

**DESIGN AND PRODUCTION OF A HYDROGEL FORMING POLYPEPTIDE:
ENGAGING HIGH SCHOOL STUDENTS IN PROTEIN DESIGN**

JAMES K. DEYLING

Bachelor of Science in Biology

Cleveland State University

May 2013

Submitted in partial fulfillment of requirements for the degree

MASTER OF SCIENCE IN BIOMEDICAL ENGINEERING

at the

CLEVELAND STATE UNIVERSITY

December 2016

We hereby approve this thesis for

James K. Deyling

Candidate for the Master of Science in Biomedical Engineering degree

for the Department of Chemical and Biomedical Engineering

and the CLEVELAND STATE UNIVERSITY

College of Graduate Studies by

Thesis Committee Chairperson, Dr. Nolan Holland
Department of Chemical and Biomedical Engineering
Cleveland State University

Date

Thesis committee member, Dr. Chandra Kothapalli
Department of Chemical and Biomedical Engineering
Cleveland State University

Date

Thesis committee member, Dr. Christopher Wirth
Department of Chemical and Biomedical Engineering
Cleveland State University

Date

Date of Defense: December 2, 2016

ACKNOWLEDGEMENTS

I would like to express my gratitude to my Professor Dr. Nolan Holland for his guidance, valuable insight, and generous volunteering of time during the research and writing of my thesis.

I would like to thank Dr. Edward Turk for his insight on the blueprint of this course, and for providing educational resources.

I would also like to thank the Department of Biomedical and Chemical Engineering for providing financial support during my studies at Cleveland State University.

I would like to thank my mother and father for all of their love and support.

I would also like to thank my coworkers and associates for their words of encouragement and guidance throughout my research, including Jim, Eric, Kevin, Jack, Adam, Richard, Adil, Aditya, James Gillihan, Nick and Prattyush.

ENGINEERING A HYDROGEL POLYPEPTIDE AS A MODEL FOR HIGH SCHOOL EDUCATION

JAMES K. DEYLING

ABSTRACT

Bioinks are a class of hydrogel that have the potential to be the ink used in the creation of printed organs, connective tissue, and other important structures within the body. One class of material that may be a suitable bioink hydrogel is elastin-like polypeptides (ELPs), which are synthetic biopolymers inspired by the naturally existing connective tissue elastin. ELPs consist of a repeat pentapeptide sequence $(GXGVP)_n$, where X is any of the 20 naturally existing amino acids other than proline. These biomolecules are capable of exhibiting environmental responsiveness when exposed to certain stimulus such as salt concentration, temperature, and pH, depending on their chain length and guest residue. This is dictated by the ELP inverse transition temperature (T_i) which affords it the property of being soluble below T_i and becoming insoluble above T_i . By changing the molecular structure, above T_i their association has the potential to result in the assembly of the ELPs into micelles and other microstructures.

The creation of these proteins requires the use of molecular biology, namely producing recombinant DNA through protein engineering approaches. Molecular biology is a complex and rapidly expanding field, typically not addressed in full until the latter half of an undergraduate biology degree. However, the basic techniques can be implemented as straightforward protocols. This work establishes a robust set of

molecular biology protocols established with the goal of making this field more approachable to younger audiences, particularly to students in high school. Typically, safety concerns and a difficult barrier to entry prevented the inclusion of younger students in biotechnology, but this protocol establishes a safe and high yield system that gives students an introduction to this field while contributing to actual research projects.

Of significance is the DNA assembly system used in this study, which uses a newer and more robust ligation technique for assembling large sequences of pentapeptides. The Gibson assembly system was exploited to expand the length of ELP molecules through subsequent rounds of cloning. Multiple round of cloning were performed on different ELP constructs, and the sequence data and system specifics are reported within. Novel triblock ELP polypeptides were created in this study, which are capable of physically crosslinking and forming hydrogels under specific conditions

TABLE OF CONTENTS

| | Page |
|---|-------------|
| ABSTRACT..... | iv |
| LIST OF TABLES..... | viii |
| LIST OF FIGURES..... | ix |
| CHAPTER | |
| I. INTRODUCTION..... | 1 |
| 1.1 Hydrogels..... | 1 |
| 1.2 ELPs..... | 7 |
| 1.3 Molecular Biology..... | 9 |
| 1.4 Educational overview..... | 20 |
| II. OBJECTIVE..... | 24 |
| III. MATERIALS AND METHODS..... | 27 |
| 3.1 Educational protocol summary..... | 27 |
| 3.1.1 Reagents and equipment to be purchased..... | 27 |
| 3.1.2 Educational protocol summaries..... | 32 |
| 3.1.2.1 Introduction to pipetting and density..... | 32 |
| 3.1.2.2 Pipetting food dyes..... | 33 |
| 3.1.2.3 Overnight culture and midiprep..... | 34 |
| 3.1.2.4 Quantification of plasmid..... | 35 |
| 3.1.2.5 Double digestion and assembly of nucleotides..... | 36 |

| | | |
|------------|---|-----------|
| 3.1.2.6 | HiFi DNA assembly..... | 38 |
| 3.1.2.7 | Bacterial transformation..... | 39 |
| 3.1.2.8 | (Optional) Colony PCR..... | 41 |
| 3.1.2.9 | Sequencing..... | 42 |
| 3.2 | Experimental design..... | 43 |
| 3.2.1 | Gene design..... | 43 |
| 3.2.2 | Overnight culture and DNA prep..... | 45 |
| 3.2.3 | DNA quantification..... | 46 |
| 3.2.4 | Plasmid and insert preparation..... | 47 |
| 3.2.5 | DNA assembly..... | 48 |
| 3.2.6 | Colony PCR..... | 48 |
| 3.2.7 | Step Replacement for Recursive Gibson Assembly..... | 50 |
| 3.3 | Digestion and ligation specifics..... | 51 |
| IV. | RESULTS AND DISCUSSION..... | 55 |
| V. | CONCLUSION..... | 64 |
| | BIBLIOGRAPHY..... | 66 |
| | APPENDIX..... | 75 |
| | Appendix A..... | 75 |
| | DNA sequences..... | 75 |

LIST OF TABLES

| Table | Page |
|---|-------------|
| Table 3.1 Classroom equipment and non-reagent consumable summary..... | 28 |
| Table 3.2 Classroom reagent summary..... | 31 |
| Table 3.3 NEB educational course support reagents..... | 32 |
| Table 3.4 Introduction to pipetting and density..... | 32 |
| Table 3.5 Pipetting food dyes..... | 33 |
| Table 3.6 Overnight culture and midiprep..... | 34 |
| Table 3.7 Quantification of plasmid..... | 35 |
| Table 3.8 Restriction digestion and assembly of oligonucleotides..... | 36 |
| Table 3.9 HiFi assembly..... | 38 |
| Table 3.10 Bacterial transformation..... | 39 |
| Table 3.11 PCR colony screening..... | 41 |
| Table 3.12 Sequencing..... | 42 |
| Table 3.13 Constructs generated..... | 45 |
| Table 3.14 Oligonucleotide resuspension..... | 48 |
| Table 3.15 Colony PCR run parameters..... | 49 |

LIST OF FIGURES

| Figure | Page |
|--|-------------|
| Figure 1.1 Decorin association with collagen fibrils..... | 4 |
| Figure 1.2 Thermoreversibility of ELPs..... | 8 |
| Figure 1.3 Sample PCR reaction first cycle..... | 13 |
| Figure 1.4 SpeI and XbaI isoschizomers,,..... | 14 |
| Figure 1.5 BioBrick assembly..... | 15 |
| Figure 1.6 Sample BsaI cut | 16 |
| Figure 1.7 Golden Gate assembly..... | 16 |
| Figure 1.8 Gibson assembly..... | 18 |
| Figure 1.9 Recursive Gibson assembly..... | 19 |
| Figure 2.1 Collagen-decorin-ELP hydrogel schematic..... | 25 |
| Figure 2.2 Cys-V(E(VEV) ₁₂) ₂ L ₂₀ hydrogel schematic..... | 26 |
| Figure 3.1 Oligonucleotide and recursive Gibson assembly ligation..... | 44 |
| Figure 3.2 PCR colony screen sample gel..... | 50 |
| Figure 3.3 NdeI and PflmI double digestion..... | 52 |
| Figure 3.4 Specifics of Gibson assembly system..... | 52 |

| | |
|--|----|
| Figure 3.5 PflMI single digestion..... | 53 |
| Figure 3.6 NdeI and BglI double digestion for recursive Gibson assembly insert..... | 54 |
| Figure 4.1 HiFi assembly and traditional ligation workflow sample..... | 60 |
| Figure 4.2 Safety, time cost, reagent cost, and equipment for molecular biology protocols..... | 60 |

CHAPTER I

INTRODUCTION

1.1 Hydrogels

Hydrogels are a class of materials that consist of crosslinked networks of hydrophilic polymers that absorb water to varying degrees. These hydrogels can absorb large amounts of water (reports of over 6500 times the weight of the polymer solution¹) and facilitate passive transport of substances when they are in the swollen state. In the case of contact lenses, hydrogels allow for the diffusion of water and dissolved gasses between the surface of the eye and the exterior environment. In wound dressings they permit the transmission of antibiotics and removal of toxins. There are many sources for the polymers that make up hydrogels, both natural and synthetic. Natural hydrogels have compatibility advantages, however, many synthetic hydrogels outperform them in swelling capacity, crosslinking methods, and mechanical strength². Their polymeric

networks are joined together by crosslinks, which are either physical or chemical in origin. Beyond these characteristics, there are many other means by which to categorize these hydrogels, including but not limited to polymeric composition, physical appearance, crystalline structure, and electrical charge.

Natural polymer constituents of hydrogels include many classes of biomolecules like proteins, polysaccharides³, and polyesters⁴. Natural polymer usage in hydrogels has the advantage of biocompatibility leading to a lower chance of toxicity or inflammation in biomedical applications, as well as being capable of synthesis within organisms to ensure high reproducibility of specific polymeric structures. However, they may not be as absorbent as many of their synthetic counterparts, as well as a lack of mechanical strength, as these properties can be tailored for a specific purpose in a synthetic hydrogel⁵.

Hydrogel crosslinking methods

Crosslinking of a hydrogel is the means by which the polymer components are held together, and are defined by either passive entanglements (physical) or permanent binding of polymers (chemical)⁵. Physical crosslinks are reversible, and are generated by the entanglements of chains or by hydrogen bonding. With these entanglements, their crosslinking is generally weak and susceptible to quick decay, unless special precautions are taken. Chemical crosslinks are added using a crosslinking agent or set of reaction conditions, which change the molecular structure of the polymers to bind chains together in an irreversible way. There are many different means to crosslink, from radiation to

introduction of free radicals. However, they typically involve the addition of a toxic agent to the hydrogel which must be thoroughly purged (if used in medical/consumer applications) and/or results in modification of the properties of the hydrogel (such as cellular adhesion, pore size, or degradation rate)⁶. Both means have their advantages, and hydrogels can be tailored to many different applications to produce the desired interactions.

An example of physical crosslinking would be the association of decorin with type I collagen. Decorin is a proteoglycan (protein with multiple sugar groups attached to it) that has an affinity for the highly prevalent protein⁷ collagen type I, which is a protein that is associated with cellular matrix stability and cell-to-surface interactions. Being a large protein, it has multiple binding sites and interacts with different biomolecules of significance. However, various sites on the proteoglycan have different affinities for collagen and it was determined that collagen binding occurs only on the inner surface of the decorin molecule. One specific sequence of amino acids in particular (LRELHLNNN), exhibited much slower diffusivity in a collagen gel compared to a hyaluronic acid gel⁸, indicating that binding to collagen molecules by this particular sequence generates a strongly crosslinked network. Further, the rheological properties of a di-block version of this amino acid chain exhibited storage and loss moduli nearly ten times higher than the control when stress was applied to pure collagen and a collagen gel infused with the di-block LRELHLNNN. This effect here is physical crosslinking across the gel, and this isolated sequence of decorin shows promise as a powerful crosslinking agent within the scope of collagen gels. Proteins like collagen are highly abundant within the body, but have the drawback of being targeted by digestive collagenase enzymes, like

all other proteins within an organism susceptible to their specific digestive enzymes. Work has been done to improve the resistance of these proteins by addition of nanoparticles⁹ to increase the life of a collagen hydrogel for drug delivery purposes.



Fig. 1.1

Schematic of staggered type collagen I fibrils (orange) associating with decorin proteins (green)¹⁰. The inner site binds tightly to one fibril while the N terminus associates with an adjacent fibril¹¹ to form a collagen fiber.

Prolastins are another interesting class of hydrogel material studied previously. They consist of blocks of silk-like proteins (consisting of GAGAGS repeats) and elastin-like proteins (GVGVP repeats) and exhibit gel properties¹². Cappello and others studied gels consisting of blocks of these groups suspended in phosphate-buffered saline (PBS) which formed stable gels *in vitro*, and *in vivo*. They form physically crosslinked gels via hydrogen bonding branching from the various amino acid groups and are stable at 37⁰C for at least one week. A hydrogel consisting of these polypeptides has the benefit of extensive biocompatibility conferred by synthesis within a living organism and subsequent purification. Also, a polypeptide gel would possess distinct tuning capabilities to desired temperature¹³, pH¹⁴, and salt concentration of gelation¹⁵, as well as decay rate *in vivo*¹⁶. ELP hydrogels have been shown to be suitable for culturing human adipose derived adult stem (HADAS) cells¹⁷. HADAS cells were suspended in a purified ELP solution containing guest residues of alanine, glycine, and valine and held above the

transition temperature of the ELP. Incubated in chondrogenic medium, the experimental colonies showed an increase of matrix synthesis and decreased proliferation, as is normal with the function of chondrocytes.

Bioprinting

Another use for hydrogels includes their application as bioinks. Bioinks are biomaterials associated with cell printing and cellular scaffold development. Bioprinting resembles other 3D-printing techniques, however operates on the smaller cellular scale. Tissue engineering techniques such as cellular scaffold development is necessary within the field of bioprinting, and research is concerned with the development of geometries, materials, and fabrication techniques that will make the printing of organs and connective tissue a reality¹⁸. Traditional fabrication methods like electrospinning¹⁹ and fiber deposition are effective for establishing flattened 2D scaffolds but when a biomedical engineer is concerned with factors such as pore size and cellular adhesion on specific geometries, more precise control of material connection is necessary. There are multiple established techniques for bioprinting, which include inkjet-based bioprinting, laser-assisted bioprinting, pressure-assisted (extrusion) bioprinting, and stereolithography, each of which have their specific advantages and disadvantages²⁰.

Inkjet-based bioprinting involves the modification of an inkjet printer to print a cellular ink slurry using a non-contact nozzle. It has the advantage of being inexpensive and minimally prone to contamination²¹ but may be limited to certain cellular types that can survive the heat and mechanical stresses generated by printing²². Laser-assisted

bioprinting is a technique using a laser focused on a glass slide containing cells suspended in a hydrogel solution. The laser generates a pulse that creates a bubble forcing cells to erupt from the cell suspension onto a collection slide laden with a hospitable hydrogel layer.²³ The advantage of this technique is that high viscosity hydrogels may be used in the preparation of the cellular suspension, as well as being able to deposit cells with high precision²⁴ Potential disadvantages include the high price of the laser-based setup and the potential destruction of cells by high heat, but no studies observed as of yet demonstrate any significant cell death. Extrusion-based bioprinting utilizes a system similar to traditional polymer extrusion in manufacturing, in which a mechanical screw or piston dispenses a constant flow of bioink onto a substrate. It is capable of full automation, being directed by CAD file to print a designed structure.²⁵ Advantages include the ability for direct control over flow rate through the nozzle and the ability to work with higher viscosities¹⁸. However, this technique may result in lower accuracy and result in shearing of cells across the nozzle.¹⁸ A stereolithographic bioprinter builds an object layer-by-layer, using light as the means to solidify the bioink used.²⁰ One advantage of this technique is that it is low cost for the resolution of printing that can be done²⁶. However, the fact that it is crosslinked by light limits its effectiveness to only photosensitive bioinks.

When developing bioinks for bioprinting applications, biocompatibility of these hydrogels is of prime concern, as the cells used in these printing techniques will be cultured for implementation *in vivo*. The natural origin of the scaffold will aid in cell adhesion, as cells are more capable of binding to an amino acid based structure with natural bioactivity compared to a synthetic hydrogel, such as poly(ethylene glycol)

(PEG)²⁷. Pore size of the cellular scaffold is also significant, as the pores generated within the structure will have an effect on cell density, cell attachment, and bulk compressive modulus²⁸, as pores within the matrix will allow for bulk transport of nutrients, waste, and oxygen to and from the cells. The ability of these networks of bioinks to maintain their structure under mechanical stresses is essential for developing a scaffold capable of supporting cellular outgrowth. Mechanical strength of the scaffold has been studied, as different elastic moduli can change how well a cell lineage adheres to a substrate, such as a hematopoietic cell line that binds well to a hydrogel²⁹. Cells in that study were found to adhere in more abundance to a more crosslinked polyethylene glycol diacrylate (PEGDA) hydrogel than to a softer and less crosslinked PEGDA gel.

1.2 ELPs

Elastin-like polypeptides (ELP) are synthetic polymers that consist of repeats of amino acids Gly-Xaa-Gly-Val-Pro, where Xaa is any amino acid except for proline and exhibit environmental responsiveness³⁰. This sequence is found in elastin, which is a naturally existing protein in mammalian connective tissue, allowing the tissues to return to their original form following stress. ELPs are of interest for their inverse transition temperature (T_i), which is a metric for polymers exiting the soluble phase and entering the insoluble phase. Based on the component amino acid and the size of the block of ELP, the transition temperature will vary among a wide range of temperatures. The transition is due to the ELP folding into a helical structure and associating with itself and other chains more than the solvent. This property is utilized for protein purification of these molecules or molecules tagged with them, as centrifugation above their transition

temperature (when they are aggregated) will result in an ELP-rich pellet, while a centrifugation below T_t will result in a ELP-rich supernatant³¹. The transition temperature of the ELP is also affected by the presence of distinct salts at varying concentrations, with multiple Hofmeister anions inducing different LCST behavior with one block of (GVGVP)₅³².

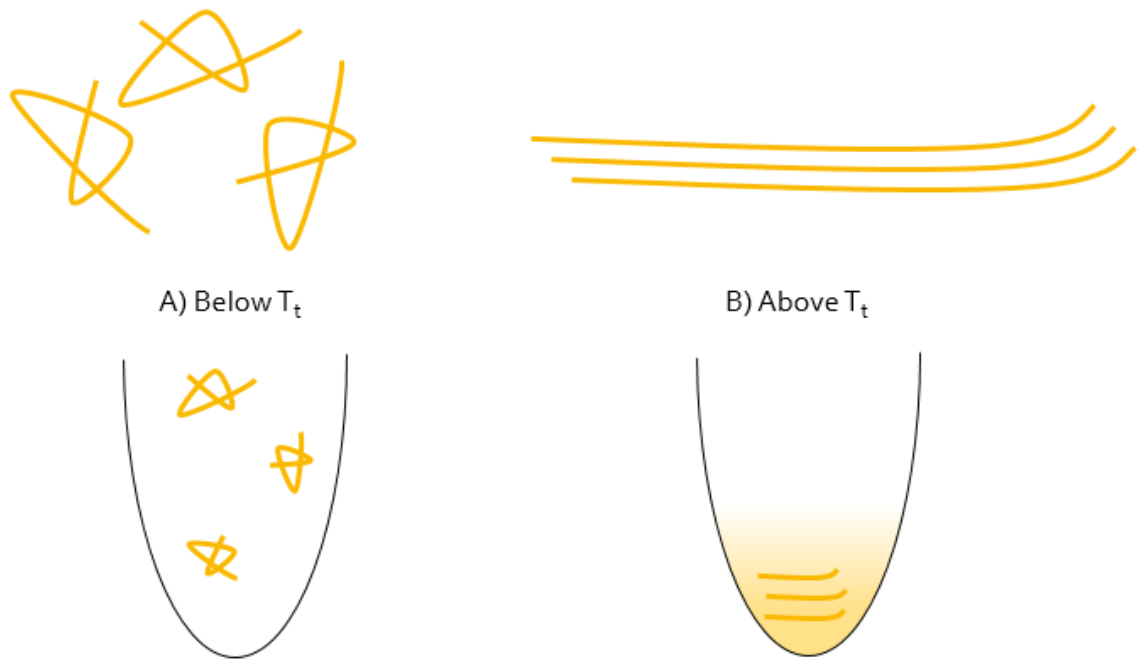


Fig. 1.2

Thermoreversibility of ELPs within solution, existing in a random coil state when (A) below the T_t and (B) aligning with each other when raised above it. Absorbance values for A will be lower than B.

Urry and others have devoted much time to the study of these elastin-like polypeptides, demonstrating that they exhibit many interesting properties when guest residues are swapped into the pentamer³³. Glutamic acid (E) is an acidic and hydrophilic amino acid, and when present as the guest residue in the form GEGVP and at low pH

(~3) it has a transition temperature near to 30°C ³³. This is in agreement with another studied construct, (VEV)₁₂, which has a T_t value near to 35°C at pH 2.7 and rises to 48°C at pH 3.8³⁴. However, when exposed to neutral pH (~7), its transition temperature increases drastically, approaching 250°C when E is the exclusive guest residue. Other residues have extreme hydrophobicity and extreme hydrophilicity, with tryptophan at -90°C and phosphorylated serine at a staggering 1000°C respectively. This is largely resulting from the increased (or decreased) exposure of hydrophobic moieties to the surrounding water of solution as well as movement of pK_a values³⁵ as a function of applied force. The ELP will never be fully separated from the solvent phase, but the formation of a dense coacervate phase will form with extended heating above the T_t of the mixture, averaging around 40% protein and 60% solvent within this dense liquid phase³⁶.

1.3 Molecular biology

Molecular biology is the field of biology that studies the molecular origin of biological function in conjunction with the major biomolecules of life. Under this umbrella is the central dogma of biology, which is an essential concept that details the pathway of how the genetic information of the cell, or deoxyribonucleic acid (DNA) becomes proteins that perform necessary cellular activities. The nucleic acids chains consists of five different molecules that contain a phosphoric acid, a deoxyribose sugar, and a purine or pyrimidine base³⁷. It is the difference in purine or pyrimidine that results in one of the four nucleic acids in DNA, which are cytosine (C), guanine (G), adenine (A), and thymine/uracil (T/U) (thymine is only present in DNA, uracil is present only in RNA). The length of the DNA polymer varies based on the species studied, however

naturally existing *E. coli* DNA is roughly 5000 kilobases in length³⁸, and has directionality. This directionality lends itself to the sense strand of DNA and is read from the 5' to 3' end, based on where the phosphate group binds to the adjacent nucleotide and the way that DNA polymerase processes the fragment. Each nucleic acid has a complementary nucleic acid in which two hydrogen bonds associate T with A; and three hydrogen bonds associate G with C. Nucleic acid triplets (referred to as codons) are associated with transferring a specific amino acid, of which there are twenty naturally occurring within the genome, and is conserved across all species. However, since there are 64 combinations available combining these bases, some combinations result in the same amino acid, and three are reserved as stop codons, which inform cellular machinery to cease adding amino acids to the protein.

DNA is wound tightly within the nucleus of the cell with the help of proteins called histones³⁹. When triggered to replicate, DNA polymerase proceeds to unwind the double helix and build complementary strands from the unwound pieces with free deoxyribonucleoside triphosphates (dNTP)³⁷. When expected to produce a protein, RNA polymerase proceeds to make a copy of the DNA piece by unwinding the helix of the DNA. The RNA polymerase binds to a specific region of the DNA, recognizing a specific piece of the sequence for a gene known as a promoter, and can begin to unwind the DNA and expose the region of interest. This is referred to as DNA transcription, and results in a single stranded piece of messenger ribonucleic acid (mRNA).

RNA is similar to DNA in structure, except for the difference of possessing a ribose sugar in the stead of the deoxyribose sugar, and the substitution of the thymine base for uracil. Three different types of RNA are used in protein synthesis, the other two

being transfer RNA (tRNA) and ribosomal RNA (rRNA), which carry amino acids and are components of ribosomes, respectively. The mRNA is transported out of the nucleus and into the cytoplasm⁴⁰, where it can associate with a ribosome. It is when complexed with a ribosome that mRNA is able to be read and translated into a protein. The 70S rRNA composed of its subunits act as the workbench for manufacture and the mRNA the blueprint for the construction, and tRNA containing individual amino acids act as raw material, recognizing each individual three letter codon group and adding the respective residue to the growing polypeptide chain. While being assembled, additional proteins may be present to assist in the folding of the protein, which may have little to no activity unless its advanced secondary or tertiary structure is present.

For protein engineering using recombinant DNA, a series of techniques are used to insert a desired sequence of nucleotides into an organism to produce a protein of a certain sequence.⁴¹ One starts with a bacterial plasmid or vector which has multiple regions of interest coded into it including for selection (antibiotic resistance, etc.), restriction enzyme cutting, and expression of proteins.⁴² Plasmid choice is important, as antibiotic resistances conferred by the plasmid, copy number, and insert size all play important roles in protein expression and cloning applications. Once chosen, one will cut the vector with restriction enzymes, generating a linearized vector which may be ligated back together with an insert. Inserts may be generated in many ways, through polymerase chain reaction (PCR), excision from another DNA sequence, or by obtaining oligonucleotides. Once an insert is prepared and vector is linearized, DNA ligase may be used to anneal the two ends of the insert to the two ends of the vector, returning it to its circular form. This recombinant plasmid is then forced into an organism (typically *E.*

coli or *S. cerevisiae*) through the process called transformation. Once the vector is transformed into the organism, the organism may be grown in media and will express the protein. In the instance of pET20b and many recombinant DNA systems, Isopropyl β -D-1-thiogalactopyranoside (IPTG) is used to induce expression. IPTG acts as a molecular mimic to a lactose metabolite, and activates the lac operon by removing a repressor that normally blocks transcription⁴³. The host organism is killed following sufficient growth of the protein, and purified out of the media/bacteria solution for use.

One can use PCR to amplify a specific piece of DNA from a source piece of DNA containing the desired sequence, and small pieces of single stranded DNA (ssDNA) called primers that would match exterior regions of the desired segment of DNA. Other standard reagents include a DNA polymerase, deoxynucleoside triphosphates (dNTP's), buffer, and magnesium salts for the PCR reaction. This reaction is done through roughly 35 cycles of amplification, each cycle nearly doubling the DNA yield of the previous cycle, ending with billions of the desired fragments. However, this is no longer a necessity with the currently applied system, which utilizes synthesized ssDNA called oligonucleotides purchased from Invitrogen. These oligonucleotides can be generated to have a nucleotide sequence desired (up to 100 bp in length), and this piece of DNA can be annealed to its complementary piece of ssDNA to produce double-stranded DNA (dsDNA) that is of similar quality to that generated by PCR technology. While not a direct replacement of PCR, annealing oligonucleotides skips the trial-and-error process typically associated with using a PCR protocol, and allows for a more robust and timely use of DNA for a high school curriculum.

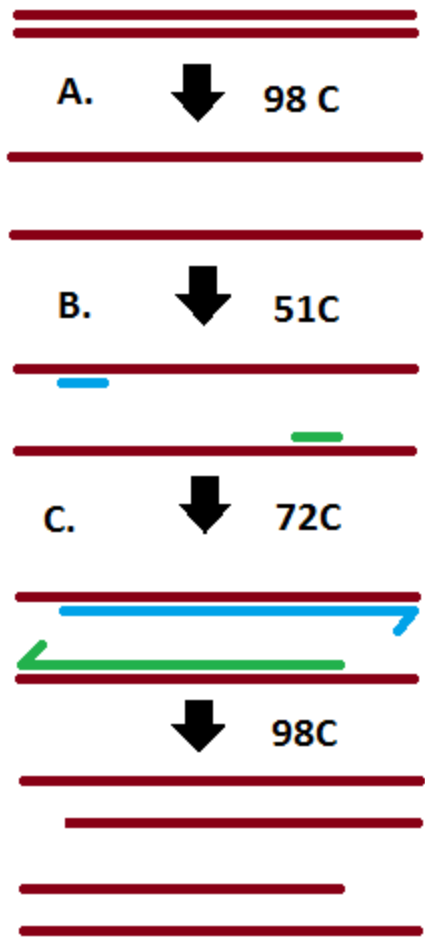


Fig. 1.3

Example of a PCR reaction. Starting with the fragment of dsDNA, during step A the DNA is denatured at high temperature. At B, the primers for the respective strands bind to their sites, and begin to build with DNA polymerase at C, ending with two pieces of dsDNA, and is repeatable for the desired number of cycles.

Molecular biology ligation advancements

Many other techniques and advancements in ligation methods have come to use in the last twenty years, many of which still use restriction endonucleases and some that no longer benefit from them. Some, like BioBrick and Golden Gate cloning, still benefit from the use of traditional endonucleases. Gibson assembly is a ligation method that

does not necessarily require the use of restriction enzymes (however they can be used in order to facilitate assembly).

BioBrick works to standardize the cloning technique, by having rule-oriented design for pieces in the assembly, such as only possessing upstream EcoRI and XbaI restriction sites and downstream SpeI and PstI sites for ease and universality of assembly⁴⁴. Promoters, reading frames, and terminators will all conform to these standardized preset pieces, allowing for ease of laboratory sharing of constructs. In order to conform to these standards, one must PCR in the respective cut sites to the 5' and 3' regions of the insert of interest. Once the cut sites have been added, one can digest the donor vector with EcoRI and SpeI to generate an insert, and the BioBrick recipient vector with EcoRI and XbaI to open the vector. The recipient vector must be purified prior to insertion, or recircularization of the vector may happen. These pieces will ligate together, generating a fusion protein possessing both the insert and the original fragment from the recipient.

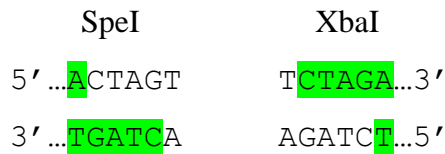


Fig. 1.4

Overhangs generated by SpeI and XbaI. Because they are isoschizomers, fragments digested in them ligate back together.

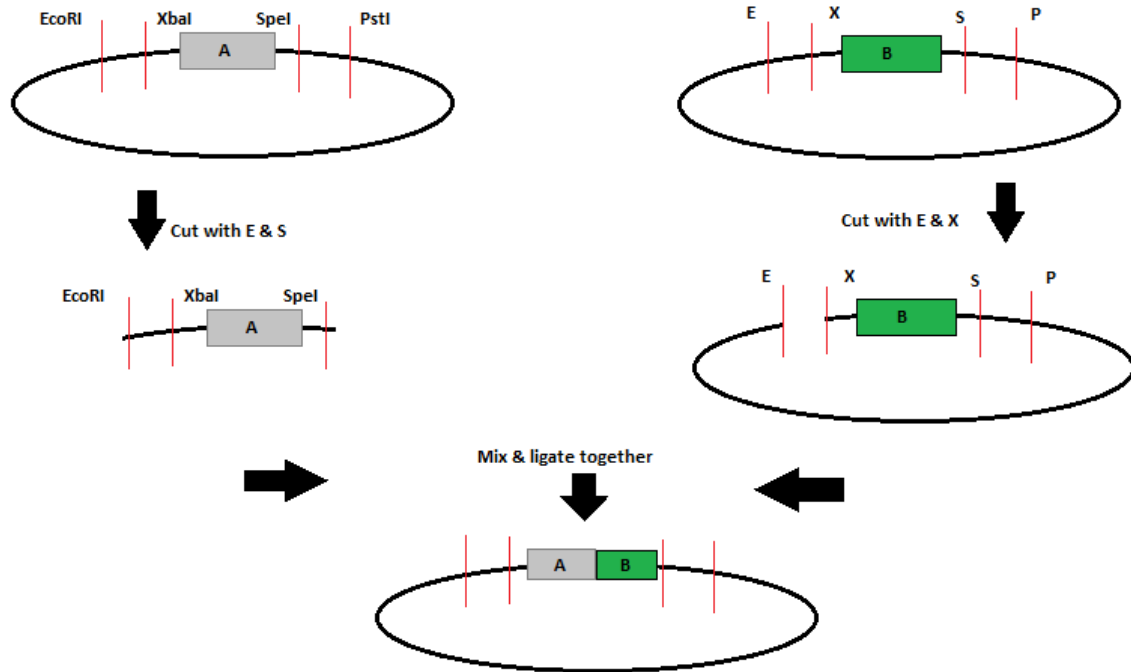


Fig. 1.5

System for annealing two fragments together within a BioBrick cassette

Golden Gate cloning exploits the activity of type II's restriction enzymes, which allow for the cleaving activity to occur outside of the recognition sequence⁴⁵. The reaction only requires a plasmid with two BsaI recognition sites and any number of inserts with similar BsaI sites, and can be carried out all at once with only the use of the BsaI restriction enzyme and T4 DNA ligase. Upon completion, only properly ligated plasmid will remain, as the last step of the assembly at 55⁰C favors digestion, and properly ligated segments delete the BsaI sites.

5' ...GGTCTCN^{NNNNN}...3'
 3' ...CCAGAG^{NNNNN}...5'

Fig. 1.6

Type II restriction activity of BsaI, which cuts downstream (green) of its recognition site (yellow)

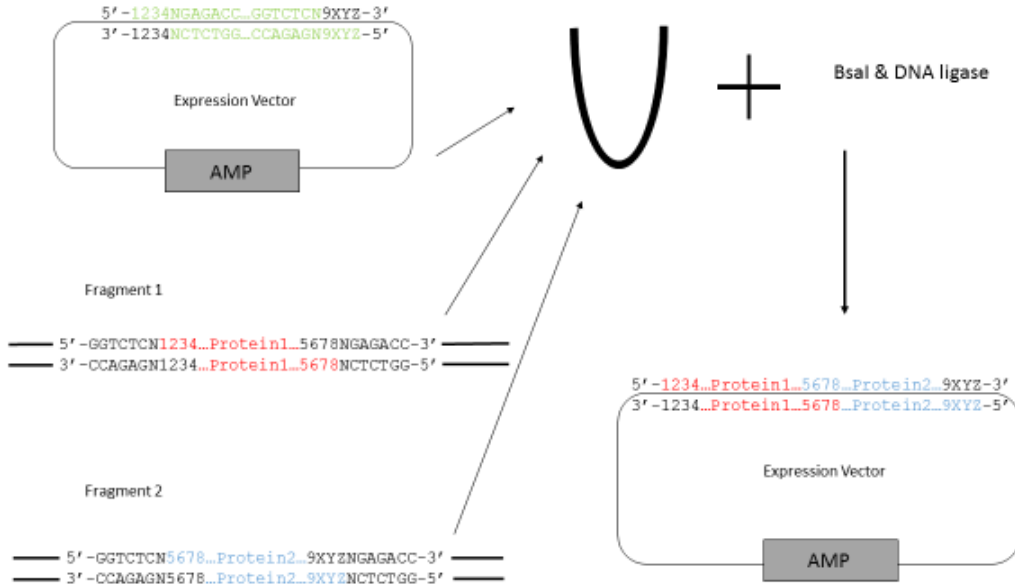


Fig.1.7

Mechanism of Golden Gate cloning occurring in a single reaction.

Gibson assembly is another system that allows for ease of use and does not necessitate the use of restriction enzymes (although they facilitate the portion of the DNA to be ligated). Previously, plasmids and inserted DNA had been digested with restriction enzymes and annealed together with classic ligases (T4, quick) which allowed only the

mating of fragments that had complementary end pieces (sticky ends). However, with this Gibson assembly system, pieces of DNA that have basic homology (fragments with similar ends) can be assembled and repaired into one piece of DNA in a single isothermal step, as the exonucleases that are used do not compete with the polymerase enzyme activity⁴⁶. Also, multiple fragments can be assembled at the same time into one plasmid, as long as each sequential addition has homology for only the next fragment, and the final fragment has homology with the downstream portion of the cut plasmid. Pieces assembled can be as large as 583 kilobases, and consist of a single reaction at 50 C simultaneously utilizing a 5' T5 exonuclease, Phusion DNA polymerase, and *Taq* DNA ligase⁴⁶. With these benefits, a multi-step ligation that does not require matching sticky ends as dictated by the limited selection of restriction enzymes can be the final product of this genetic engineering process.

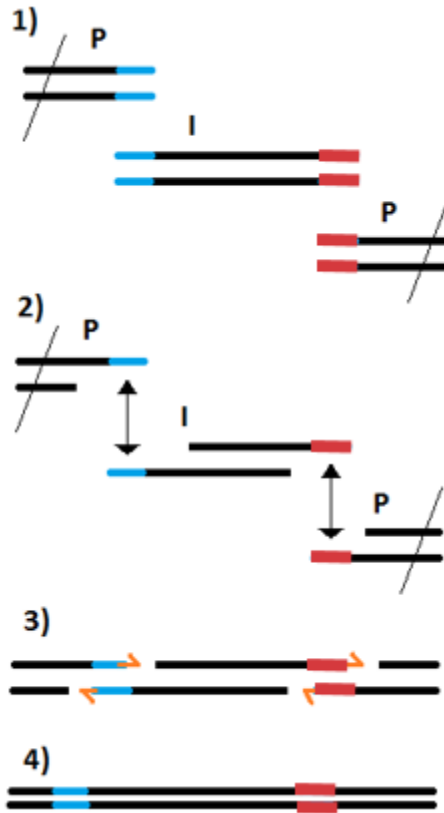


Fig. 1.8

1) Example of Gibson assembly activity annealing an insert into a linearized plasmid. 2) 5' exonuclease chews back 5' portion of dsDNA, and homologous 3' regions anneal. 3) DNA polymerase rebuilds in the 5' -> 3' direction, and a DNA ligase fills in gaps caused by exonuclease. 4) Insert has circularized the plasmid.

In Gibson assembly systems, one can anneal one or multiple inserts at the same time into a plasmid. However, when designed in a specific way, one can design a Gibson assembly system, in which a specific insert can have its length increased through successive rounds of cloning. When this system is applied, the length of the encoded repetitive polypeptide chain can be doubled with every successive round of cloning. To design a system like this, one will need two different plasmid inserts, both with homology at the 5' region and another site on the plasmids, as well as two unique restriction enzyme

sites near to the 5' end and one near to the 3' end. For this purpose, both had 3' homology in repetitive sequences of VEV ELP pentapeptides deeper in the sequence, as well as PflMI and NdeI cut sites near the start of the sequence and a BglI cut site near the end. Piece A will be removed from its plasmid by means of an NdeI and BglI double digestion, which cuts at the 5' homologous region as well as the 3' region still within the homology but nearing the end of the coding region of the protein. Piece B will be singly digested with PflMI, which has homology with piece A at the 5' region of the cut and also has the beginning portion of the repetitive VEV pentapeptide sequence on the 3' region of the cut. When assembled with the Gibson reaction, both homologous regions assemble, ligating the entirety of piece A into the 5' region of piece B.

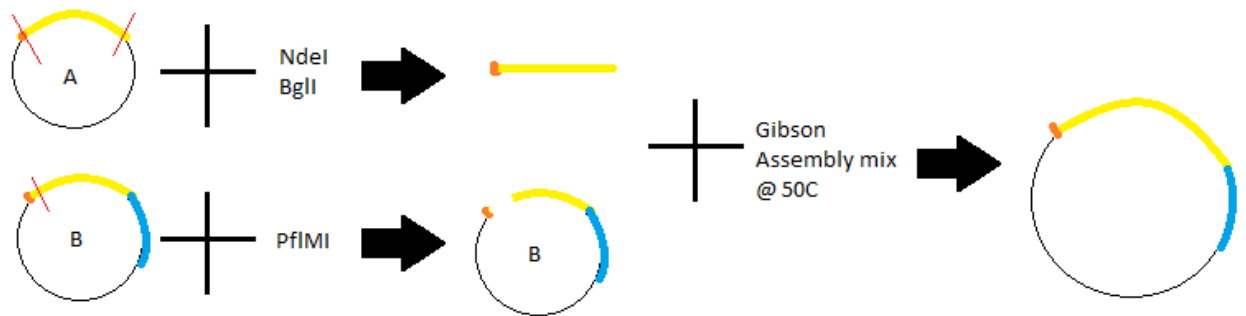


Fig. 1.9

Demonstration of the Gibson Assembly ligation system, where orange is the oligonucleotide induced 5' homology, yellow is the repetitive (VEV)₁₂ sequence, and blue is a different ELP region with different properties.

1.4 Educational Overview

As a whole, molecular biology and all other life sciences are fast-growing fields, with vast depths of knowledge necessary for success, and many branches that are highly cross-disciplinary ranging from philosophy to physics. Due to the diversity of this coursework, additional earlier teaching will be beneficial to preparing students for the career paths of this field. Focusing on the methods and skills associated with lab technique will be of primary concern when educating these students, encouraging them to not only act like a scientist, but to think and communicate like one as well⁴⁷. Five core concepts have been suggested by Vision and Change in Undergraduate Biology Education, which are: 1) evolution; 2) pathways and transformations of energy and matter; 3) information flow, exchange, and storage; 4) structure and function; and 5) systems⁴⁷.

One primary issue that can keep this subject from reaching the masses of students wishing to develop this skill is the large barrier-to-entry of knowledge between the experts and the laymen. This field incurs thought on many different size scales, from the angstrom level of atomic interactions to the macroscopic features readily visible by phenotypic expression in populations. Few professional fields deal with such a wide range of topics, and at varying levels of complexity. A large portion of this gap is associated with the diction of the trade. There are numerous analogies and similes available for describing molecules that resemble each other (left handed vs right handed glove, mirror images), but to understand what chiral molecules are and what separates a diastereomer from an enantiomer may require more than a simple metaphor. This is

further challenged by the structure-function relationship that has been established as a core-concept in biology and biochemistry education⁴⁷.

Growth of the field

Another large gap in the education of high school students is the breadth of information that is rapidly expanding faster than educational techniques for teaching these concepts. With the vast field of bioinformatics and information technologies allowing for much faster processing of entire genomes, protein structures, and expression profiles, it is difficult for teachers to keep up to date on what is the most current information on relevant topics. Further, with these quickly growing fields, pinpointing exactly what to teach may become muddled as topics increase in sophistication and what were commonly accepted as accurate dogmas may change with new advances. Determining what to teach a student will also vary based on what level of mastery they need to attain to enter the field of their choice, be it more basic or advanced.

With the advent of advanced computational technologies, data that had previously been collected and tabulated by hand is now readily processed in seconds with software packages. These programs allow for vast sums of collected values to be displayed in a concise way, which is a great tool for those initiated to the field but a blockade to a novice trying to interpret such figures. The problem with representing these complex systems is that their model must be equally complex.

Attempts are being made to offset these difficulties associated with education of high school students, by various interested agencies. Through funding by the NSF, the

American Society for Biochemistry and Molecular Biology (ASBMB) have established multiple resources for the lecturer tasked with teaching these difficult subjects, ranging from concept inventories for scientific inquiry to attitude inventories to assess learning and appreciation of course material^{48,49}. More and more websites from sources like the American Society for Cell Biology and the National Association of Biology Teachers are also springing up regularly with resources for relating difficult-to-teach concepts to students, including hands-on activities and analogies for lectures.

The primary reason for establishing a curriculum of protein engineering is to expand on the learning style of students. In a traditional laboratory experience, students will perform a series of steps from a lab book that will generate a specific result at the end if all stages of the experiment are followed correctly (referred to in literature as “cookbook” laboratories)⁵⁰. When offered an authentic research experience where results are not known, students are forced to think like scientists. That is to say, they will analyze results from a set of experiments and attempt to draw conclusions from them. The student will be then tasked to communicate what happened by their hand, and have the potential of becoming a published author by merit of their novel construct^{47,51}. An effective laboratory experience will not only result in an answer, but will inspire the student to ask questions, formulate their own future experiments, and effectively communicate the result to others, which this protein engineering course offers.

The scope of high school student polypeptide design is not restricted to only hydrogel structures. Other ELP types that will benefit from increased pentapeptide repeats include drug delivery vehicles and antifreeze protein applications. Micelle formation from ELP repeats form the basis of one type of nanoparticle that carries the

potential to be “loaded” with chemical structures when brought above their T_i . The carry volume and temperature at which this occurs can be modulated by multiple condition changes including salinity, pH, and chain length^{52,53}. Antifreeze proteins (AFP) are proteins which lower the temperature at which ice freezes and are isolated from many species of fish, plants, and insects. Trimerized constructs of fish antifreeze proteins have been studied and shown to be more effective than AFP monomers⁵⁴, but the study of trimerized insect AFP is limited, and so the attachment of foldon domains to these insect AFP present another student research opportunity.

CHAPTER II

OBJECTIVE

The objectives of this project were to design a new potential hydrogel forming polypeptide using protein engineering techniques. Sequentially progressing through rounds of cloning allow one to also generate constructs with different properties by exploiting our newly established protocol using the Gibson assembly system. Proven here is the effectiveness of such a system, by using an oligonucleotide insert for the first round to introduce homologous sequences on constructs. One such construct generated possess both trimerizing (foldon) and collagen-binding domains. Following this, entire constructs can be excised and fused together to generate larger proteins that will potentially demonstrate different gelation properties.

The proteins designed (Decorin(V)₁₉-foldon, Decorin(V)₃₇-foldon, and Decorin(V)₇₃-foldon) have the potential to bind strongly to collagen in solution, utilizing the strongest binding leucine repeat region of the Decorin protein⁸. The trimerizing

foldon domain is derived from a segment of fibrin from T4 bacteriophage⁵⁵, and is a sequence of 27 amino acids that associates with two other foldon domains to intertwine into a ribbon-like structure⁵⁶. Another generated construct is Cys-V(E(VEV)₂₄)L₂₀, which benefits from physically crosslinking hydrophilic leucine-rich GLGVP blocks into micelles and generated di-sulfide bridges from cysteine-cysteine interactions⁵⁷.

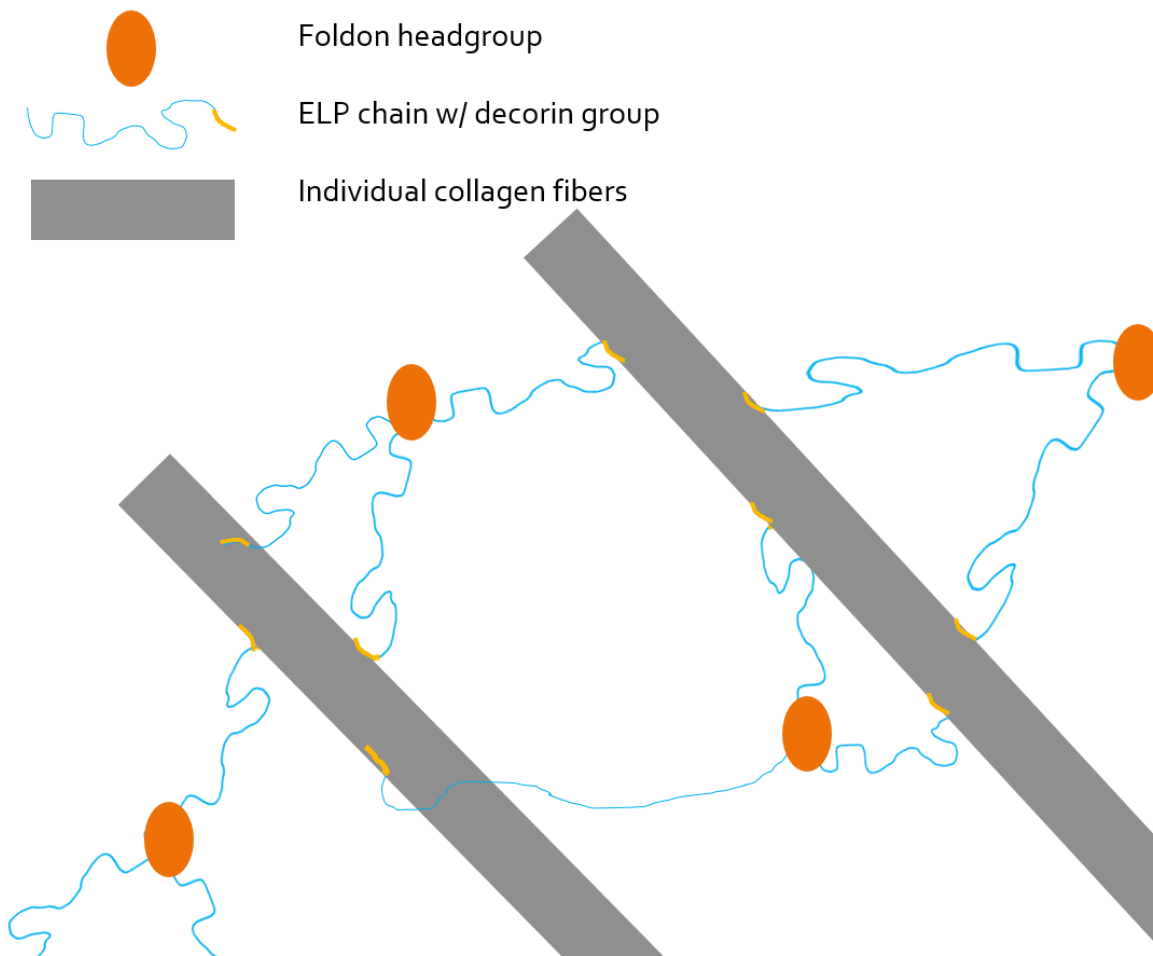


Fig. 2.1

Collagen-ELP hydrogel proposed activity. The decorin domain binds parallel to collagen fibrils

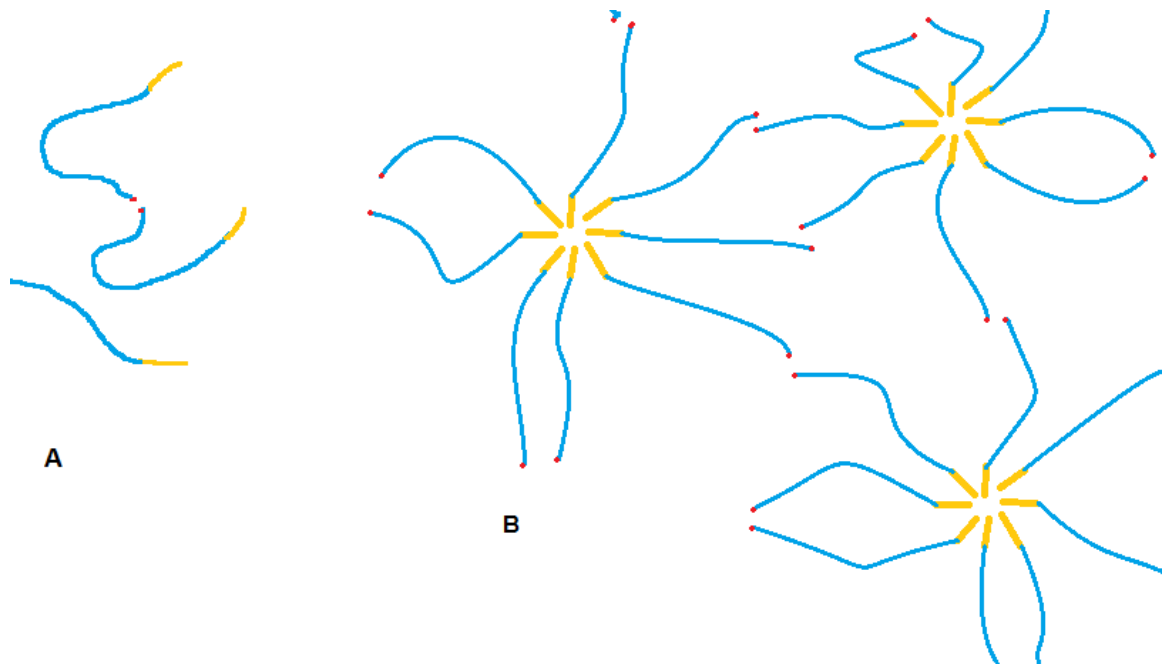


Fig. 2.2

Proposed association between chains of $(VEV)_{24}L_{20}$ ($(VEV)_{12}$ featured in blue, L_{20} in gold and cysteine groups in red). A) Below the T_i of the polymer chain, these chains are soluble in solution and do not associate strongly. B) Above T_i these chains are no longer soluble in solution and associate with each other, the hydrophobic lysine groups gather while the hydrophilic glutamic acid-valine regions associate with the solution. The cysteine groups provide a passive form of natural crosslinking.

Beyond this, the protein engineering steps were designed specifically with high school education in mind, in order to teach the students basic biotechnology skills.

Following the protocols described within this thesis, students with minimal biotechnology backgrounds were able to successfully accomplish both oligonucleotide and modular Gibson ligation cloning.

CHAPTER III

MATERIALS AND METHODS

3.1 Educational Protocol Summary

3.1.1 Reagents and Equipment to be purchased

To teach high school students this curriculum, certain equipment must be purchased, however some may be borrowed or shared with a higher learning institution, such that the cost of starting this class may be made more accessible to a lower budget institution. Table 3.1, 3.2, and 3.3 list are the bare minimum pieces of equipment necessary for setting up a lab.

| Equipment name | Serial number (Fisher Scientific) | Frequency of use | Cost |
|---|--------------------------------------|---|---|
| 13 ft ³ -86 ⁰ C Freezer | IU1386A | Store competent cells & frozen cultures | \$11835.19 new, best when shared with an institution |
| Gilson™ PIPETMAN™ G Starter Kit | F167900 | Every step will require pipetting | \$930.00 new, refurbished are okay with calibration |
| Sorvall Legend XTR benchtop centrifuge | 75-210-063 | Used for midiprep protocol (need rotor that can spin 50 mL tubes spin at 15,000g) | \$13,245 new, best when shared with an institution |
| Pressure cooker | N/A | Used as replacement for autoclave for sterilization of materials and equipment | Varies, around \$100.00 |
| T100 Thermal cycler | (Biorad) 1861096 | PCR unit for heating and cooling applications, colony PCR protocol | \$2495.00 with promotions, refurbished offered by manufacturer |
| Prep station enclosure | 15-338-365 | Used for all steps, this is where students will prepare all samples and perform all pipetting | \$2554.29 |
| Qubit 3.0 starter kit | Q33218 | Quantifies plasmid, used 3-4 times | \$2446.25 |
| High performance microcentrifuge | S35879 | Used for small centrifuging applications, and colony PCR protocol | \$1612.90 |

| | | | |
|---|-------------------|---|---------------------|
| Fridge/freezer | N/A | Storage of samples, reagents | Variable |
| Dry bath | 88-860-021 | Used as substitute for PCR unit, required for restriction digest and transformation | \$427.00 |
| Pureyield plasmid midiprep bundle 100 reaction | PRA6740 | Used twice per student per course | \$1086.73 |
| E-gel EX starter kit 1% | G6511ST | Used for colony screening, may not be necessary if confident in ligation results | \$997.50 |
| Multi-size tube rack 5PK | 03-448-15 | Holds 1.5 mL, 15 mL, and 50 mL tubes | \$37.17 |
| PCR tube cooler | 05-403-00 | Used w/ PCR steps, keeps samples chilled | 62.09 |
| Mini vortexer | 14 955 151 | Used every step that involves pipetting | \$179.4 |
| Incubator shaker | 07-202-157 | Used for transformation, can also be used for incubating plates overnight in the absence of a standard incubator | \$2624.16 |
| Erlenmeyer flasks 250 mL x 12 | FB500250 | Used for overnight cultures of midiprep | \$62.43 |
| Autoclavable glass media bottles 500 ml x 10 | FB800500 | Storage of DI water, LB media | \$87.99 |
| SimpleSeq Webless Sequencing Kit Premixed 96 reaction | Eurofins Genomics | Sequencing completed plasmids | \$384.00 |
| 0.1-20 μ L sterile filter tips 960/PK | 02-707-474 | Pipette tips are disposable after one use, sterile with a filter help to reduce contamination and prevent drawing of fluid into mechanical parts. | \$77.75 |
| 20-200 μ L sterile filter tips 960/PK | 02-707-478 | | \$77.62 |
| 200-1000 μ L sterile filter tips 560/PK | 02-707-480 | | \$62.48 |
| PCR tubes w/ attached cap 8/strp 250 strips | AB2000 | Used if thermocycler is being used | 312.00 |
| Serological pipet and pipet tips 500PK | 13683C S68228D | Disposable tips for pipetting large volumes | \$27.09 \$180.59 |

| | | | |
|-------------------------------------|------------|--|----------|
| Small/medium/large gloves 1000/PK | 191301597B | Every step | \$114.40 |
| | 191301597C | | \$114.40 |
| | 191301597D | | \$114.40 |
| 50 mL falcon tubes 500PK | 14-959-49A | Disposable tubes for up to 50 mL volumes | \$126.54 |
| 1.5 mL microcentrifuge tubes 500/PK | 05-408-136 | Disposable plasticware used regularly | \$12.62 |
| Petri plates x500 | FB0875713A | Used for transformation of bacteria | \$228.58 |

Table 3.1

Equipment and non-chemical consumables used within the high school curriculum

| Reagents | Fisher Serial Number | Frequency of use | Cost |
|--|----------------------|---|----------|
| E-gel Sample Loading Buffer, 1X 4x1.25 mL | 10482055 | Used for loading gels for the colony PCR | \$50.96 |
| Sodium chloride 1 kg | BP358-1 | Used for media and bacterial plates | \$55.21 |
| Ampicillin 5 g | BP1760-5 | Used for media and bacterial plates | \$77.48 |
| Yeast extract 500 g | BP1422-500 | Used for media and bacterial plates | \$137.48 |
| Peptone 500 G | BP1420-500 | Used for media and bacterial plates | \$126.27 |
| LB agar 500 g | BP1425-500 | Used for making bacterial growth plates | \$138.96 |
| Ethanol 96% 1000 mL | BP82021 | Used to prepare midiprep reagents, also for sterilization | \$41.91 |
| Isopropanol 500 mL | BP2618500 | Used to prepare midiprep reagents | \$18.28 |
| S.O.C. Media 10 mL x 10 | 15 544 034 | Used once per student per course | \$100.62 |

Table 3.2

Consumable reagents used within the high school curriculum.

| Reagents provided by New England Biolabs Education Course Support Program: | Inventory Number | Purpose |
|--|------------------|--|
| NEBuilder HiFi DNA Assembly Master Mix | E2621S | Ligation of homologous DNA fragments |
| Q5 Hot Start High-Fidelity 2X Master Mix | M0494S | PCR mix containing dNTPs, MgCl ₂ , DNA polymerase |
| NEB 5-alpha Competent <i>E. coli</i> (High Efficiency) | C2987I | Cloning strain for bacterial transformation |
| NdeI | R0111S | Restriction enzyme |
| PfIMI | R0509S | Restriction enzyme |
| BglI | R0143S | Restriction enzyme |

Table: 3.3

Reagents obtained from New England Biolabs Educational Course Support Program

3.1.2 Educational Protocol Summaries

3.1.2.1 Introduction to pipetting & density

| Safety concerns: | Setup time: | Equipment and reagents: |
|---|-------------|---|
| Gloves will need to be worn, whenever handling pipettes | 5-10 min | Micropipettes & tips Analytical balance 1.5 mL tube rack and tubes Beaker of water |

Table 3.4

Safety concerns, setup time, and equipment/reagents consumed during the introduction to pipetting & density protocol.

During the first laboratory setting, students are introduced to micropipetting which is verified through the concept of density. Students will be tasked with pipetting a

specific amount of water into a tared tube on the analytical balance, and report the value back to the instructor. Taking turns, each student will provide data for a chart on the chalk board on their value to be added and the reading obtained from the balance. Once all students are finished with this, the instructor will query the students on what the density of water is, and back calculating from the mass measured with the balance, determine the volume they actually managed to pipette compared to the requested value. Further, the instructor can introduce the premise of standard error, and allow the students to calculate the error in their pipetting.

3.1.2.2 Pipetting Food Dyes

| Safety concerns: | Setup time: | Equipment and reagents: |
|--|-------------|---|
| Gloves will need to be worn, whenever handling pipettes Chemicals are used so proper safety precautions must be taken | 10-15 min | Micropipettes & tips Vortexer 1.5 mL tube rack and tubes 4 pack of food dyes |

Table 3.5

Safety concerns, setup time, and equipment/reagents consumed during the pipetting food dye protocol

This experiment will further test the ability of the student to follow a specific protocol and use the micropipettes correctly. Students will be tasked with adding various food dyes in various combinations to make a final solution that resembles the final product that the instructor has made. Using proper pipetting technique will yield tangible

results, as when a student does not follow the instructions precisely, the final result will not have the same color or volume as the demonstration tube.

3.1.2.3 Overnight Culture & Midiprep

| Safety concerns: | Setup time | Equipment and reagents: |
|--|--|--|
| Gloves will need to be worn, whenever handling pipettes Chemicals will be handled so proper safety precautions should be made Centrifuge and air compressor will be used so student will need to be briefed on safety precautions associated with using these Students will be working with live bacterial cultures | Media needs to be made (3 h), overnight cultures must be prepared (grown overnight) | Micropipettes & tips 50 mL falcon tubes Midiprep kit 1.5 mL tube rack and tubes Vortexer Serological pipettes 250 mL Erlenmeyer flasks Incubator shaker LB media + ampicillin Autoclave Centrifuge |

Table 3.6

Safety concerns, setup time, and equipment/reagents consumed during the overnight culture/midiprep protocol.

This is the first step for performing the protein engineering part of this course, and students will need to treat all steps with appropriate care. The instructor will prepare LB media with ampicillin at a 50 mL volume in an Erlenmeyer flask, and inoculate with a frozen stock of a bacterial construct that contains the plasmid that the students will work with. Grown overnight at 37°C and 250 rpm in the incubator shaker, the cultures will be removed from the shaker and centrifuged according to the midiprep protocol. Following

completion of this protocol, the students will possess purified bacterial plasmid that will serve as the framework for their genetic manipulation. Success rates should be moderate to high, as the midiprep leaves room for user error and will likely result in at least a small amount of plasmid for all students. Since this step requires minimal micropipette use, students still lacking in fundamental technique should be safe for this step until they can gather more micropipette practice.

Note: Ensure that you dry the binding column for at least an hour prior to elution; failure to do so may result in residual ethanol contamination preventing proper elution and interfere with sequencing data.

3.1.2.4 Quantification of plasmid

| Safety concerns: | Setup time: | Equipment and reagents: |
|--|-------------|---|
| Gloves will need to be worn, whenever handling pipettes Chemicals will be handled so proper safety precautions should be made, specifically a photosensitive dye which needs to be kept in away from light as much as possible Students will be working with live bacterial cultures | N/A | Micropipettes & tips Qubit fluorometer starter kit 1.5 mL tube rack and tubes Vortexer |

Table 3.7

Safety concerns, setup time, and equipment/reagents consumed during the plasmid quantification protocol.

In this protocol, students will quantify the plasmid that they have obtained from the previous step, so that they can set up a downstream reaction with the proper concentration of plasmid, and to determine if their midiprep worked correctly. Tasked with making a working solution of Qubit dye and buffer, the students will use fluorometry to test their midiprep sample and get a concentration of plasmid in their eluted sample in ng/ μ L format. This protocol calls for precise micropipetting, as well as making standards that will be the reference point of their quantification, and initially may require more supervision.

3.1.2.5 Double digestion and assembly of oligonucleotides

| Safety concerns: | Setup time: | Equipment and reagents: |
|--|---|--|
| Gloves will need to be worn, whenever handling pipettes Chemicals will be handled so proper safety precautions should be taken Hot block/thermocycler will be used so student will need to be briefed on safety precautions associated with using these Students will be working with live bacterial cultures | Oligonucleotides need to be resuspended to 100 μ M concentration (10 min), and hot block needs to ramp to 98 ^o C (~20 min), and denaturing of the restriction enzymes at 65 ^o C (overnight) | Micropipettes & tips Restriction enzymes Enzyme buffer 1.5 mL tube rack and tubes Vortexer Hot block/thermocycler Forward and reverse complementary oligonucleotides Quantified plasmid |

Table 3.8

Safety concerns, setup time, and equipment/reagents consumed during the double digestion and assembly of oligonucleotide protocol.

For this step, students will take their oligonucleotides, mix them in equimolar concentration with the respective reagents to get them to 2 μM , and heat them to 98 $^{\circ}\text{C}$ for 5 minutes, and then let them cool to room temperature for 45 minutes to allow them to anneal. During that cooling, the students can set up their restriction digests that are appropriate for their specific plasmid and insert. The instructor will take care to monitor that all reagents should be vortexed, except for the restriction enzymes, as it may result in their denaturing and not working any longer. Also, they need to be kept on ice as much as possible, for the same reason. This is another protocol that involves extensive pipetting, and attention to detail will give students the best outcomes. This should be successful, however, as long as all of the reagents are added to the restriction mixture and left to incubate overnight at 37 $^{\circ}\text{C}$.

3.1.2.6 HiFi DNA assembly

| Safety concerns: | Setup time: | Equipment and reagents: |
|---|---|---|
| Gloves will need to be worn, whenever handling pipettes Chemicals will be handled so proper safety precautions should be made Hot block/thermocycler will be used so student will need to be briefed on their safety precautions Students will be working with live bacterial cultures | Hot block need to be warmed up to 50°C (10 min) | Micropipettes & tips Annealed oligos Digested plasmid Hot block 1.5 mL tube rack and tubes Vortexer HiFi assembly mix |

Table: 3.9

Safety concerns, setup time, and equipment/reagents consumed during the HiFi DNA assembly protocol.

During the HiFi DNA assembly, the double digested isolated plasmid from the frozen stock receives the double stranded oligonucleotide insert. This is technically the final step of assembly, as the student will have their modified plasmid if they followed all of the steps correctly. However, downstream steps are needed to put the plasmid into a fresh strain of bacteria to replicate it, and to sequence the plasmid to ensure that it was assembled correctly. This is not a difficult step, it just requires attention to detail that all reagents are added, and as long as they are, this protocol should be fairly robust.

3.1.2.7 Bacterial Transformation

| Safety concerns: | Setup time: | Equipment and reagents: |
|---|--|---|
| <p>Gloves will need to be worn, whenever handling pipettes</p> <p>Chemicals will be handled so proper safety precautions should be made</p> <p>Hot block/thermocycler and incubator shaker will be used so student will need to be briefed on safety precautions associated with using these instruments</p> <p>Students will be working with live bacterial cultures</p> | <p>Hot block need to be warmed up to 42^oC (5 min), competent cells may need to be obtained and kept on dry ice until use.</p> | <p>Micropipettes & tips</p> <p>Incubator shaker</p> <p>Hot block</p> <p>1.5 mL tube rack and tubes</p> <p>Vortexer</p> <p>Competent cells (DH5α)</p> <p>Ampicillin LB plates</p> <p>S.O.C. Media</p> <p>Annealed plasmid</p> <p>Crushed ice</p> |

Table 3.10

Safety concerns, setup time, and equipment/reagents consumed during the bacterial transformation protocol.

At this point in the curriculum, the student has successfully modified their plasmid and is now performing the steps to ensure that it has assembled correctly. To do this, the plasmid is placed within a strain of bacteria that is does not have any innately conferred antibiotic resistance and is able to take up different pieces of DNA (the plasmids). This bacteria may either express large quantities of the plasmid (cloning strain) or express the proteins edited into the genetic code of the plasmids. Ampicillin or another antibiotic resistance is also coded for in the plasmid to allow for antibiotic plate-based selection of bacteria that have taken in the plasmid. Using a series of hot and cold

cycles, the bacteria take in the plasmid, and when presented with favorable growing conditions, are able to start expressing the ampicillin resistance gene before being placed onto ampicillin plates, and following an overnight incubation, observed colonies should contain the plasmid of interest. Positive and negative controls can also be imparted to this protocol, with the transformation of both cut and uncut unmodified plasmids, to test the effectiveness of the restriction enzymes and other steps followed up to the transformation. This is not a challenging protocol, just time consuming, and as such, should have relatively high success rates.

3.1.2.8 (Optional) Colony PCR

| Safety concerns: | Setup time: | Equipment and reagents: |
|---|---|--|
| <p>Gloves will need to be worn, whenever handling pipettes</p> <p>Chemicals will be handled so proper safety precautions should be made</p> <p>Hot block/thermocycler and incubator shaker will be used so student will need to be briefed on safety precautions associated with using these instruments</p> <p>Students will be working with live bacterial cultures</p> | <p>Preparation of 10 μM standard T7 primers for PCR (~10 min), boiling a beaker of water is necessary, as well as the wait for the PCR to finish (~ 2 h)</p> | <p>Micropipettes & tips</p> <p>Hot plates</p> <p>Thermocycler</p> <p>PCR tubes</p> <p>1.5 mL tube rack and tubes</p> <p>Vortexer</p> <p>Ampicillin LB plates</p> <p>E-gel EX unit and 1%/2% agarose gels</p> <p>E-gel 1kb+ ladder</p> <p>E-gel loading buffer</p> <p>Overnight transformed plates</p> <p>Q5 2x PCR master mix</p> <p>Forward and Reverse PCR primer</p> <p>DI water</p> <p>Microcentrifuge</p> |

Table 3.11

Safety concerns, setup time, and equipment/reagents consumed during the colony PCR protocol.

This is the only optional procedure of the entire protocol, but operates as another verification step to offer insight as to whether the construct was assembled correctly. This is quicker than the process of picking various colonies off of plates, midprepping them, quantifying them, and sending them to be sequenced, hoping for the best. Students will instead pick colonies from their transformed plates, inoculate a new plate with that colony as well as a PCR tube that will undergo a PCR reaction to amplify their fragment

of DNA that is in the primed region of the plasmid (standard T7 promoter and reverse primers work for the currently established system). Once amplified, the resultant PCR DNA can be run on a gel, and compared to a control of uncut unmodified plasmid, to determine if there was a visible size increase between the experimental and control. This process takes roughly three hours (although two hours are waiting on the PCR reaction) and yields information towards which colonies possess the correct plasmid such that unnecessary sequencing reactions are not performed. This is a pipetting-heavy protocol, and as such, it will be important to ensure that all components of reactions are put together correctly (PCR mix, E-gel buffer & PCR resultant).

The next two steps are repeats of the protocols for midiprep and quantification, which are 3.1.2.3 and 3.1.2.4 respectively.

3.1.2.9 Sequencing

| Safety concerns: | Setup time: | Equipment and reagents: |
|---|--|--|
| Gloves will need to be worn, whenever handling pipettes Chemicals will be handled so proper safety precautions should be made Students will be working with live bacterial cultures | Resuspending T7 primers to 2 μ M concentration (~10 min) | Micropipettes & tips Vortexer Sequencing primers Sequencing kit |

Table 3.12

Safety concerns, setup time, and equipment/reagents consumed during the sequencing protocol

The final protocol, sequencing is the step where the student likely receive a definitive result on the constructs that they have attempted to build. Ideally, the plasmid concentration for the sequencing reaction will be over 100 ng/ μ L, however one can receive clean sequence reads with values above 10 ng/ μ L. Once the primers and purified plasmid are mixed, the samples are sealed and sent out by UPS overnight to Eurofins Genomics, and results are typically emailed within 24 hours.

3.2 Experimental Design

3.2.1 Gene Design

The goal of this thesis is to establish a viable course outline that would give students a working knowledge of how molecular biology works and how to perform the techniques of protein engineering. Listed here is a detailed protocol of the steps involved from the start of project with a plasmid to the gathering of sequence data from a successful construct.

Starting this project, one must have an idea of what they would like to create. Gene design is a significant and challenging barrier to entry in molecular biology. For the purposes of this paper, we aimed to generate a tri-block ELP that would have amphiphilic properties, and had the potential to form a hydrogel when subjected to the correct salt and temperature conditions. Our initial gene design called for the installation of an oligonucleotide insert with a design that would induce homology for self and different nucleotide sequences. Doing this allows for the doubling in size of protein sequences by different single and double digestions around the region of homology.

This was accomplished with a front-flanking oligonucleotide region with a cysteine group and two GVGVP pentapeptides with a GEGVP in the center using HiFi assembly to add in this region by use of two annealed complementary oligonucleotides. Once this homologous region was added to the front of (VEV)₁₂ and (VEV)₁₂L₂₀ to generate the constructs Cys-EV(VEV)₁₂ and Cys-EV(VEV)₁₂L₂₀, we could proceed with the second round of cloning. The second round of cloning called for excising the entirety of the Cys-EV(VEV)₁₂ construct from its plasmid, and ligating it into the front of another Cys-EV(VEV)₁₂ and also the front of Cys-EV(VEV)₁₂L₂₀, we generated constructs that were able to be doubled in size with every round of ligation.

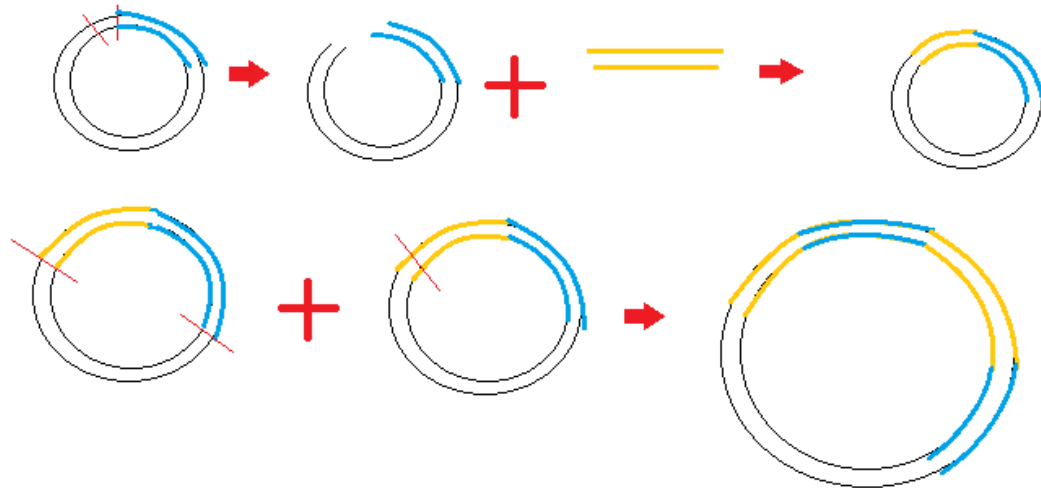


Fig. 3.1

Demonstrates of the ability to add a double stranded oligonucleotide to a double digested construct, and then utilize the homology generated by the oligonucleotide to ligate in the first constructed piece to the front flanking region of a single digested construct. Red lines are restriction enzymes, gold lines are oligonucleotide insert, and blue lines are original plasmid content

| Construct name | Amino acid composition | Starting material | Insert used | Restriction enzymes used |
|---|--|--|--------------------------|--------------------------|
| Cys-(VEV) ₁₂ | MGCNHGEGVPGVGVVP (GVGVPGEGVPGVGVVP) ₁₂ GWP | (VEV) ₁₂ | dsDNA oligonucleotide | NdeI, PflMI |
| Cys-(VEV) ₁₂ L ₂₀ | MGCNHGEGVPGVGVVP (GVGVPGEGVPGVGVVP) ₁₂ (GLG VP) ₂₀ GWP | (VEV) ₁₂ L ₂₀ | dsDNA oligonucleotide | NdeI, PflMI |
| Cys-(VEV) ₂₄ | MGCNHGEGVPGVGVVP (GVGVPGEGVPGVGVVP) ₂₄ GWP | Cys-(VEV) ₁₂ | Cys-(VEV) ₁₂ | NdeI, PflMI, BglII |
| Cys-(VEV) ₂₄ L ₂₀ | MGCNHGEGVPGVGVVP (GVGVPGEGVPGVGVVP) ₂₄ (GLG VP) ₂₀ GWP | Cys-(VEV) ₁₂ , Cys-(VEV) ₁₂ L ₂₀ | Cys-(VEV) ₁₂ | NdeI, PflMI, BglII |
| Decorin-(V) ₁₉ - foldon | MGHGVVPHLRELHLNKNKL(GVGVVP) ₁₉ GWPGYIPEAPRDGQ AYVRKDGEWVLLSTFL | E(V) ₁₉ - foldon | dsDNA oligonucleotide | NdeI, PflMI |
| Decorin-(V) ₃₇ - foldon | MGHGVVPHLRELHLNKNKL(GVGVVP) ₃₇ GWPGYIPEAPRDGQ AYVRKDGEWVLLSTFL | E(V) ₃₇ - foldon | dsDNA oligonucleotide | NdeI, PflMI |
| Decorin-(V) ₇₃ - foldon | MGHGVVPHLRELHLNKNKL(GVGVVP) ₇₃ GWPGYIPEAPRDGQ AYVRKDGEWVLLSTFL | E(V) ₇₃ - foldon | dsDNA oligonucleotide | NdeI, PflMI |

Table 3.13

Constructs generated and their starting material

3.2.2 Overnight Culture and DNA prep

Next, one must have a base genetic material to work with, and that is the plasmid generated from a frozen stock of bacteria. Frozen stocks of ELP constructs were grown in lysogeny broth (LB) media containing ampicillin for up to 24 hours. The amount of plasmid obtained from these overnight cultures is largely determined by the type of DNA prep kit to be used, i.e. whether a miniprep (Qiagen) or a midiprep (Promega). Both were

used for this project, but the ideal unit for the classroom setting is the midiprep, as it yields a much larger quantity of DNA from an overnight culture (100-200 μg) as compared to the miniprep which generates 20 μg or less. Minipreps were used for their cost-effectiveness, but when it comes to a classroom, the larger yield of DNA leaves room for student error when performing a DNA prep procedure. For the miniprep procedure, the volume of LB media was 10mL with a scraping of the frozen stock of interest added, and grown overnight in an incubator/shaker (G24, New Brunswick Scientific Co.) The spin-column protocol was followed for the miniprep kit, and the elution volume was 34 μL for the constructs. For the midiprep procedure, the volume of LB was 50mL, and the vacuum protocol was followed, with an elution volume of 400 μL .

3.2.3 DNA Quantification

Once the plasmid was successfully eluted, the quantity of DNA obtained was determined, using the Qubit 3.0 fluorometer. Following the instrument's protocol, 1 μL of DNA was used to the 199 μL Qubit working solution in one of the provided 0.5 mL Qubit tubes, and was compared against two standards as provided by the kit, using the high-specificity program. Once a concentration of DNA was obtained from the sample, DNA manipulation can be performed.

3.2.4 Plasmid and Insert Preparation

Restriction digest using restriction enzymes was performed to facilitate the process of our downstream assembly protocol. Using a specific quantity of DNA to generate a reaction concentration of 20 ng/ μ L, plasmids were single digested with PflMI (New England Biolabs) or double-digested with both NdeI (New England Biolabs) and BglI (New England Biolabs) or NdeI and PflMI in a 10 μ L reaction. This reaction consisted of 1 μ L 10X NeBuffer 3.1 (New England Biolabs), 1 μ L PflMI (or 1 μ L NdeI and 1 μ L BglI) and the remaining volume consisted of nuclease-free water (Promega). This reaction was performed for overnight at 37 $^{\circ}$ C in the thermocycler, and then the restriction enzymes were denatured at 65 $^{\circ}$ C for twenty minutes. The plasmid DNA is then left with two sticky overhangs. Care was taken to not destroy the restriction enzymes, as they are prone to unfolding and subsequent loss of activity following extended exposure to room temperature as well as through vortexing, so the reaction was mixed by pipetting the 10 μ L mixture up and down.

Oligonucleotides (Invitrogen) were obtained and annealed to form the insert DNA that would be put inside of the plasmid. First, the oligonucleotides were spun down briefly to collect all of the lyophilized material at the bottom, and were resuspended in nuclease free water to a concentration of 100 μ M. Following this, the oligonucleotides were annealed together at 95 $^{\circ}$ C for 5 minutes, and left to cool at room temperature in a reaction (Table 3.14). This annealed oligonucleotide is now ready to be inserted into the plasmid double-digested with NdeI and PflMI using the Gibson assembly protocol.

| Component: | Volume (μL) |
|----------------------------------|--------------------------|
| 100 μM forward primer | 10 |
| 100 μM reverse primer | 10 |
| 1M Tris (pH 8.0) | 1 |
| 50 mM MgCl_2 | 10 |
| Nuclease free water | 69 |
| Total volume | 100 |

Table 3.14

Reaction conditions for annealing complementary oligonucleotides.

3.2.5 DNA Assembly

The Gibson assembly protocol consists of the addition of a double-stranded oligonucleotide at the 5' region of a previously designed construct³⁴. The constructs to be modified included MGH-(VEV)₁₂-GWP and MGH-(VEV)₁₂L₂₀-GWP, where the single letter abbreviations signify the guest residue of the GVGXP format. Cut plasmids were mixed with the insert DNA at 2 ng/ μL and 0.2 μM concentrations respectively. NEBuilder HiFi Master Mix (2x) was added to this solution and nuclease-free water to a volume of 20 μL , and was held at 50⁰C for 15 minutes to allow for assembly. After this, the assembled construct was transformed into NEB 5-alpha competent *E. coli* cloning strain cells to incubate on an ampicillin-resistant plate overnight.

3.2.6 Colony PCR

Once the transformed bacteria was grown overnight, individual colonies were screened to ensure that the plasmid present in the colony was the assembled construct and

not undigested original plasmid. To perform this, individual colonies were stabbed with a pipette tip, used to inoculate a new plate, and mixed into a tube containing 25 μ L of DI water. These samples were boiled for two minutes, then centrifuged at 16000xg for 2 min. A 2 μ L aliquot of supernatant from these tubes was taken and used in a standard PCR reaction containing T7 promoter and terminator primers and Q5 Hot Start High-Fidelity 2x Master Mix (NEB) and was placed in a thermocycler (Table 3.15)

| Step | Temperature ($^{\circ}$ C) | Time | Repeat 34x |
|----------------------|-----------------------------|----------|------------|
| Initial denaturation | 98 | 30s | |
| Denaturation | 98 | 8s | |
| Annealing | 51 | 30 | |
| Extension | 72 | 60s | |
| Final extension | 72 | 120s | |
| Hold | 4 | Infinite | |

Table 3.15

PCR step times for Colony PCR with standard T7 primers

Once the cycles are completed, 5 μ L of PCR reactant is taken from the PCR tube and added to 15 μ L of E-gel loading buffer, loaded into a 1% or 2% E-gel agarose pre-cast gel, and run against E-gel 1Kb+ DNA ladder and an uncut, unmodified plasmid for comparison (Fig. 3.2). Successful constructs as determined by the correct number of base pairs relative to their parent plasmid are then grown from their respective stabs overnight

and are DNA prepped. Once the prep has been performed, plasmids from the constructs are submitted to Eurofins genomics for overnight sequencing.

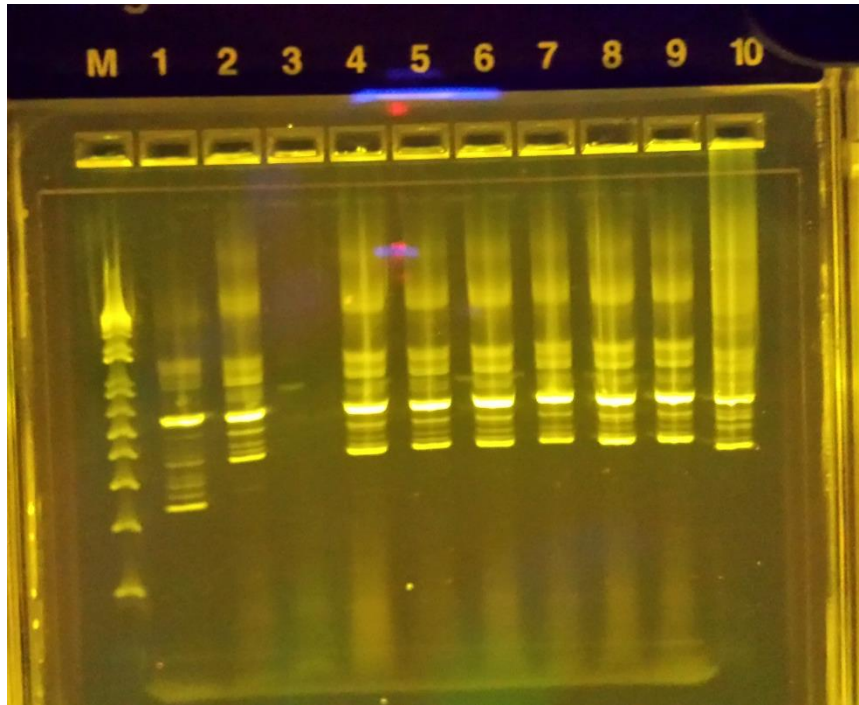


Fig. 3.2

Example of a completed PCR screening, ran on a 2% E-Gel agarose gel. Well M contains 1kb+ marker, wells 1-9 contain colony stabs, and well 10 contains unmodified control plasmid.

3.2.7 Step Replacement for recursive Gibson Assembly

Following the first successful round of DNA insertion, a second round was performed using the Cys-EV(VEV)₁₂ with homologous insert as a library of sorts, in that the entirety of that construct is capable of being excised from its parent plasmid and ligated into a single digested segment of the same construct or another construct (such as the Cys-EV(VEV)₁₂L₂₀) that still possesses the flanking homologous regions. For this protocol, all of the steps essentially remain unchanged up to the digestion, where Cys-

EV(VEV)₁₂ was double digested with both NdeI and BglI to remove the entirety of the construct, and the same construct or Cys-EV(VEV)₁₂L₂₀ was single digested with PflMI to expose the homologous segment that the two ends of the double-digested Cys-EV(VEV)₁₂ would ligate to.

3.3 Digestion and ligation specifics

DNA sequences were modified, and depicted below is the series of modifications performed on each construct. When performing a double digestion on (VEV)₁₂ and (VEV)₁₂L₂₀ for oligonucleotide insertion, a small fragment of roughly 10 bp in length was excised from the coding sequence (Fig. 3.3). This fragment did not need to be removed by any cleanup steps to prevent its reinsertion, as the 5' exonuclease present in Gibson assembly chews back the 5' region of DNA. When the removed fragment associated with exonuclease, it was broken down completely by the enzyme as it proceeds to catalyze DNA disassembly to expose homology for assembly.

TATACATATGGGCCACGGCGTGGGTGTT
 ATATGTATACCCGGTGCCGCACCACAA

+ PflMI & NdeI yields:

TATACA TATGGGCCACGGC GTGGGTGTT
 ATATGTAT ACCCGGTG CCGCACCCACAA

Fig. 3.3

Starting sequence of the (VEV)₁₂ and (VEV)₁₂L₂₀ constructs, and the action of double digesting with two restriction enzymes. Underlined regions are sites recognized by the restriction enzymes used, and red highlights are the overhangs generated by digestion.

After the double digestion and annealing of the single stranded oligonucleotides, the HiFi assembly activity performs the ligation reaction (Fig. 3.4).

AGAAGGAGATATACATATGGGTGTAACCAACGGTGTGGGAGTGCCGGGCGAAGGTGTTCCCTGGCGTGGGTGTTCCG
 TCTTCCTCTATATGTATACCCAACATTTGGTGCCACACCCCTCACGGCCCGCTTCCACAAGGACCGCACCCACAAGGC

+ HiFi Assembly & Double Digested vector yields:

5' exonuclease chews back 5' ends
 ACTTTAGAAGGAGATATACAGAAGGAGA TATACATATGG & TTCCTGGCGTGGGTGTTCCG GTGGGTGTTCCGGGGGAAGG
 CGAAATCTTCCTCTATATGTATCTTCCTCTATATGTATACC AAGGACCGCACCCACAAGGC CACCCACAAGGCCCCCTTCC

Fragments annealed, DNA polymerase extends 3' ends
 ACTTTAGAAGGAGATATACATATGG & TTCCTGGCGTGGGTGTTCCGGAAGG
 CGAAATCTTCCTCTATATGTATACC AAGGACCGCACCCACAAGGCCCCCTTCC

DNA Ligase fills in nicks and gaps

Fig. 3.4

HiFi assembly system working to insert an oligonucleotide into a plasmid, where white letters are the chewed back nucleotides

Once the first round of cloning is finished, the assembled product was sent for sequencing to verify its sequence. Once verified, the recursive Gibson assembly will be performed. Portions of isolated plasmid of both Cys-VE(VEV)₁₂ and Cys-VE(VEV)₁₂L₂₀ were digested with PflMI to linearize the vector near the 5' region of the construct

(Fig. 3.5). Another portion of isolated Cys-VE(VEV)₁₂ was double digested with both NdeI and BglI, which effectively excised the entirety of the protein sequence. This was annealed back into both portions of plasmid that were single digested, following the same homology system as demonstrated with the oligonucleotide (Fig. 3.6).

Cys-E₃₈ & Cys-E₃₈L₄₀

TATACATATGGGCCACGGCGTGGGTGTT
 ATATGTATACCCGGTGCCGCACCCACAA

+ PflMI yields:

TATACATATGGGCCACGGC _____ GTGGGTGTT
 ATATGTATACCCGGTG _____ CCGCACCCACAA

Fig. 3.5

When single digested with PflMI, both vectors are linearized near the 5' region of the protein coding region.

Cys-E₃₈

TATACATATGGGTTGTAACCACGGTGTGG...GGAGTGCCGGGCTGGCCG
ATATGTATACCCAACATTGGTGCCACACC...CCTCACGGCCCGACCGGC

+ NdeI & BglI yields:

TATACA _____ TATGGGTTGTAACCACGGTGTGG...GGAGTGCCGGGC _____ TGGCCG
ATATGTAT _____ ACCCAACATTGGTGCCACACC...CCTCACGGC _____ CCGACCGGC

Insert piece has homology with recipient single digested plasmids, creating:

Cys-E₇₅ & Cys-E₇₅L₄₀

Fig. 3.6

Double digesting Cys-VE(VEV)₁₂ with NdeI and BglI effectively removes 95% of the coding sequence, and can be used as an insert to expand the single digested plasmids. This process may be repeated to obtain a desired chain length.

CHAPTER IV

RESULTS AND DISCUSSION

Student Outcomes

One goal of this project was to establish a curriculum that would give students a solid foundation in basic molecular biology and protein engineering concepts, giving them a foothold to move onto more challenging concepts in higher education. Five students from a high school located on the Cleveland State University Campus (MC²STEM High School) were taught in a trial version of this class. They were taught basic principles of molecular biology and provided with an opportunity to perform the lab work associated with protein engineering techniques. Undergraduate and graduate students were also presented with the same course material and allowed to complete their

own projects, yielding their own unique engineered proteins through the same steps that the high school students pursued.

Students should be able to develop this knowledge through a combination approach of hands-on activities and a series of lecture to supplement these activities, and through this combined approach should come out with a more thorough understanding of biological processes and the challenges involved with making modifications to genomes. A series of protocols were established to make each step “bite-sized” in essence, i.e. capable of being performed in a single classroom session, with available equipment and reagents. While financial limitations may limit equipment purchasing, defined here is the bare minimum required to proctor this course. With a dedicated educator and sister institutions able to share equipment (centrifuges, -80°C freezer), this quick course can be performed at low cost and high speeds such that repeat rounds of cloning may be performed during one semester.

Initially, students are to be introduced to a simple but precise instrument (micropipette) and are explained the intricacies of the instrument, as well as made aware to the small scale with which they will be operating. After understanding the technique of using this instrument and undergoing a protocol involving precise additions of food dyes to create a specific final product, the students were challenged to work in groups to develop videos to educate a lay audience on proper technique for using these micropipettes. Following this brief introduction, students were taught necessary basic DNA and genetic theory and vocabulary, and introduced to one of the most important instruments for molecular biology, the polymerase chain reaction (PCR) unit (T100, Bio-rad). The teacher was tasked with explaining the specific reagents required for using this

instrument properly, in addition to the condition under which the PCR unit operates and what actually occurs when the reagents are added to the unit. After understanding what the device does, the students were asked to prepare a basic demonstration as to how PCR works, using differently-colored paperclips as individual nucleotides to represent the entire reaction.

The start of the molecular biology protocols portion of the curriculum was started with an overnight culture of a previously developed ELP construct that will benefit from modification. The overnight culture was incubated and shaken overnight in antibiotic-treated media to allow for selection of bacteria that possess the plasmid to be modified. This step is immediately followed by a midiprep protocol, which served the purpose of excising plasmids from their bacterial hosts. Students were exposed to the concepts of bacterial growth as well as the concept of what a plasmid is and how plasmids can benefit strains of bacteria.

The next step of this process includes the use of a fluorometer (Qubit 3.0, Thermo Fisher Scientific) to quantify the dsDNA that has been obtained from the PCR step. The concept of fluorescence was described to students, as well as how fluorometry works, and what happens with the reaction set up for the Qubit instrument. Following the instruction of quantification, the next step was the DNA assembly step, in which students used their inserts that they have generated from the previous steps and insert them into a vector. Students were tasked to explain what a plasmid is, what restriction enzymes are and how they act on strands of dsDNA, as well as explain how the Gibson assembly system works, as far as what the enzymes are in the master mix and what role they play in the annealing of an insert into a plasmid. This is a culmination step, in which knowledge of pipetting,

use of the thermocycler, and careful sterile technique yielded the product that they are looking for, but a few additional steps are needed to verify their final product.

The next step required by this protocol was the transformation of this plasmid into competent cells, which are able to take up the plasmid that they have modified. By careful monitoring of the growth conditions of these bacteria, they are able to grow exponentially, and with this growth, also produce nearly as many copies of this plasmid as cells themselves. Students were able to explain what *E. coli* is, the different strains of it, how it is used in the laboratory setting, and the proper care of a strain of bacteria. Students plated their bacterial strains on antibiotic plates containing ampicillin, and learned what additional properties the plasmid confers to the naïve competent cell.

Students at this point were treated to a “cooking show” of sorts, where the plates that the *E. coli* are growing on are put into the incubator overnight, and are pulled the next morning, to show off whether their transformation was successful. Students only need one colony on each plate to have a successful transformation, but were shown plates with normal growth, which ranged from 20-50 colonies from an oligonucleotide ligation to a lawn of bacteria from an insert digestion/ligation reaction.

When more than one colony was present, the students were presented with the opportunity to perform a colony screening procedure. This protocol called for the use of sterile technique to select each colony and perform a PCR reaction on the plasmid DNA present in each colony. Students prepared their own PCR reactions as well as a plate with their isolated screened colonies, and proceed to start the run, with the teacher staying after class to collect the reaction tubes and store them for the next class period. During that next class period, the class set up 1%/2% E-gel agarose gel electrophoresis runs with

their PCR resultant. This protocol reinforces the idea of how PCR is an essential toolkit of the molecular biologist, as well as agarose gel usage as a comparative tool. When the students can show a successful construct has been made, they will be given the opportunity to perform a midiprep on an overnight culture consisting of the colonies that have shown the correct DNA insertion. They will proceed to quantify their plasmid DNA again, and will be instructed on how exactly a sequencing result is obtained by the Sanger sequencing method.

Many of the students found the content challenging but liked being challenged in a scientific setting. Some students found that the teaching material was too complex and the lecturing may need to be restructured to compensate for this, however they did find the biological implications of the material relevant to their college education goals. Further, the students liked the idea of working in a laboratory setting and the ability to perform research at their educational level. The students also received a better grasp of how research is done, and realized that an experiment is not always successful the first time in research settings, unlike what they had experienced in previous basic chemistry and biology laboratories. This data was obtained during interviews with the students from an educational post-baccalaureate project that James Gillahan worked on cooperatively while I taught the students

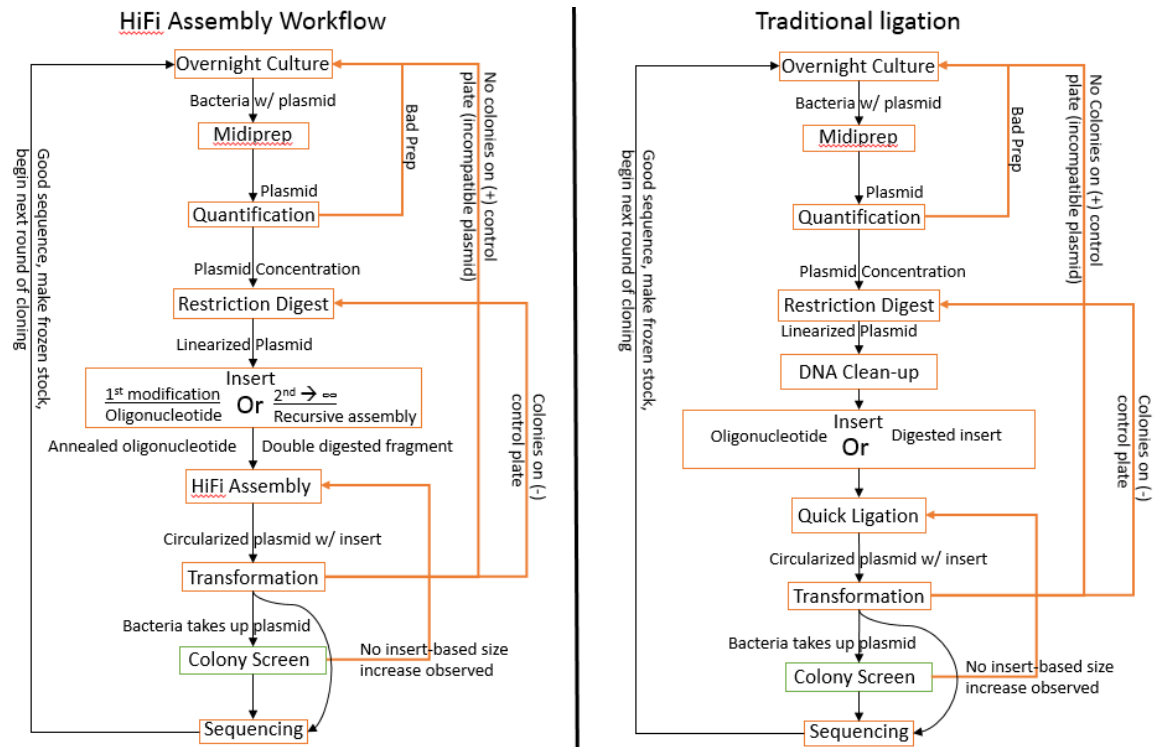


Fig. 4.1

Sample workflow for two different ligation methods, where green boxes are optional steps, black lines denote progress, and red lines are failure indications. Traditional ligation requires an additional DNA clean-up step to prevent re-ligation of linearized plasmid.

| | Safety | Consumable | Time-student | Time-educator | Equipment |
|--------------------|--------|------------|--------------|---------------|------------------------|
| Overnight culture | 2 | 1 | 1 | 4 | A F I |
| Midiprep | 2 | 3 | 2 | 2 | C _t M |
| Quantification | 3 | 1 | 1 | 3 | Q |
| Restriction Digest | 1 | 1 | 1 | 3 | T / D |
| HiFi Assembly | 1 | 1 | 1 | 1 | T / D |
| Transformation | 2 | 2 | 3 | 3 | A D F I |
| Colony Screen | 2 | 4 | 3 | 2 | A C _m D E T |
| Sequencing | 1 | 1 | 1 | 1 | |

Scale for safety, consumable cost, time cost for student and teacher, and necessary equipment rated from (1-4)

Equipment needed for step (**Bold is preferred**, *italicized is second choice*)
 A = autoclave
 C_t = tabletop centrifuge
 C_m = *microcentrifuge*
 D = dry bath
 E = E-gel agarose gel kit
 F = -86°C freezer
 I = incubator shaker
 M = Midiprep kit
 Q = Qubit 3.0 fluorometer
 T = thermocycler

Fig. 4.2

Costs associated with each step of the HiFi assembly protocol.

Combining this robust system with a straightforward protocol that has eliminated most difficult and potentially unsafe steps has made this a viable system for providing recombinant DNA education to those not yet in higher education. Safe protocols for these techniques are a priority for high school education, as there are many hazardous and carcinogenic chemicals that were traditionally applied to molecular biology protocols. Phenol-chloroform extraction is a technique used to remove proteins and other cellular mass from a solution and leave the DNA to elute to the top with the water in a separate phase. This is replaced by the midiprep DNA prep column, and works off of similar principles, namely ethanol washes to remove cellular debris and a final water elution step to detach isolated DNA from the column for use. Acrylamide and ethidium bromide are components of the gels for electrophoresis, for use of separating charged molecules⁵⁸ (like DNA) and for staining nucleic acids ran on a gel when exposed to ultraviolet light. The E-Gel system does not require either of these components, and instead used agarose gels and SYBR-gold for staining the nucleic acids, both of which are less toxic than their counterparts.

Design outcomes

Another goal of this project was to generate an ELP capable of gelling under certain conditions. Five different ELP hydrogel polypeptides were successfully designed as a result of Gibson assembly. This was verified in part by PCR colony screening and conclusively by DNA sequencing performed by Eurofins Genomics. The colony screening was effective in verifying which colonies valid for sequencing, as base pair

length bands depicted on the agarose gel should be 36 base pairs longer for oligonucleotide insertion and 540 base pairs longer for (VEV)₁₂ insertion.

The modular Gibson assembly system is a powerful tool for recombinant DNA synthesis. It is useful for generating large proteins of repetitive sequences which serves as basis of ELP research, as size and guest residue incorporation impart different properties to their T_i values. When adding to the chain length of an ELP, the transition temperature has a tendency to drop as long as the guest residue for the added chains remains the same⁵⁹. Developing ELP hydrogels with different chain lengths offers the potential for study of different gelation temperatures as well as the different mechanical and physical properties of the gel (modulus, pore size, etc.), the objective being finding one that offers clinical relevance. Future projects will include expanding the VEV groups further on many other constructs that possess hydrophobic regions, such as V₂₀ and F₃₂. This assembly method is possible with a full library of ELP constructs, as long as the proper homologous oligonucleotide sequence is introduced to the sequence during the first round of cloning. ELP constructs can now be “mixed and matched” to append new moieties to existing constructs to allow for new microstructure studies that were now previously available.

While not necessary, combining restriction digestion with the Gibson assembly system has improves the likelihood of successful cloning. First, double digestion involves generating two linearized plasmids from one original construct. For this system, a small piece roughly 12 nucleotides in length is removed and the remaining 4 kilobases of the plasmid remain. However, with a double digestion featuring complementary sticky ends on both pieces of plasmid, ligation of these pieces back together is a potential issue

without a cleanup step required with traditional ligation. All pieces of DNA within the reaction have large portions of their 3' region removed, as the Gibson system requires anywhere from 15-40 nucleotides of homology in order to have successful ligation. Using the Gibson system, the smaller piece is completely destroyed during the assembly reaction by the activity of the exonuclease that chews back the 3' region of all DNA fragments within the reaction volume. The fact that there are linearized plasmids massively increases the likelihood that successful cloning will occur, due to the higher chance of a plasmid being chewed back by the nuclease at the correct region. Also, if all of the plasmids are not exposed to the nuclease and were not digested with a restriction enzyme, the likelihood of an unmodified plasmid being transformed into a bacteria is much higher, bringing the need for more PCR screens, which is arguably one of the most time-consuming and expensive protocols. Having a higher probability of exposed homologous regions increases the likelihood of inserts ligating and recircularizing the plasmid, and the likelihood of a chemically competent cell picking up this assembled plasmid. This benefits the cloning process as the ligation is more likely to occur when restriction enzymes are used.

CHAPTER V

CONCLUSION

Several genes encoding potential polypeptide based bioinks have been successfully prepared as demonstrated by DNA sequencing. Further, this is a proof of concept that the Gibson assembly ligation method will work for other protocols, and the doubling reaction performed will work to further expand on the size of various ELP constructs given the right conditions. The modular Gibson assembly system is a powerful tool that we have established for use in any protein engineering application in which precise doubling of protein size is desired and one is limited by large repeated segments. Inserting digested fragments from a plasmid into a single digestion of the same plasmid or any compatible plasmid is granted by this Gibson assembly system, following the insertion of an oligonucleotide with homologous sequence. Further, it does not require any additional purification step following digestions, saving time and increasing the chance of success compared to traditional digestion-ligation reactions.

This gene assembly setup has the potential to double sequences or mix and match indefinitely, so long as the desired homology can be generated by addition of oligonucleotides and the restriction sites necessary for excising the desired larger insert piece can be preserved across subsequent rounds of doubling. This technique has extensive use in studies relevant to ELPs, as changes in size and composition allow for many different transition temperatures and micellar structures to be measured and observed. These ELPs have use not only as potential hydrogels, but also as drug delivery vehicles and as a tool for protein purification.

For high school students, this protocol provides a viable and realistic research experience. It allows them to think critically about laboratory processes, use the instrumentation of researchers, and participate in real experimentation with an outcome similar to authentic research. Students that participate in a course designed around this protocol will be better prepared for undergraduate coursework in biologically inspired fields, and learn what it feels like to be a scientist. As more learning institutions begin to participate in this coursework, the associated costs with establishing a laboratory within a school may also drop, allowing for expansion of the program and for more students to benefit from advanced scientific education.

BIBLIOGRAPHY

1. Bhuniya SP, Rahman S, Satyanand AJ, Gharia MM, Dave AM. Novel route to synthesis of allyl starch and biodegradable hydrogel by copolymerizing allyl-modified starch with methacrylic acid and acrylamide. *Journal of Polymer Science Part A: Polymer Chemistry*. 2003;41(11):1650.
2. Peppas NA, Hilt JZ, Khademhosseini A, Langer R. Hydrogels in biology and medicine: From molecular principles to bionanotechnology. *Advanced Materials*. 2006;18(11):1345.
3. Coviello T, Matricardi P, Marianecchi C, Alhaique F. Polysaccharide hydrogels for modified release formulations. *J Controlled Release*. 2007;119(1):5.
4. Finne A, Albertsson A. Polyester hydrogels with swelling properties controlled by the polymer architecture, molecular weight, and crosslinking agent. *Journal of Polymer Science Part A: Polymer Chemistry*. 2003;41(9):1296.
5. Maitra J, Shukla VK. Cross-linking in hydrogels - A review. *American Journal of Polymer Science*. 2014;4(2):25.
6. van Dijk-Wolthuis WNE, Hoogeboom JAM, van Steenberghe MJ, Tsang SKY, Hennink WE. Degradation and release behaviour of dextran based hydrogels. *Macromolecules*. 1997;30:4639.

7. Di Lullo GA, Sweeney SM, Korkko J, Ala-kokko L, San Antonio JD. Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *Journal of Biological Chemistry*. 2002;277:4223.
8. Federico S, Pierce BF, Piluso S, Wischke C, Lenglein A, Neffe AT. Design of decorin-based peptides that bind to collagen I and their potential as adhesion moieties in biomaterials. *Angewandte Chemistry*. 2015;54:10980.
9. Nistor MT, Vasile C, Chiriac AP, Tartau L. Biocompatibility, biodegradability, and drug carrier ability of hybrid collagen-based hydrogel nanocomposites. *Journal of Bioactive and Compatible Polymers*. 2013;28(6):540.
10. Weber IT, Harrison RW, Iozzo RV. Model structure of decorin and implications for collagen fibrillogenesis. *J Biol Chem*. 1996;271(50):31767-31770.
11. Schonherr E, Hausser H, Beavan L, Kresse H. Decorin-type I collagen interaction: PRESENCE OF SEPARATE CORE PROTEIN-BINDING DOMAINS. *J Biol Chem*. 1995;270(15):8877-8883.
12. Cappello J, Crissman JW, Crissman M, Ferrari FA, Textor G, Wallis O, Whitley JR, Zhou X, Burman D, Aukerman L, Stedronsky ER. In-situ self-assembling protein polymer gel systems for administration, delivery, and release of drugs. *Journal of Controlled Release*. ;53:3/1/2016-105-117.

13. Bromberg LE, Ron ES. Temperature-responsive gels and thermogelling polymer matrices for protein and peptide delivery. *Advanced Drug Delivery Reviews*. 1998;31:197.
14. Khalily MA, Goktas M, Guler MO. Tuning viscoelastic properties of supramolecular peptide gels via dynamic covalent crosslinking. *Org Biomol Chem*. 2015;13(7):1983-1987.
15. Ozbas B, Kretsinger J, Rajagopal K, Schneider JP, Pochan DJ. Salt-triggered peptide folding and consequent self-assembly into hydrogels with tunable modulus. *Macromolecules*. 2004;37(19):7331-7337.
16. Chaterji S, Kwon IK, Park K. Smart polymeric gels: Redefining the limits of biomedical devices. *Prog Polym Sci*. 2007;32(8-9):1083-1122.
17. Betre H, Ong SR, Guilak F, Chilkoti A, Fermor B, Setton LA. Chondrocytic differentiation of human adipose-derived adult stem cells in elastin-like polypeptide. *Biomaterials*. 2006;27(1):91.
18. Dababneh AB, Ozbolat IT. Bioprinting technology: A current state-of-the-art review. *Journal of Manufacturing Science and Engineering*. 2014;136.
19. Lim SH, Mao HQ. Electrospun scaffolds for stem cell engineering. *Adv Drug Deliv Rev*. 2009;61(12):1084-1096.
20. Mandrycky C, Wang Z, Kim K, Kim D. 3D bioprinting for engineering complex tissues. *Biotechnology Advances*. 2016;34(4):422.

21. Xu T, Zhao W, Zhu JM, Albanna MZ, Yoo JJ, Atala A. Complex heterogeneous tissue constructs containing multiple cell types prepared by inkjet printing technology. *Biomaterials*. 2013;34(1):130-139.
22. Xu T, Gregory C, Molnar P, Cui X, Jalota S, Bhaduri S, Boland T. Viability and electrophysiology of neural cell structures generated by the inkjet printing method. *Biomaterials*. 2006.
23. Koch L, Gruene M, Unger C, Chichkov B. Laser assisted cell printing. *Curr Pharm Biotechnol*. 2013;14(1):91-97.
24. Barron JA, Ringeisen BR, Kim H, Spargo BJ, Chrisey DB. Application of laser printing to mammalian cells. *Thin Solid Films*. 2004;453-454:383-387.
25. Ozbolat IT, Yu Y. Bioprinting toward organ fabrication: Challenges and future trends. *IEEE Trans Biomed Eng*. 2013;60(3):691-699.
26. Wang Z, Abdulla R, Parker B, Samanipour R, Ghosh S, Kim K. A simple and high-resolution stereolithography-based 3D bioprinting system using visible light crosslinkable bioinks. *Biofabrication*. 2015;7(4):045009-5090/7/4/045009.
27. Zhu J, Marchant RE. Design properties of hydrogel tissue-engineering scaffolds. *Expert Rev Med Devices*. 2011;8(5):607-626.
28. Matsiko A, Gleeson JP, O'Brien FJ. Scaffold mean pore size influences MEsenchymal stem cell chondrogenic differentiation and matrix deposition. *Tissue Engineering Part A*. 2015;21(3):486.

29. Lee-Thedieck C, Rauch N, Fiammengo R, Klein G, Spatz JP. Impact of substrate elasticity on human hematopoietic stem and progenitor cell adhesion and motility. *Journal of Cell Science*. 2012;125:3765.
30. Urry DW, Shaw RG, Prasad KU. Polypentapeptide of elastin: Temperature dependence of ellipticity and correlation with elastomeric force. *Biochemical and Biophysical Research Communications*. 1985;30(1):50.
31. Meyer DE, Trabbic-Carlson K, Chilkoti A. Protein purification by fusion with an environmentally responsive elastin-like polypeptide: Effect of polypeptide length on the purification of thioredoxin. *Biotechnology Progress*. 2001;17:720.
32. Cho Y, Zhang Y, Christensen T, Sagle LB, Chilkoti A, Cremer PS. Effects of Hofmeister anions on the phase transition temperature of elastin-like polypeptides. *Journal of Physical Chemistry B*. ;112(44):3/2/16-13765-13771.
33. Urry DW, Gowda DC, Parker TM, Luan C, Reid MC, Harris CM, Pattanaik A, Harris RD. Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers*. 1992;32(9):1243.
34. Tang M. SYNTHESIS AND CHARACTERIZATION OF pH-RESPONSIVE ELASTIN-LIKE POLYPEPTIDES WITH DIFFERENT CONFIGURATIONS. . 2013. http://rave.ohiolink.edu/etdc/view?acc_num=csu1376784321.

35. Urry DW, Peng SQ, Hayes L, Jaggard J, Harris RD. A new mechanism of MEchanochemical coupling: Stretch-induced increase in carboxyl pK_a as a diagnostic. *Biopolymers*. 1990;30(1-2).
36. Urry DW. Entropic elastic processes in protein mechanisms. I. elastic structure due to an inverse temperature transition and elasticity due to internal chain dynamics. *J Protein Chem*. 1988;7(1):1-34.
37. Prescott LM, Harley JP, Klein DA. *Microbiology*. Dubuque, IA: Wm. C. Brown Publishers; 1996.
38. Welch RA, Burland V, Plunkett G, 3rd, Redford P, Roesch P, Rasko D, Buckles EL, Liou SR, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HL, Donnenberg MS, Blattner FR. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic escherichia coli. *Proc Natl Acad Sci U S A*. 2002;99(26):17020-17024.
39. Isenberg I. Histones. *Annual Review Biochem*. 1979;48:159.
40. Zhao J, Jin SB, Bjorkroth B, Wieslander L, Daneholt B. The mRNA export factor Dbp5 is associated with balbiani ring mRNP from gene to cytoplasm. *EMBO J*. 2002;21(5):1177-1187.
41. Olivera BM, Lehman IR. Linkage of polynucleotides through phosphodiester bonds by an enzyme from escherichia coli. *Proc Natl Acad Sci U S A*. 1967;57(5):1426-1433.

42. Filipe JM, Monteiro GA, Cabral JMS, Taipa MA. Design of bacterial vector systems for the production of recombinant proteins in *escherichia coli*. *Journal of Microbiology and Biotechnology*. 2003;14(1):1.
43. Marbach A, Berrenbrock K. Lac operon induction in *escherichia coli*: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *Journal of Biotechnology*. ;157(1):82-88.
44. Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *Journal of Biological engineering*. 2008;2(5).
45. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: A one-pot DNA shuffling method based on type II restriction enzymes. *PLoS ONE*. 2009;4(5).
46. Gibson DG, Young L, Chuang R, Venter JC, Hutchinson CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature*. 2009;6(5):343.
47. Bauerle C, DePass A, Lynn D, O'Connor C, Singer S, Withers M, Anderson CW, Donovan S, Drew S, Ebert-May D, Gross L, Hoskins SG, Labov J, Lopatto D, McClatchey W, Varma-Nelson P, Pelaez N, Poston M, Tanner K, Wessner D, White H, Wood W, Wubah D. Vision and change in undergraduate biology education: A call to action. *American Association for the Advancement of Science*. 2009.
48. Ghose R, Sewer M, Booker S, Sumter T. Workshop: ASBMB mentoring program for early career scientists.
https://www.nsf.gov/awardsearch/showAward?AWD_ID=1426133. Updated 2014.

49. Stevens-Truss R. Grants aim to enhance STEM education in middle and high schools. http://www.asbmb.org/asbmbtoday/asbmbtoday_article.aspx?id=14596. Updated 2011.
50. Bybee RW. Learning science and the science of learning. . 2002.
51. Handelsman J, Ebert-May D, Beichner R, Bruns P, Chang A, DeHaan R, Gentile J, Lauffer S, Stewart J, Tilghman SM, Wood WB. Scientific teaching. *Science*. 2004;304(5670):521-522.
52. Ghoorchian A, Holland NB. Molecular architecture influences the thermally induced aggregation behavior of elastin-like polypeptides. *Biomacromolecules*. 2011;12(11):4022-4029.
53. Ghoorchian A, Cole JT, Holland NB. Thermoreversible micelle formation using a three-armed star elastin-like polypeptide. *Macromolecules*. 2010;43(9):4340-4345.
54. Can O. Novel antifreeze protein constructs for improved activity. . 2008. http://rave.ohiolink.edu/etdc/view?acc_num=csu1229703788.
55. Frank S, Kammerer RA, Mechling D, Schulthess T, Landwehr R, Bann J, Guo Y, Lustig A, Bachinger HP, Engel J. Stabilization of short collagen-like triple helices by protein engineering. *J Mol Biol*. 2001;308(5):1081-1089.
56. Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG. Structure of bacteriophage T4 fibritin: A segmented coiled coil and the role of the C-terminal domain. *Structure*. 1997;5(6):789-798.

57. Anumolu SS, Menjoge AR, Deshmukh M, Gerecke D, Stein S, Laskin J, Sinko PJ. Doxycycline hydrogels with reversible disulfide crosslinks for dermal wound healing of mustard injuries. *Biomaterials*. 2011;32(4):1204-1217.
58. Raymond S, Weintraub L. Acrylamide gel as a supporting medium for zone electrophoresis. *Science*. 1959;130(3377):711.
59. Meyer DE, Chilkoti A. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules*. 2004;5(3):846-851.

APPENDIX

APPENDIX A

DNA Sequences

Cys-VE oligonucleotide forward:

M G C N H G V G V P G E G V
AGAAGGAGATATACATATGGGTTGTAACCACGGTGTGGGAGTGCCGGGCGAAGGTGTT
P G V G V P
CCTGGCGTGGGTGTTCCG

Cys-VE oligonucleotide complement

CGGAACACCCACGCCAGGAACACCTTCGCCCGGCACTCCCACACCGTGGTTACAACCCA
TATGTATATCTCCTTCT

(VEV)₁₂

M G H
5' -ATGGGCCAC
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P

G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G W P
GGCTGGCCG-3'

Cys-V (E (VEV)₁₂)_{2L20}

Anterior region:

M G C N H G V G V P G E G V P
5'-ATGGGTGTGTAACCACGGTGTGGGAGTGCCGGGCGAAGGTGTTCCCT
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG

G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG

G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG

G E G V P
GGCGAAGGTGTTTCCT

G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P

GGCGTGGGTGTTCCGGGGGAAGGTGTACCNGGTGTGGGAGTGCCG-3'

Posterior region:

G V G V P G E G V P G V G V P
5'-GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G V G V P G E G V P G V G V P
GGCGTGGGTGTTCCGGGGGAAGGTGTACCAGGTGTGGGAGTGCCG
G L G V P G L G V P G L G V P
GGCTTGGGTGTTCCGGGCCTGGGTGTCCCAGGTCTGGGCGTACCG
G L G V P G L G V P G L G V P
GGCCTCGGTGTTCCCTGGTCTGGGCGTGCCGGGCTTGGGTGTTCCG
G L G V P G L G V P G L G V P
GGCCTGGGTGTCCCAGGTCTGGGCGTACCGGCCTCGGTGTTCCCT
G L G V P G L G V P G L G V P
GGTCTGGGCGTGCCGGGCTTGGGTGTTCCGGGCCTGGGTGTCCCA
G L G V P G L G V P G L G V P
GGTCTGGGCGTACCGGCCTCGGTGTTCCCTGGTCTGGGCGTGCCG

G L G V P G L G V P G L G V P
GGCTTGGGTGTTCCGGGCCTGGGTGTCCCAGGTCTGGGCGTACCG
G L G V P G L G V P
GGCCTCGGTGTTCCCTGGTCTGGGCGTGCCG
G W P
GGCTGGCCG-3'

When extrapolated and combined, gives us:

MGCNH-V (E (VEV)₁₂)₂L₂₀

Decorin-V₁₉-foldon

M G H G V V P H L R E L H L N N N K
5'-ATGGGCCACGGCGTGGTACCTCATTTACGTGAATTACACTTGAATAATAATAAA
L G V G V P G V G V P G V G V P
CTAGGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGTGCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGTGCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G W P G Y I P E A P
GGCGTGGGTGTGCCGGGCTGGCCGGGTTACATCCCGGAAGCTCCG
R D G Q A Y V R K D G E W V L
CGTGACGGTCAGGCTTACGTTTCGTAAAGACGGTGAATGGGTTCTG

L S T F L
CTGTCTACCTTCCTG

Decorin-V₃₇-foldon

M G H G V V P H L R E L H L N N N K
5'-ATGGGCCACGGCGTGGTACCTCATTACGTGAATTACACTTGAATAATAATAAA
L G V G V P G V G V P G V G V P
CTAGGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P

GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G W P G Y I P E A P
GGCGTGGGTGTGCCGGGCTGGCCGGGTACATCCCGGAAGCTCCG
R D G Q A Y V R K D G E W V L
CGTGACGGTCAGGCTTACGTTTCGTAAAGACGGTGAATGGGTTCTG
L S T F L
CTGTCTACCTTCCTG-3'

Decorin-73-foldon:

M G H G V V P H L R E L H L N N N K
5'-ATGGGCCACGGCGTGGTACCTCATTTACGTGAATTACACTTGAATAATAATAAA
L G V G V P G V G V P G V G V P
CTAGGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGTGCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGTGCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGTGCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P

GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P
GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT
G V G V P G V G V P G V G V P
GGCGTGGGTGTGCCGGGCGTTGGTGTTCCTGGCGTAGGTGTCCCA
G V G V P G V G V P G V G V P
GGCGTCGGTGTTCCTGGTGTCCGGGCGTGGGTGTGCCG
G V G V P G V G V P G V G V P

GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTTCCT
G V G V P G W P G Y I P E A P
GGCGTGGGTGTGCCGGGCTGGCCGGGTTACATCCCGGAAGCTCCG
R D G Q A Y V R K D G E W V L
CGTGACGGTCAGGCTTACGTTTCGTAAAGACGGTGAATGGGTTCTG
L S T F L
CTGTCTACCTTCCTG-3'