (Ligation reaction insert to vector ratio calculator for a given insert and backbone size)
- MUSCLE Sequence Aligner (Tool for aligning multiple sequences)

Appendix H

Other Work: PLA Cloning and Expression

Although not the main focus of this thesis, a considerable amount of time was spent developing a project to clone, express, and purify Phytocyanin-like-Arabinogalactan (PLA) proteins. These proteins have been observed to have significant "hormone-like" effects when used as a plant tissue culture media additive, nearly tripling the embryo count seen from non-PLA-treated explants (Figure 32).

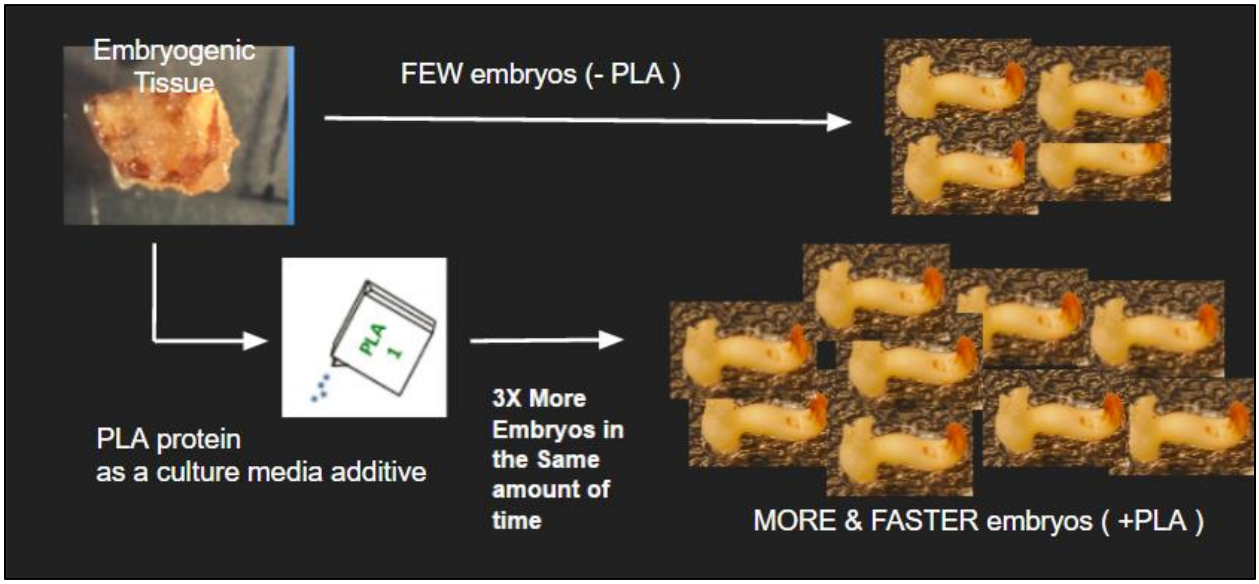

Figure 32. Phytocyanin-like Arabinogalactan (PLA) Proteins as Enhancers of Somatic Embryogenesis (SE)

The development of this project was accomplished mostly by Tina Shing Li Lai, a former plant biology graduate student in the lab[30]. Her work focused primarily on cloning, expressing, and purifying *Theobroma cacao* PLA (TcPLA) proteins via nickel resin affinity chromatography. As was discussed earlier, the pitfalls of using polyhistidine tag purification are numerous, but among the most detrimental are cost and scalability. In order to circumvent these purification obstacles, my contribution to the project was to clone these TcPLA genes into the pET14B expression vector without polyhistidine tags in the hopes of pursuing size-exclusion purification

via Tangential Flow Filtration (TFF). The basic premise of TFF is to simply pump the cell lysate tangentially across a series of two filtration membranes to remove large, extracellular debris and proteins and small cytotoxins. Depending on the molecular weight cutoff of the membrane, small proteins will pass through into the permeate while larger ones are retained in the retentate (Figure 33)[41]. This purification strategy is sound for PLA proteins because they are quite small (ranging from 11-14 kDa) compared to other proteins that might be in the cell lysate, but still quite large compared to the toxins found in the cytoplasm of *E. coli*.

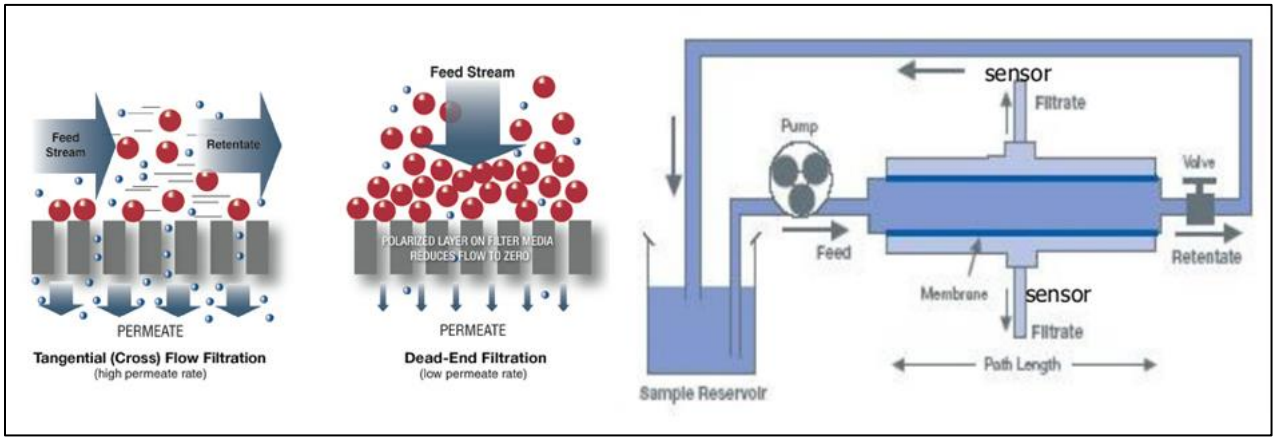

Figure 33. Theory and Schematic of Tangential Flow Filtration Protein Purification

The details of the cloning will be skipped for simplicity, but two of the most promising TcPLA genes (TcPLA1 and 3) were successfully cloned into pET14B without polyhistidine tags utilizing the same general procedure outlined in Chapter 3 (see Figure 34 for the target plasmid map of pET14B-TcPLA1 His(-)).

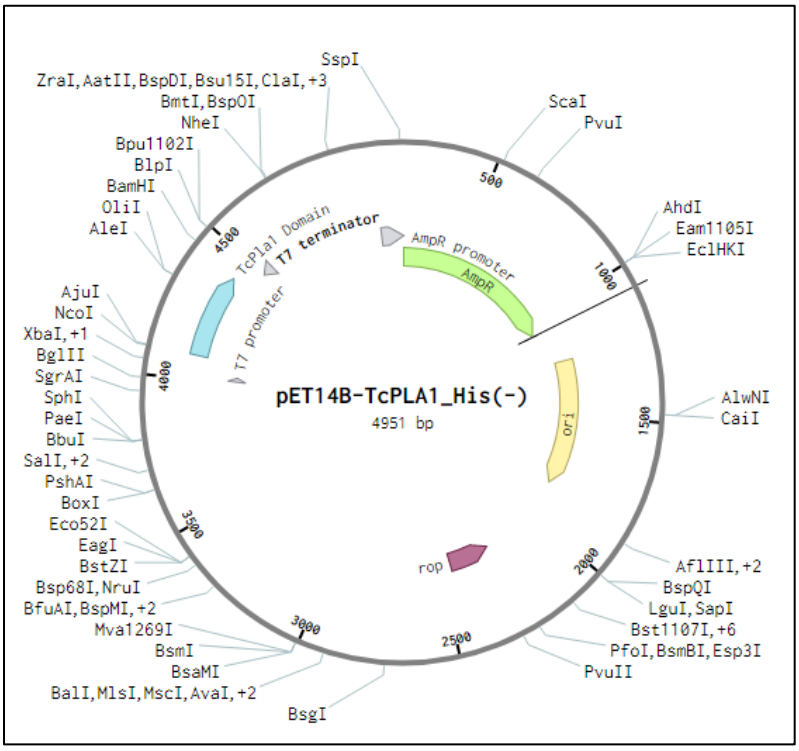

Figure 34. pET14B-TcPLA1 His(-) Plasmid Map

The final, sequence-verified constructs were transformed into BL21 (DE3) *E. coli*, and protein was expressed via growth in auto-induction media (essentially standard Luria Broth bacteria media supplemented with lactose to induce activation of the lac operon via glucose deprivation). As negative controls, pET14B native plasmid in BL21 cells and empty (no DNA) BL21 cells were also inoculated in auto-induction media to show expected background protein. The *E. coli* was harvested via centrifugation, and protein was extracted via freeze-thaw cycles on dry ice. The soluble and insoluble protein fractions were subsequently visualized on a 16% Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis Gel (SDS-PAGE) (Figure 35).

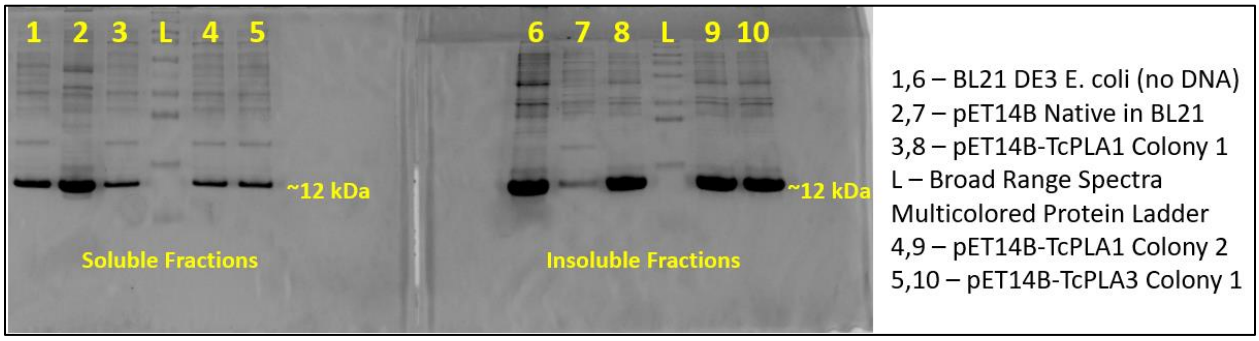

Figure 35. SDS-PAGE Analysis of TcPLA1 and 3 Protein Expression

SDS-PAGE analysis of the cell lysate revealed protein contamination very close to the size of the recombinant protein (~12 kDa) in the negative controls (lanes 1, 2, 6, and 7), thus making it impossible to conclusively determine if the TcPLA proteins were successfully expressing. It was hypothesized that perhaps there is an endogenous 12 kDa protein in the BL21 (DE3) *E. coli*, but no literature corroborated this. Several corrective measures were taken in an attempt to alleviate the issue, including making all fresh media and reagent stocks, sterilizing all benchtop equipment, and testing BL21 cells from other labs, but all the gels looked the same as Figure 35. After several weeks of deliberation, it was realized that the enzyme lysozyme, which was used in the lysis buffer for homogenizing the harvested cells, is 14.3 kDa, which was similar enough to the target protein so therefore likely indecipherable on an SDS gel. The lysis procedure was adjusted accordingly to omit lysozyme and include an ethanol-dry ice bath for freeze-thaw to facilitate cell wall degradation without lysozyme. The resulting lysate was run on a 16% SDS-PAGE gel (Figure 36).

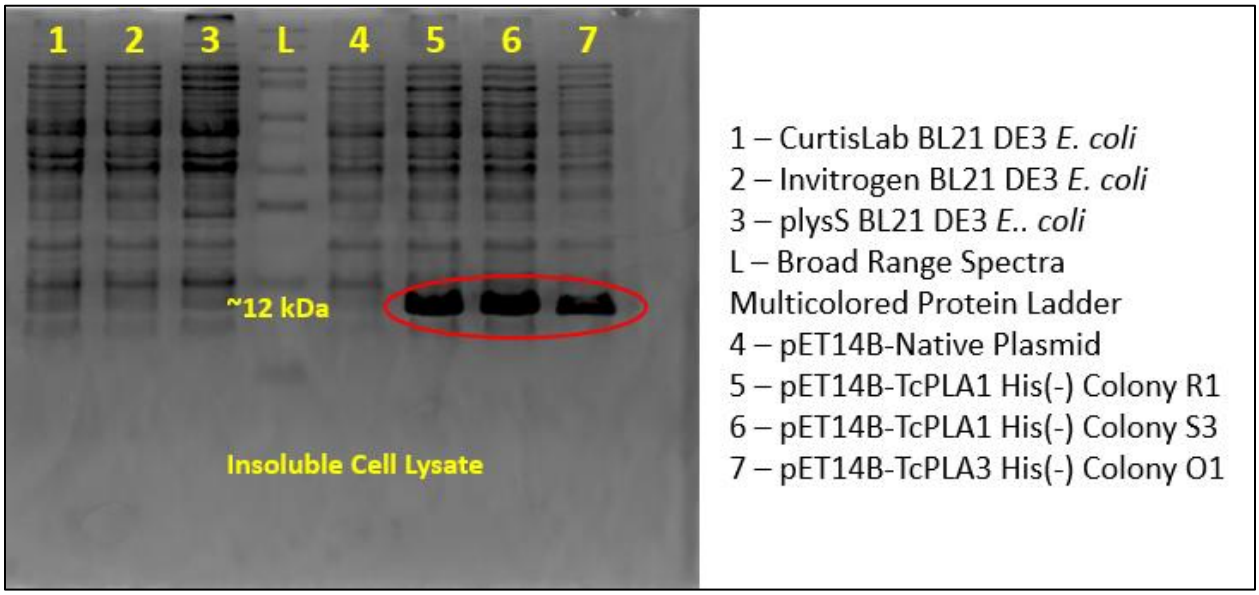

Figure 36. SDS-PAGE Analysis of TcPLA1 and 3 Protein Expression Withholding Lysozyme as Harvesting Reagent

With the removal of lysozyme, the negative controls (lanes 1-4) only contained background protein, and overexpression of the TcPLA1 and 3 proteins is readily apparent (lanes 5-7). Despite this success, however, these proteins remained in the insoluble fraction of the cell lysate, indicating the formation of inclusion bodies around the recombinant proteins, thus altering their size, and making it impossible to purify them via size-exclusion Tangential Flow Filtration. From this point on, the lab began to heavily pursue the ELP+INT aggregation tag as the desired purification technique.

BIBLIOGRAPHY

1. "What Is HIV?" *U.S. Department of Veterans Affairs*, 21 June 2005, www.hiv.va.gov/patient/basics/what-is-HIV.asp.
2. "Global Statistics." *HIV.gov*, 27 Feb. 2018, www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics.
3. Kharsany, Ayesha B.M., and Quarraisha A. Karim. "HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities." *The Open AIDS Journal*, Bentham Open, 8 Apr. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4893541/.
4. Dewi Rokhmah, Khoiron. "The Role of Sexual Behavior in the Transmission of HIV and AIDS in Adolescent in Coastal Area." *Procedia Environmental Sciences*, Elsevier, 31 Jan. 2015, www.sciencedirect.com/science/article/pii/S187802961500016X.
5. "Fact 7: About Three Quarters of the African Population Live of Less than $2, Half of the Population of Less than $1.25 per Day. People in Rural Areas Are More Often Struck by Poverty." *Global Growing*, global-growing.org/en/content/fact-7-about-three-quarters-african-population-live-less-2-half-population-less-125-day.
6. Rosen, Sydney, and Lawrence Long. "How Much Does It Cost to Provide Antiretroviral Therapy for HIV/AIDS in Africa?" *Center for International Health*.
7. "Cyanovirin-N - an Overview." *ScienceDirect Topics*, www.sciencedirect.com/topics/medicine-and-dentistry/cyanovirin-n.
8. Koharudin, Leonardus M. I., and Angela M. Gronenborn. "Antiviral Lectins as Potential HIV Microbicides." *NCBI*, U.S. National Library of Medicine, 8 July 2014, www.ncbi.nlm.nih.gov/pmc/articles/PMC4149946/.

9. "Freshwater Harmful Algal Blooms." *EPA*, Environmental Protection Agency, 1 Nov. 2016, www.epa.gov/research-grants/freshwater-harmful-algal-blooms. "The Science of Harmful Algal Blooms." *USGS.gov*, 24 Oct. 2016, www.usgs.gov/news/science-harmful-algae-blooms.
10. "The Science of Harmful Algal Blooms." *USGS.gov*, 24 Oct. 2016, www.usgs.gov/news/science-harmful-algae-blooms.
11. Rastogi, et al. "Bloom Dynamics of Cyanobacteria and Their Toxins: Environmental Health Impacts and Mitigation Strategies." *Frontiers*, Frontiers, 28 Oct. 2015, www.frontiersin.org/articles/10.3389/fmicb.2015.01254/full.
12. Anderson, Donald M, et al. "Estimated Annual Economic Impacts from Harmful Algal Blooms (HABs) in the United States." *Woods Hole Oceanographic Institution*, Sept. 2000.
13. Illustration of nutrient sources of harmful algal blooms. Digital image. *USGS science for a changing world*. 24 October 2016. https://www.usgs.gov/media/images/illustration-nutrient-sources-harmful-algal-blooms-usgs-image
14. Mori, Toshiyuki, and Michael Boyd. "Cyanovirin-N, a Potent Human Immunodeficiency Virus-Inactivating Protein, Blocks Both CD4-Dependent and CD4-Independent Binding of Soluble gp120 (sgp120) to Target Cells, Inhibits sCD4-Induced Binding of sgp120 to Cell-Associated CXCR4, and Dissociates Bound sgp120 from Target Cells ." *American Society for Microbiology*, 1 Mar. 2001, aac.asm.org/content/45/3/664.full.
15. Coffin, John M. "Immunopathogenic Mechanisms of HIV Infection." *NCBI*, U.S. National Library of Medicine, 1 Jan. 1997, www.ncbi.nlm.nih.gov/books/NBK19451/.

16. Potential mechanisms for HIV-1 transmission across the mucosal epithelium. Digital image. *The Lancet.* www.thelancet.com/journals/lancet/article/PIIS0140-6736(07)60202-5/fulltext.

17. Cyanovirin-n disruption of the CD4 receptor via gp120. Digital image. *Slideplayer.* http://slideplayer.com/slide/8177874/

18. Alexandre, Kabamba B., et al. "The Lectins Griffithsin, Cyanovirin-N and Scytovirin Inhibit HIV-1 Binding to the DC-SIGN Receptor and Transfer to CD4+ Cells." *Virology*, U.S. National Library of Medicine, 20 Feb. 2012, www.ncbi.nlm.nih.gov/pmc/articles/PMC3274779/.Balzarini, Jan, and Lut Van Damme. "Microbicide Drug Candidates to Prevent HIV Infection." *The Lancet*, 14 Feb. 2007, www.thelancet.com/journals/lancet/article/PIIS0140-6736(07)60202-5/fulltext.

19. O'Keefe, Barry R., et al. "Scaleable Manufacture of HIV-1 Entry Inhibitor Griffithsin and Validation of Its Safety and Efficacy as a Topical Microbicide Component." *PNAS*, National Academy of Sciences, 14 Apr. 2009, www.pnas.org/content/106/15/6099.

20. Brichacek, B, et al. "In Vivo Evaluation of Safety and Toxicity of a Lactobacillus Jensenii Producing Modified Cyanovirin-N in a Rhesus Macaque Vaginal Challenge Model." *PloS One.*, U.S. National Library of Medicine, 12 Nov. 2013, www.ncbi.nlm.nih.gov/pubmed/24265721.

21. Huskens, Dana, et al. "Microvirin, a Novel α(1,2)-Mannose-Specific Lectin Isolated from Microcystis Aeruginosa, Has Anti-HIV-1 Activity Comparable with That of Cyanovirin-N but a Much Higher Safety Profile." *Journal of Biological Chemistry*, 6 Aug. 2010, www.jbc.org/content/285/32/24845.long.

22. Kehr, Jan-Christoph, et al. “A Mannan Binding Lectin Is Involved in Cell–Cell Attachment in a Toxic Strain of Microcystis Aeruginosa.” *Molecular Microbiology*, Wiley/Blackwell (10.1111), 5 Dec. 2005.

23. Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology, 5*, 172. http://doi.org/10.3389/fmicb.2014.00172

24. Inclusion body formation in *E. coli* from overexpression of recombinant protein. Digital image. *Chemical and Engineering News*. https://pubs.acs.org/cen/science/86/8641sci1.html

25. Zhou, Jie, et al. “Discovery of a Super-Strong Promoter Enables Efficient Production of Heterologous Proteins in Cyanobacteria.” *Nature News*, Nature Publishing Group, 28 Mar. 2014, www.nature.com/articles/srep04500.

26. Curtis, E. (2016). Development of the Marine Cyanobacterium *Synechococcus* sp. PCC 7002 as a Scalable Low-cost Photoautotrophic Protein Production Platform. The Pennsylvania State University Schreyer Honors College.

27. Horton, Robert M., et al. “Engineering Hybrid Genes without the Use of Restriction Enzymes: Gene Splicing by Overlap Extension.” *Gene*, Elsevier, 15 Jan. 2003, www.sciencedirect.com/science/article/pii/0378111989903594.

28. Xu, Y, et al. “Expression of Genes in Cyanobacteria: Adaptation of Endogenous Plasmids as Platforms for High-Level Gene Expression in Synechococcus Sp. PCC 7002.” *Methods in Molecular Biology (Clifton, N.J.).*, U.S. National Library of Medicine, 24 Sept. 2010, www.ncbi.nlm.nih.gov/pubmed/20960136.

29. Hoffmann, A, and R G Roeder. “Purification of His-Tagged Proteins in Non-Denaturing Conditions Suggests a Convenient Method for Protein Interaction Studies.” *Nucleic Acids Research*, U.S. National Library of Medicine, 25 Nov. 1991, www.ncbi.nlm.nih.gov/pmc/articles/PMC329157/.

30. Shing Li Lai, T. (2016). Enhancement of Somatic Embryogenesis in *Theobroma Cacao* using *Phtyocyanin-like- 1 Arabinogalactan (PLA)* Protein Domains. The Pennsylvania State University Graduate School.

31. Wu, W Y, et al. “Recombinant Protein Purification by Self-Cleaving Aggregation Tag.” *Nature Protocols.*, U.S. National Library of Medicine, www.ncbi.nlm.nih.gov/pubmed/17406465.

32. Brand, Jerry J. “Cryopreservation of Cyanobacteria.” *Cyanosite.bio.purdue.edu*, www-cyanosite.bio.purdue.edu/protocols/cryo.html.

33. “*Microcystis Aeruginosa*.” Microbewiki, microbewiki.kenyon.edu/index.php/Microcystis_aeruginosa.

34. Azevedo, Flávio, et al. “Colony PCR.” *SpringerLink*, Springer, New York, NY, 1 Jan. 1970, link.springer.com/protocol/10.1007/978-1-4939-7060-5_8.

35. Hoiczyk, Egbert, and Alfred Hansel. “Egbert Hoiczyk.” *Journal of Bacteriology*, 1 Mar. 2000, jb.asm.org/content/182/5/1191.full?site=JBacteriol&utm_source=TrendMDJBacteriol&utm_medium=TrendMDJBacteriol&utm_campaign=trendmdalljournals_0.

36. Weisburg, W G, et al. “16S Ribosomal DNA Amplification for Phylogenetic Study.” *Journal of Bacteriology.*, U.S. National Library of Medicine, Jan. 1991, www.ncbi.nlm.nih.gov/pubmed/1987160.

37. Rasmussen, J.P., Barbez, P.H., Burgoyne, L.A. and Saint, C.P. (2008), Rapid preparation of cyanobacterial DNA for real-time PCR analysis. Letters in Applied Microbiology, 46: 14–19. doi:10.1111/j.1472-

38. Chavali, Sreenivas, et al. “Oligonucleotide Properties Determination and Primer Designing: a Critical Examination of Predictions.” *BIOINFORMATICS*, vol. 21, no. 20, pp. 3918–3925., doi:10.1093/bioinformatics/bti633.

39. Legenski, K. (2016). Rice Embryogenic Protein Production and Distribution. The Pennsylvania State University Schreyer Honors College.

40. M.H., Soliman, et al. “New Approach for Generating Transgenic Tomato Plants Expressing the Cyanobacterium GRX-2 Gene.” *Genetics Department. Faculty of Agriculture - Cairo University*.

41. Van Reis, Robert, and Andrew Zydney. “Membrane Separations in Biotechnology.” *Current Opinion in Biotechnology*, Elsevier Current Trends, 30 Mar. 2001, www.sciencedirect.com/science/article/pii/S0958166900002019.

ACADEMIC VITA

Ryan Jones

Education

The Pennsylvania State University **University Park, PA**
Schreyer Honors College *Graduation: Spring 2018*
Bachelor of Science in Chemical Engineering | Minor in Engineering Leadership Development

- Chemical Engineering thesis in the expression, purification and applications of HIV-binding lectin proteins

Relevant Experience

Regeneron Pharmaceuticals, Inc. **Tarrytown, NY**
Preclinical Manufacturing Process Development Analytics Intern *May 2017-August 2017*

- Developed and optimized stability and robustness assays for REGN2477 Protein A pools, varying loading and pool concentration, temperature, High Molecular Weight (HMW) aggregate levels, and time
- Utilized JMP statistical software to create Designs of Experiment and model monoclonal antibody capture conditions
- Piloted an effort to transition the Analytics team from manual extraction of residual host cell CHO DNA from protein batches to an automated method, reducing hands-on time from six to three hours
- Provided daily laboratory support including high-throughput turnaround of cell culture titer batches and buffer prep

Chemical Engineering Integration of Biology and Material Science REU **University Park, PA**
Research Mentor *May 2016-August 2016*

- Achieved successful molecular cloning and expression of plant embryogenic genes via a $0.6MM grant from the Gate's Foundation to aid crop improvement programs in Africa
- Developed and troubleshot a large-scale production pathway for PLA proteins to aid in the fight against genetic variation in seed-born plant species through *in vitro* up regulation of known embryogenic transcription factors
- Mentored an inexperienced peer in molecular biology techniques and laboratory protocols

CurtisLab Research Group **University Park, PA**
Undergraduate Thesis Researcher *November 2015-May 2018*

- Dynamically researched the expression and applications of an HIV binding lectin protein with anti-transmission capabilities
- Independently maintained a dozen sterile plant tissue culture lines used for evaluating the effectiveness of various plant embryogenic proteins and implementation in Temporary-Immersion Bioreactors (TIBs)

Antler Ridge Winery **Ulster, PA**
Manufacturing Intern *May 2015-August 2015*

- Facilitated all aspects of industrial wine production from grape harvest and batch fermentation to bottling and sale
- Operated the mechanical bottling line and addressed any unit malfunctions, bottling over seven hundred cases daily
- Interacted with hundreds of customers and connoisseurs at on-site wine bar and local wine and cheese festivals

Leadership

Pollock Dining Commons - The Mix C-Store **University Park, PA**
Student Manager *September 2014-May 2018*

- Achieved three promotions in less than two years for outstanding display of leadership, ultimately obtaining the highest student position available, working nearly 30 hours per work
- Supervised and mentored nearly one hundred student employees in fast-paced, made-to-order food preparation and customer service, collectively earning $1.2MM in profits during the 2015-2016 fiscal year
- Organized training sessions, meetings, and international appreciation events for new and promoted student employees
- Generated monthly schedules for fifty upper level employees and resolve disputes accordingly

Presentations

Expression and Purification of Plant Embryogenic Proteins

- Chemical Engineering REU Symposium *August 2016*
- American Chemical Society Local Chapter UG Research Symposium *September 2016*

Awards

- **First Place** 2016 American Institute of Chemical Engineers Undergraduate Student Poster Competition (Food/Pharma/Biotech Division)
- Recipient of The Penn State Chemical Engineering Bioendowment Fellowship ($6K to fund thesis research)
- Recipient of The President's Freshman Award for Academic Excellence for obtaining a 4.0 GPA in the Fall 2014 semester
- Recipient of The H. Stethers CH E Family Scholarship (Junior and Senior Years)
- Recipient of The Katherine Jennings Jones CH E Family Scholarship (Senior Year)