



Spring 4-2015

## Synthesis and Screen of a Proline-Rich Combinatorial Library Towards the Identification of Sickle Cell Hemoglobin Polymerization Inhibitors

Laura Steenberge  
*Illinois Wesleyan University*

Follow this and additional works at: [https://digitalcommons.iwu.edu/chem\\_honproj](https://digitalcommons.iwu.edu/chem_honproj)



Part of the [Chemistry Commons](#)

---

### Recommended Citation

Steenberge, Laura, "Synthesis and Screen of a Proline-Rich Combinatorial Library Towards the Identification of Sickle Cell Hemoglobin Polymerization Inhibitors" (2015). *Honors Projects*. 36.

[https://digitalcommons.iwu.edu/chem\\_honproj/36](https://digitalcommons.iwu.edu/chem_honproj/36)

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Commons @ IWU with permission from the rights-holder(s). You are free to use this material in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This material has been accepted for inclusion by faculty at Illinois Wesleyan University. For more information, please contact [digitalcommons@iwu.edu](mailto:digitalcommons@iwu.edu).

©Copyright is owned by the author of this document.

**SYNTHESIS AND SCREEN OF A PROLINE-RICH  
COMBINATORIAL LIBRARY TOWARDS THE  
IDENTIFICATION OF SICKLE CELL  
HEMOGLOBIN POLYMERIZATION  
INHIBITORS**

**A Research Honors Project by  
LAURA STEENBERGE**

**Illinois Wesleyan University  
Department of Chemistry**

**Defended on 20 April 2015**

## *Acknowledgements*

First and foremost, I would like to thank my research mentor, Dr. Brian Brennan, for his constant support and guidance over the past three years. I started off doing research as just another check off of my resume on my unrelenting path to medical school, and he opened my eyes to the world of science and helped me develop a love for research. Thank you for pointing me towards an MD/PhD track and preparing me for an exciting time in graduate school and beyond. All of the Mugsy's times, my constant appearances at your office, and the time spent talking about how cool proteins are will not be forgotten.

I would also like to thank the members of my thesis committee, Dr. Ram Mohan, Dr. Melinda Baur, and Dr. David Bollivar, for their support during this project over the past year and their work to help develop this work into a final project. I would like to extend a general thanks to the entire IWU Chemistry department for providing a space for me to cultivate my knowledge and passion for science and for putting up with my shenanigans over the past four years.

I of course need to thank all of the members of my lab group who have supported me through the countless hours of lab time. Special shout-outs go to Ben Becker for discovering how to make peptides with me, Cody Wynn for being the go-to person for questions, the two Zacks for spicing up lab time, and Ariella Tavor for letting me express all of my emotions during this past year. I would like to especially thank Sam Baginski for experiencing with me all of my failures and successes, always listening to my lab stories, and being by my side for the past four years.

Finally, I would be remiss if I did not thank all of my family and friends for their love and support. Thank you for pretending to be interested in hemoglobin even though you know nothing about science and for always supporting through this entire journey.

## *Materials Used*

Fmoc protected peptides were purchased from Synbiosci. CLEAR-Amide Resin and TentaGel R NH<sub>2</sub> were purchased from Peptides International. Goat anti-hemoglobin  $\alpha$  stock solution of 200  $\mu$ g/mL and rabbit anti-goat IgG with AP conjugated stock solution of 400  $\mu$ g/mL were purchased from Santa Cruz Biotech. A stock solution of 18.75 mg/mL nitro blue tetrazolium chloride (NBT) and 9.4 mg/mL 5-bromo-4-chloro-3-indolyl-phosphate toluidine-salt in 67% DMSO was purchased from Roche Diagnostics. N,N-dimethylformamide was purchased from Aapptec. Morpholine at over 99% purity, triethylamine at over 99.5% purity, human ferrous hemoglobin S, sodium chloride at over 99% purity, tris(hydroxymethyl)aminomethane at over 99.8% purity, glycine at over 99% purity, sodium dodecyl sulfate at 98% purity, Tween-80 from Sigman-Aldrich, and trifluoroacetic acid at over 99% purity was purchased from Sigma-Aldrich. Triisopropylsilane at 98% purity from Acros Organics was used. A Ph.D.-7 Phage Display Peptide Library Kit containing *E. coli* K12 ER2738 and Ph.D.-7 Phage Display Peptide Library was used to conduct the phage display experiment.

## *List of Abbreviations in the Report*

AP: Alkaline phosphate buffer (100 mM NaCl, 100 mM Tris-Cl at pH 9.5, 50 mM MgCl<sub>2</sub>)  
BCIP: 5-bromo-4-chloro-3-indolyl-phosphate  
ELISA: Enzyme linked immunosorbent assay  
Fmoc: Fluorenylmethyloxycarbonyl chloride  
Hb S: Sick cell hemoglobin  
HBTU: *N,N,N',N'*-Tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate  
HOBt: Hydroxybenzotriazole  
NBT: Nitro blue tetrazolium chloride  
PBS: Phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, and 2 mM KH<sub>2</sub>PO<sub>4</sub>)  
PBST: PBS with 0.1% Tween-80  
SPPS: Solid phase peptide synthesis  
TBS: Tris buffered saline (10 mM Tris at pH 8, 150 mM NaCl)  
TBST: TBS with 0.1% Tween-80  
TEA: Triethylamine

## TABLE OF CONTENTS

### ***Chapter I: Sickle Cell Disease***

Sickle Cell Disease Overview	1
Structure of Hemoglobin	1-3
The Binding of Hemoglobin to Oxygen	3-4
Structure of Sickle Cell Hemoglobin and its Polymers	4-6
Phenotypes of Sickle Cell Disease	6-7
Current Therapeutics for the Treatment of Sickle Cell Disease	7-9
Our Approach	9-10

### ***Chapter II: Phage Display***

Phage Display Overview	11-13
Phage Display Libraries	13-16
Experimental	16-18

### ***Chapter III: Peptide Combinatorial Library Synthesis***

Overview of Peptides as Therapeutics	19-20
The Structure of Peptides	21-22
The Synthesis of Peptides	22-28
Combinatorial Peptide Libraries	28-30
Goal of Combinatorial Proline-Rich Peptide Library	30-31
Experimental	31-32

### ***Chapter IV: ELISA Screen and Polymerization Assays***

ELISA Overview	33-36
ELISA Experimental	36-39
Polymerization Assay Overview	39-40
Polymerization Assays Experimental	41-43
Conclusions and Future Work	43-44

<b><i>References</i></b>	45-46
--------------------------	-------

## **Abstract**

Sickle cell disease is a genetic disorder that affects the hemoglobin within red blood cells. A point mutation in the gene coding for the  $\beta$ -subunit of hemoglobin allows the mutant chain to interact with a hydrophobic pocket of another hemoglobin in a deoxygenated environment, causing polymerization of the proteins. This creates the characteristic sickle-shape of the diseased blood cells that can clog capillaries, leading to tissue damage and cell death. Currently, there are limited options for those affected with sickle cell disease. The research to be presented is focused on discovering peptides that can interact with the mutated hemoglobin and prevent aggregation. A novel proline-rich peptide ligand, ZSF39, was identified through a phage display against deoxygenated sickle cell hemoglobin. A combinatorial peptide library based on the structure of ZSF39 was synthesized and screened for binding affinity using an ELISA. A tightly binding peptide, LHS1, was discovered through the ELISA and found to have a significant inhibitory effect on the polymerization of sickle cell hemoglobin. This work represents a novel approach for the discovery of therapeutics for this debilitating disorder.

## **Chapter I: Sickle Cell Disease**

### *A. Sickle Cell Disease Overview*

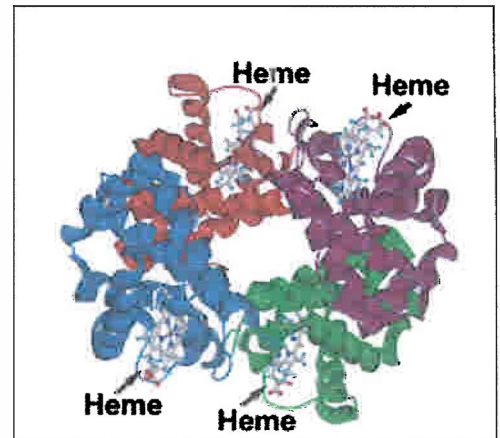
Sickle cell disease, one of the most common monogenic disorders, is an inherited blood disease. It encompasses a variety of disorders that all result in the polymerization of hemoglobin, the majority of which are due to the sickle cell hemoglobin mutation Hb SS. The disorder was first described by Herrick in 1910, and the connection with a hemoglobin abnormality was first suggested by Linus Pauling in 1945.<sup>1</sup> Today, there are more than 280,00 children born worldwide with sickle cell disease annually.<sup>2</sup> Although the disease affects many people every year, it is characterized as a rare disease in the United States because the majority of the people affected are in sub-Saharan areas of Africa. This disease causes significant mortality, with a median survival age of 42 years for men and 48 years for women in the United States,<sup>3</sup> and also causes a variety of symptoms that impact the lives of patients. Although this disease affects many people, it is not widely studied, and hence is an important area of study for the identification of new treatments.

### *B. Structure of Hemoglobin*

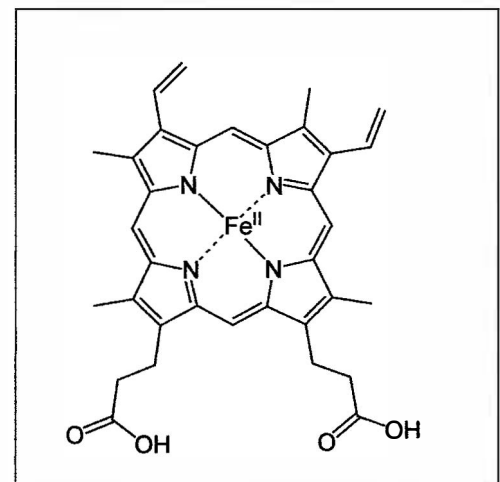
The majority of sickle cell disease cases are caused by a single point mutation in the  $\beta$ -globin subunit of hemoglobin, resulting in a structure that can polymerize under deoxygenated conditions and form fibers. The general structure of normal hemoglobin, the oxygen-carrying protein, is a tetramer that is composed of two pairs of subunits that can carry up to four molecules of oxygen. The subunits can be formed from seven different polypeptide chains ( $\alpha$ ,  $\beta$ ,  $A\gamma$ ,  $G\gamma$ ,  $\delta$ ,  $\epsilon$ , and  $\zeta$ ), each of which arises in some form of hemoglobin during different points in development.<sup>4</sup> Adult hemoglobin (Hb A) is comprised of  $\alpha$  and  $\beta$  subunits, with an overall structure of  $\alpha_2\beta_2$ . Another common type of hemoglobin, fetal hemoglobin (Hb F), is present

primarily in fetuses and has the structure  $\alpha_2\gamma_2$ . Fetal hemoglobin has a stronger affinity for oxygen than adult hemoglobin so it can capture oxygen molecules from the maternal hemoglobin.

Two pairs of two subunits come together to make the overall tetramer, forming the structure seen in Figure I.1. Each subunit contains a heme group, whose structure is shown in Figure I.2. It has a ferrous atom in the center that coordinates with five ligands when not bound to oxygen, or six ligands when bound to oxygen. Four of the coordinate covalent bonds of the ferrous atom are to the pyrrole nitrogen atoms of the porphyrin, which lies in a plane. The fifth bond is to the imidazole of a histidine in one of the  $\alpha$  helices, His F8, which is perpendicular to the plane of the porphyrin ring. This residue is known as the proximal histidine. If oxygen is bound to heme, it occupies the sixth coordination site and also coordinates with the distal histidine (His E7). These interactions are shown in Figure I.3. The coordination of heme to the various ligands is important for the cooperative binding of oxygen to hemoglobin.

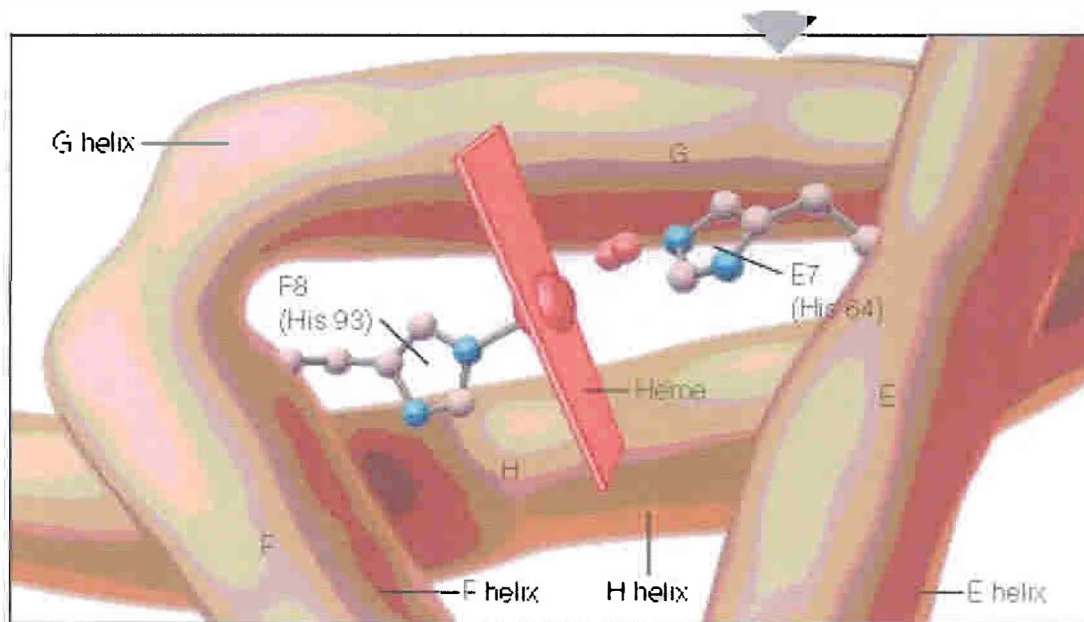


**Figure I.1** A ribbon diagram structure of adult hemoglobin bound to four oxygen molecules with the heme groups indicated. The two  $\alpha$  subunits are colored blue and green and the two  $\beta$  subunits are colored red and purple.<sup>5</sup>



**Figure I.2** The structure of the heme group.





**Figure I.3.** A cartoon of the heme molecule coordinating with His F8 and His E7 through oxygen, with the oxygen molecule shown in red.<sup>6</sup>

### *C. The Binding of Hemoglobin to Oxygen*

Hemoglobin exhibits positive cooperative binding to oxygen, which means that the binding of each oxygen molecule to one subunit makes it easier for the next oxygen molecule to bind to another subunit, and the dissociation of each oxygen molecule makes it easier for the next oxygen molecule to dissociate. Myoglobin, which essentially can be viewed as one subunit of hemoglobin, requires a 81 fold change of concentration of oxygen to increase the percentage associated with oxygen from 10% to 90%; however, hemoglobin only needs a 4.8 fold change to attain this percentage change.<sup>5</sup> This cooperativity is due to the effect of the geometry of the heme group on the conformation of the protein. When heme is not bound to oxygen, it is slightly bent, with the ferrous atom out of the plane of the rings. However, when it binds to oxygen, the heme adopts a flat conformation with the ferrous atom in the plane. This is relatively difficult due to the large size of the  $\text{Fe}^{2+}$  atom and can only happen when the free energy of the bond formation

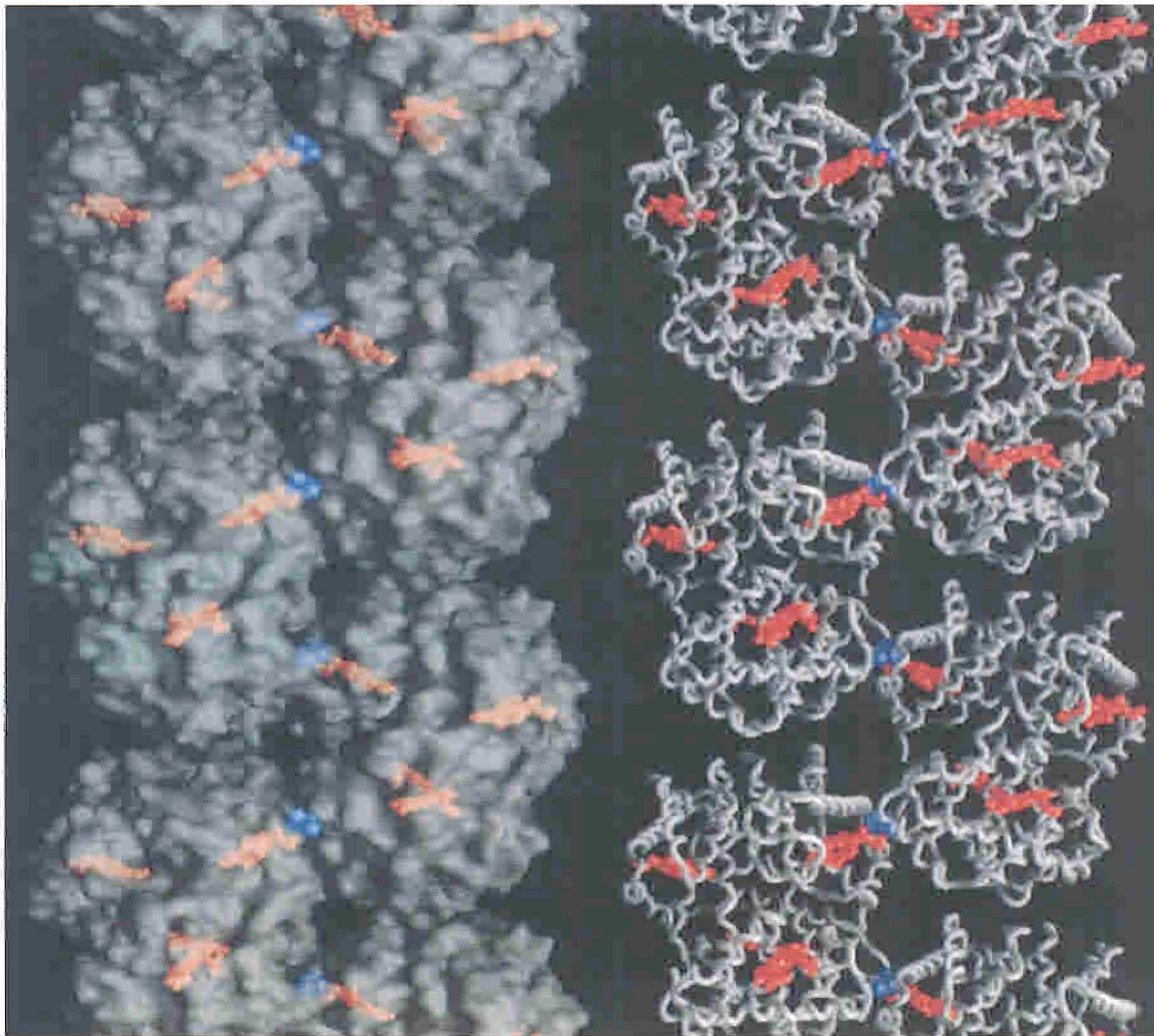
overcomes the ring strain of the flat heme conformation.<sup>5</sup> This change causes the His F8 residue to be pulled towards the porphyrin ring, which moves the F helix and causes subsequent destabilization of the interaction of the subunit with the adjacent subunit. This results in a conformational change of the nearby subunit that allows it to bind to oxygen easier. The oxygenated conformation of hemoglobin is referred to as the “R” or “relaxed” state, while the deoxygenated conformation is the “T” or “tense” state.<sup>5</sup>

Many factors can influence the equilibrium between the T and R states of hemoglobin. For example, hemoglobin releases protons when shifting from the T state to the R state, so a decrease in pH causes a shift in the equilibrium back to the T state. This ability to respond to pH allows hemoglobin to adequately deliver oxygen in oxygen-deprived muscles, as the pH changes due to biological cues such the dissociation of carbonic acid or the formation of lactic acid. Another factor that affects the conformations of hemoglobin is the molecule 2,3-bisphosphoglycerate (BPG), which stabilizes the T state of hemoglobin and is used by the body to regulate oxygenation levels.

#### *D. Structure of Sickle Cell Hemoglobin and Its Polymers*

Sickle cell hemoglobin (Hb S) results from a point mutation in the gene coding for the  $\beta$  subunit that allows the protein to polymerize in the deoxygenated state. The gene for the  $\beta$  subunit in humans is on chromosome 11, and a change from an adenine base to a thymine base creates a codon that codes for valine (V) (GTG) instead of glutamate (E) (GAG).<sup>7</sup> This mutation takes place at the sixth position of the  $\beta$  chain, and therefore is known as an E6V mutation. This mutation replaces a negatively charged residue with a nonpolar residue that can interact with residues of another hemoglobin protein when hemoglobin is deoxygenated, resulting in polymerization. The crystal structure of polymerized deoxygenated hemoglobin reveals that the

valine interacts with a hydrophobic pocket containing alanine, phenylalanine, and leucine between the E and F helices of another hemoglobin.<sup>8</sup> This lateral contact produces a double strand of hemoglobin tetramers, as shown in Figure I.4. This polymer, which is also stabilized through axial contacts between two adjacent proteins in one strand, is 210 Å thick and has a slight helical twist. The end result of this polymerization is a polymer with fourteen filament strands that has decreased solubility in aqueous environments.

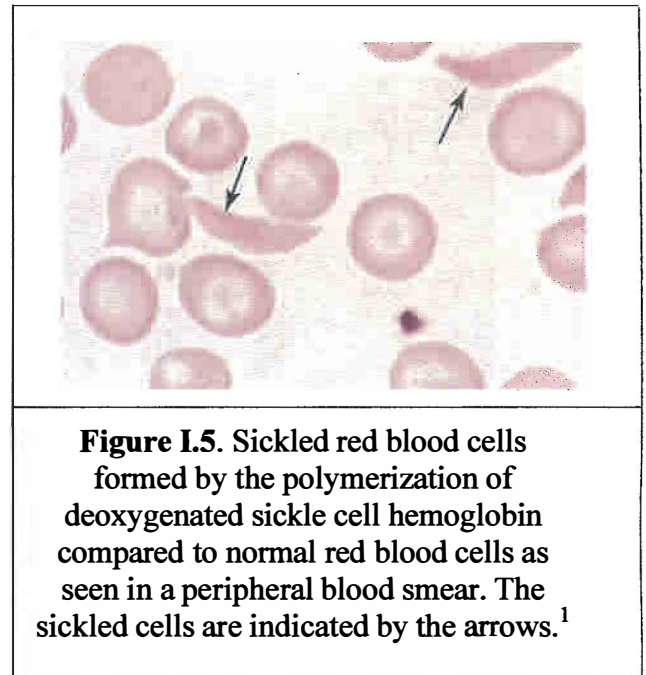


**Figure I.4** The structure of polymers of deoxygenated sickle cell hemoglobin. On the diagram on the left, the heme groups are shown in red and the valine mutations involved in the lateral contacts connecting the two strands are shown in blue, and the rest of the protein is shown as a transparent molecular surface. On the diagram on the right, the same coloring system is used, and the backbone of the protein is shown in white.<sup>9</sup>

The lateral contacts holding the two strands together are not just held together through hydrophobic interactions, but are also stabilized by hydrophilic interactions with nearby residues. These other interactions are also important for the polymerization of Hb S and subsequent consequences of sickle cell disease.

### *E. Phenotypes of Sickle Cell Disease*

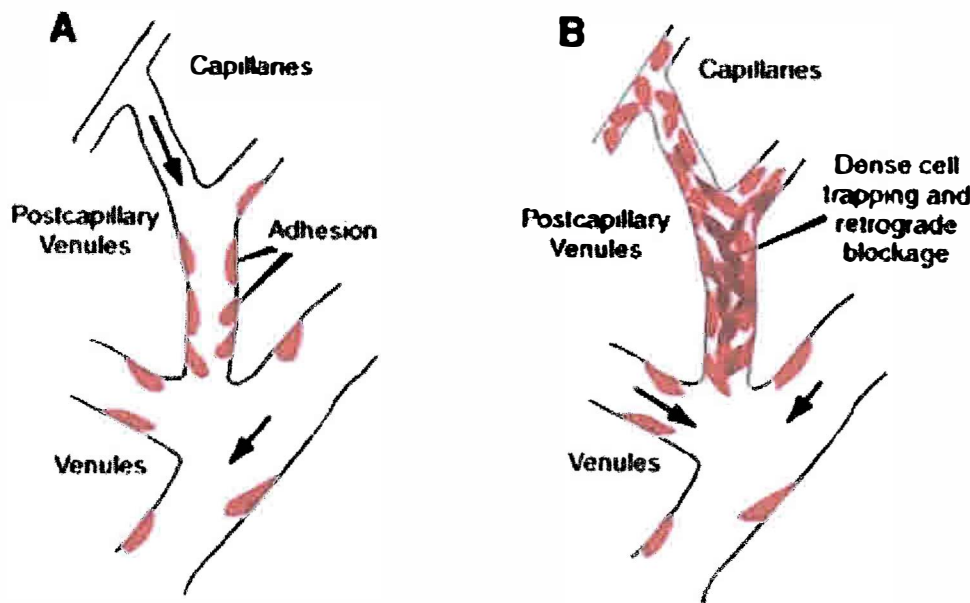
The polymerization of hemoglobin has a variety of effects in the body. As the hemoglobin in the red blood cells polymerizes, the erythrocyte deforms into the characteristic sickle shape that is associated with the disease. Figure I.5 shows a comparison of sickled red blood and normal red blood cells. This causes a number of complications. One of the main problems that can arise is that the abnormal shape and increased expression of surface



adhesion proteins on the red blood cells can block small blood vessels and cause vaso-occlusion, as shown in Figure 1.6. This results in tissue ischemia, which causes pain and organ damage.<sup>3</sup>

The pain is a major hindrance to the normal life of patients, and the organ damage can build up to severe levels. Another issue with hemoglobin polymerization is that it can cause hemolysis, or the rupture of red blood cells, because of the deformation of the cells when polymerization occurs. The normal life of a red blood cell is about 100 to 120 days, while sickle red blood cells live for only a little more than a week.<sup>3</sup> Hemolysis causes anemia, pulmonary hypertension, and jaundice. Another common issue in sickle cell patients that usually arises in childhood is splenic sequestration. The spleen has narrow blood vessels and is overworked due to the large volume of

damaged red blood cells, so it frequently undergoes infarction and is nonfunctional in sickle cell patients.<sup>3</sup> This leads to patients that are more susceptible to infection. Another problem in sickle cell disease patients is acute chest syndrome, which arises from a variety of complex issues such as infections, pulmonary emboli, and pulmonary infarctions and is a major cause of mortality. Additionally, many other patients die from compound organ damage caused by many vaso-occlusion events. Due to these symptoms, sickle cell disease is the cause of many hospital visits, morbidity, and mortality of the world population.



**Figure I.6** The process of vaso-occlusion. The red blood cells adhere to the wall of the blood vessel and build up until the flow of blood through the vessel is stopped, causing pain and organ damage.<sup>10</sup>

#### *F. Current Therapeutics for the Treatment of Sickle Cell Disease*

One of the few predictors of how severe the sickle cell disease will present itself is the amount of fetal hemoglobin in the blood. The greater the amount of Hb F, the better the outcome for the patient. Fetal hemoglobin production usually drops nine to eighteen months after birth, and stabilizes after three to four years. However, a small amount of Hb F still remains in the

blood, around 0.5-2% of total hemoglobin in regular adults and 3-20% in sickle cell disease patients.<sup>9</sup> Because the structure of Hb F is  $\alpha_2\gamma_2$ , it is not affected by the sickle cell mutation in the  $\beta$  subunit. Higher concentrations of Hb F work to decrease the concentration of Hb S and results in less polymerization and therefore fewer complications of the disease. The only current FDA-approved treatment for sickle cell disease is hydroxyurea, which raises the levels of Hb F in the blood. Hydroxyurea is usually used as an anti-cancer agent, and is a ribonucleotide reductase inhibitor. Through its cytotoxic activity, it can produce stress on the production of red blood cells, which triggers the body to increase the expression of fetal hemoglobin over adult hemoglobin.<sup>11</sup> Unfortunately, the response to the treatment greatly varies between patients, and it frequently has undesirable side effects, making it unusable in some patients. The only other treatments available for SCD patients are those that manage the symptoms, such as pain relievers, hydrating agents, and organ transplantation to replace damaged organs. It can be cured with a bone marrow transplant, but this is a risky and expensive procedure, and it is difficult to find suitable donors that do not produce an immunogenic response.

There are a variety of treatments that are currently being investigated as better alternatives to hydroxyurea for sickle cell disease. One area of research is finding molecules that can prevent polymerization. For example, there is a class of compounds that form Schiff-base adducts with Hb S and increase its affinity for oxygen, decreasing the levels of hemoglobin in the T state and reducing the amount of polymerization.<sup>12</sup> Another area of research is focused on decreasing the adhesion of the red blood cells to the endothelium of the blood vessels, which would reduce the occurrence of vaso-occlusion events.<sup>13</sup> A third area is working to restore nitric oxide bioavailability. Nitric oxide, or NO, is a vasodilator that is made bioavailable by binding to hemoglobin.<sup>5</sup> This function is disrupted in polymerized hemoglobin and free heme scavenges



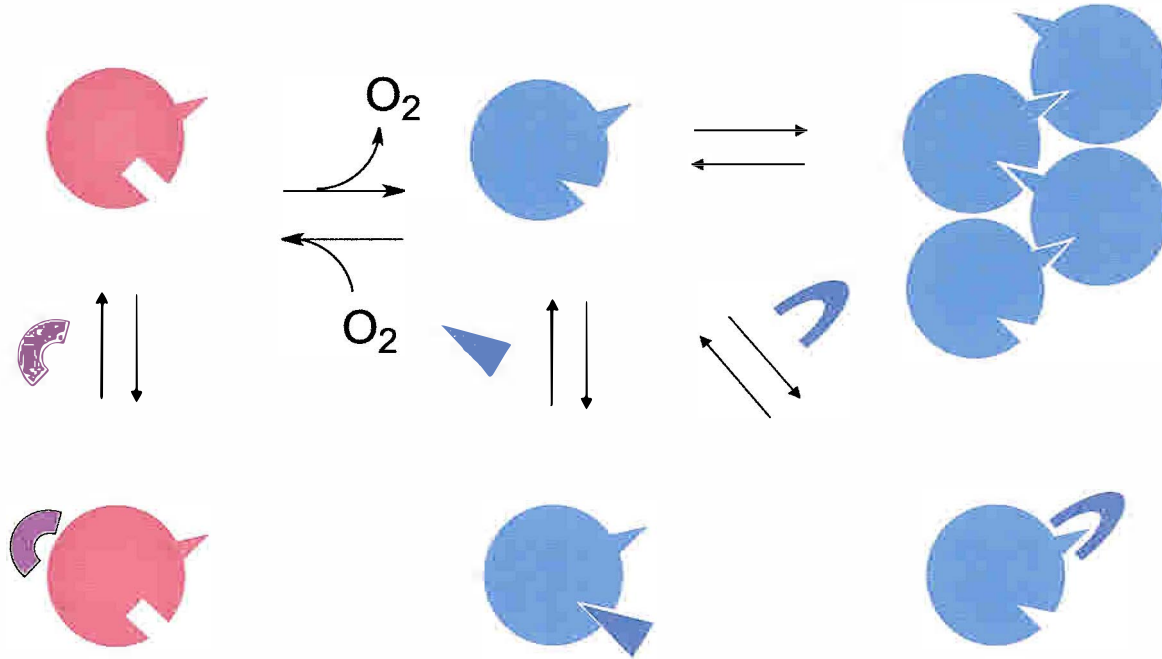
NO; therefore, there is less NO available to the blood vessels. This causes the vessels to be more constricted than usual and results in an aggravation of vaso-occlusion.<sup>2</sup> There is also quite a bit of work being done on gene therapy. For example, Purumbet et al. were able to insert a lentiviral vector with gamma globin exons and beta globin regulatory units and increase the concentration of fetal hemoglobin in a murine model to a level where the amount of polymerization decreased significantly.<sup>14</sup> However, gene therapy comes with a variety of problems, including gene delivery and immune responses to the gene. Therefore, an effective treatment for sickle cell disease is still needed.

#### *G. Our Approach*

Our group is attempting to find a ligand that will bind to sickle cell hemoglobin and inhibit polymerization or increase the delay time before polymerization. We utilize screening techniques such as phage display or one-bead-one-compound library screens to identify novel tightly binding ligands, then test these high affinity ligands to determine whether they can affect the polymerization of hemoglobin.

Ligands that have an effect on polymerization could operate through two distinct mechanisms. One way is through interaction with the residues in either the hydrophobic pocket or the mutated valine and directly blocking polymerization. Akbar et al. developed a 15 amino acid peptide that mimicked the hydrophobic pocket in the EF helix acceptor region and inhibited polymerization *in vitro* by interacting with the valine mutation,<sup>9</sup> showing the feasibility of this approach. Another approach is through interaction with hemoglobin molecule in a way that stabilizes the R state and results in decreased deoxygenated hemoglobin. An example of this method is 5-hydroxymethyl-2-furfural, a Schiff base that works as mentioned above and was shown to inhibit sickling of SS cells at low drug concentrations.<sup>12</sup> These different mechanisms

are shown in Figure 1.7. To find ligands that can work through either of these two mechanisms, the general approach of our group is to identify novel peptidic ligands for hemoglobin, synthesize them using solid phase peptide synthesis techniques, and test them against hemoglobin to determine whether an effect on polymer formation is found.



**Figure 1.7** The possible mechanisms of inhibiting the polymerization of sickle cell hemoglobin. In the first mechanism, a ligand binds somewhere on the hemoglobin molecule and stabilizes the T state. In the second mechanism, a ligand binds to the hydrophobic pocket and prevents it from interacting with the mutated valine. In the third mechanism, a ligand binds to the mutated valine and prevents it from interacting with the hydrophobic binding pocket.



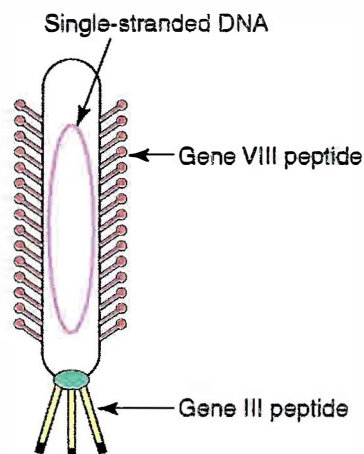
## Chapter II: Phage Display

### A. Phage Display Overview

Previously in our group, a phage display selection was used to identify novel high affinity peptides for deoxygenated hemoglobin. Phage display is a combinatorial technique that allows for rapid screening of a diverse library of biological compounds such as peptides. A population of phage expressing peptides on their surfaces are used to screen a library against target molecules and select for the most tightly binding ligands. This technique was first introduced by G.P. Smith in 1985, and has since then become a powerful tool for elucidating the structures of tightly binding ligands with a wide array of applications.

The phages used in displays are filamentous bacteriophages that are engineered to express a library of peptides on their surfaces. These rod-shaped phage are usually from the *Inoviridae* family and the genus *Inovirus* and include members such as M13, f1 and fd.<sup>16</sup> M13 is the most commonly used phage. These phage infect Gram negative bacteria that contain F-plasmids. The M13 phage has a circular genome comprised of ss-DNA (single-stranded DNA) made up of 6000 to 8000 bases that code for eleven proteins, including coat proteins, DNA replication proteins, and assembly proteins.<sup>17</sup> These phage are used because they can tolerate the insertion of foreign DNA and are able to express the resulting fusion protein without disrupting the normal function of the phage. They are also stable under a variety of harsh conditions such as extreme temperatures and different pH ranges. Finally, they are non-lytic and do not cause bacteria to burst, so they can accumulate to a high concentration within the bacterial cells. These phage are genetically modified to express a library of compounds, such as peptides or proteins, as a fusion protein with one of their coat proteins.

The M13 phage have five different coat proteins named pIII, pVI, pVII, pVIII, and pIX.<sup>16</sup> These proteins have all been made to express peptide libraries on their N-termini or C-termini. The most commonly used coat proteins for phage display are pIII, the minor coat protein, and pVIII, the major coat protein, because they are the most accessible proteins for screening. pVIII is present in high quantities on the outside of the phage, with about 2700 copies tightly packed expressed on the surface of the phage. These proteins can only handle peptide fusions of about ten residues or less onto their N-termini, as longer peptides interfere with the assembly process of the phage. In pVIII phage displays, the affinity of the ligands obtained is usually decreased because so many copies of pVIII are displayed in a small space on the phage. Therefore, sometimes phagemids, which are plasmids that result in the expression of a mixture of wild type and fusion proteins, are used to reduce the number of recombinant pVIII,<sup>18</sup> but these require the use of helper phage to ensure proper working of the screen. The more commonly used capsid protein is pIII. This protein is usually involved in interacting with the F-pilus of the bacterial cells and modulating infection.<sup>19</sup> It can tolerate larger insertions and can handle fusion proteins with peptides of up to 50 residues inserted before its normal function is disrupted. It has a different expression pattern than pVIII and is instead present in three to five copies on the cap of one phage. The locations and expression patterns of these two coat proteins are shown in Figure II.1. Because pIII has reduced valency compared to pVIII, using it for phage display usually results in higher affinity ligands with  $K_d$  values around 1-10  $\mu\text{M}$ , compared to pVIII  $K_d$  values of around 10-100  $\mu\text{M}$ .<sup>16</sup> In our phage display, we utilized M13 phage with the M13KE cloning vector, which results in the expression of peptides on the pIII proteins of the phage.



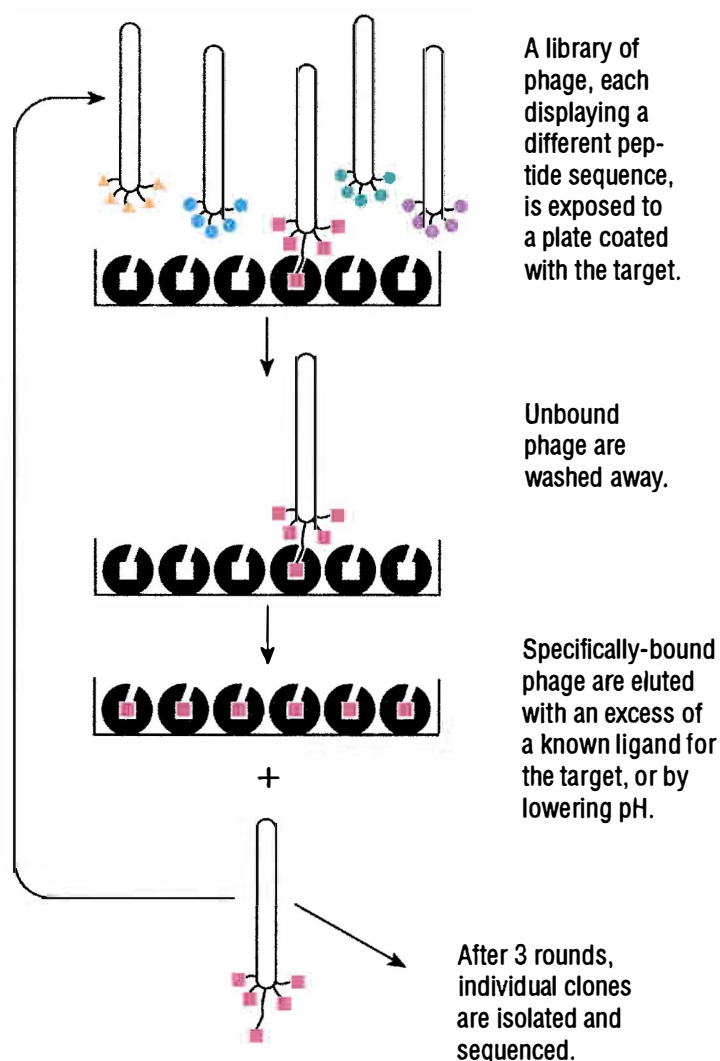
**Figure II.1** An illustration of the location of pVIII and pIII coat proteins. As can be seen in this cartoon, pVIII is expressed at high levels all over the phage, and pIII is expressed in low levels at one end of the phage.<sup>20</sup>

### *B. Phage Display Libraries*

The libraries expressed on the surface of the phage can consist of various types of compounds. The most common are peptide libraries. The combinatorial library is made through recombinant techniques to produce random oligonucleotides which are then introduced into the phage genome and result in the random expression of amino acids.<sup>16</sup> All of the twenty natural amino acids can be used which produces  $20^n$  possible peptide sequences, where  $n$  is the number of residues in the peptide chain. Practically, due to stop codons and degeneracies in the genetic code, only about  $10^9$  peptides of any length can be displayed in a phage library. A common way to produce the random oligonucleotides is through an (NNK) codon degeneracy, where N is a mixture of equal amounts of the four DNA nucleotides and K is an equal mixture of guanine and thymine.<sup>16</sup> This reduces the number of stop codons that arise randomly. There are also a few selective pressures that result in biases in the displayed peptide library. For example, arginine and cysteine residues can interfere with the secretion of pIII, so they can be selected against.<sup>19</sup> In spite of these biases, which can be avoided through extra steps if desired, a massive number of

peptides can be displayed and screened. The large number of peptides produced that can be screened rapidly makes this technique very powerful and ensures the selection of very tightly binding ligands. Using this method, peptides can be found that can be as specific and tightly binding as antibodies. Since peptides are much easier and cheaper to produce than antibodies, this is an exciting avenue for searching for novel drugs and pharmaceutical targets.<sup>20</sup>

After the phage are engineered to express this wide diversity of compounds, they are screened against targets to find the highest affinity ligand. The most common method for *in vitro* phage display screening is biopanning. In the panning method, the target is first immobilized to a plate or well. Then, the phage are added and allowed to bind. The unbound phage are removed through a series of washing steps, each with increasing stringency to ensure that the phage that remain bound are only the high affinity phage.<sup>17</sup> After the washing, the phage is eluted with either the ligand of the enzyme, by a pH change, by using a denaturant, or some other condition that will disrupt the interaction between the phage and the enzyme. Because filamentous phage are very stable, they can undergo this step without being damaged. Finally, the eluted phage is cultured with bacteria to amplify it, and the screening process is repeated one to three times to ensure that only the highest affinity phage are selected. Finally, the DNA sequences of the top hits are obtained and the peptide sequence can be determined. The overview of this process can be seen in the diagram in Figure II.2.



**Figure II.2** The overview of the phage display cycle. The phage library is allowed to bind to the target in step one. In step two, the phage with lesser affinity are washed away. In step three, the tighter binding phages are eluted and amplified. In step four, the cycle is repeated two more times with the amplified phage. After the third cycle, the phage DNA is isolated and sequenced.<sup>19</sup>

Phage display selection experiments can be carried out with a variety of biological compounds. While peptides are the most common, proteins and antibodies can also be expressed in a combinatorial fashion. Unnatural amino acids, such as D-amino acids, and cyclic peptides can also be expressed in a phage display. Phage display can be done *in vivo* to find ligands that can target a specific organ or tissue of the body. During an *in vivo* phage display, the phage is

delivered into a live animal, the animal is sacrificed, and the target organ or tissue is lysed and the phage present identified.<sup>21</sup> Finally, organisms besides phage can also be used to screen diverse libraries. Two common examples are bacteria and yeast, both of which can be engineered to express peptides or proteins fused with their cell surface proteins.<sup>22</sup> However, these organisms are more complex than phage and the other proteins on their surface are more likely to have undesirable interactions with the expressed fusion proteins.

Phage display selection experiments have many advantages over other screening processes. One advantage is the wide diversity of compounds that are expressed. As mentioned, there are usually about  $10^9$  possible peptides displayed; in comparison, chemical libraries that are used in high-throughput screens consist of around tens of thousands of compounds.<sup>23</sup> Another advantage is that this large library can be very rapidly screened, as the phage are small and easily amplified, which allows for the quick determination of high affinity ligands. A third advantage is that the genotype and the phenotype of the phage are linked, since the peptide is displayed on the surface of the phage with the DNA containing the sequence inside of the phage. This makes it applicable for versatile functions while still allowing the easy identification of peptide sequences. Because of these advantages, phage display has been used in a wide variety of applications with great success.

### *B. Experimental*

Previously in our group, a phage display was performed against deoxygenated sickle cell hemoglobin to identify novel high-affinity ligands of the protein. All of the screens were done in an anaerobic chamber. First, a solution of 2 mg/mL of Hb S and 5 mg/mL of sodium dithionite in PBS was shaken with a plate for three hours to coat the plate with deoxygenated Hb S protein. The supernatant was removed and analyzed by ultraviolet visible spectroscopy to ensure that

the hemoglobin was deoxygenated. The spectrum showed a single peak at 554 nm, which is indicative that the hemoglobin was in its deoxygenated state. Then, a buffer comprising of 0.1 M  $\text{NaHCO}_3$  and 5% evaporated milk (blocking buffer) was added to plate and incubated with the Hb S for an hour. The supernatant was removed and the plate was firmly slapped on a paper towel to remove residual solution. The plate was rapidly washed six times with 0.1% TBST and was firmly slapped between each washing. A solution of 1.0 mL of 0.1% TBST and  $2 \times 10^{11}$  of M13 phage was added to the plate and incubated for one hour. The non-binding phage were discarded by removing the supernatant and washing ten times with 0.1% TBST as described above. The bound phage was eluted by shaking 1 mL of a 0.2 M glycine-HCl with a pH of 2.5 on the plate for ten minutes. After elution, the solution was neutralized with 150  $\mu\text{M}$  of 1 M Tris-HCl.

The eluted phage was titered to determine the concentration of phage eluted. To titer, serial dilutions of the phage were prepared and added to a culture of ER2738 (an *E. coli* strain). These were cultured on plates with antibiotics and X-gal. This allowed plaques to be counted and the number of phage in the sample determined.

The eluted phage were amplified by incubating them in a culture of mid-log ER2738 for 4.5 hours at 37°C. Using centrifugation and PEG/NaCl precipitation, the phage was isolated from the bacterial cells and titered. The amplified phage was used for a second round of panning, using 0.5% TBST in the wash steps but keeping the other steps constant. The eluted phage was titered, amplified, then titered again. Finally, a third round of panning with the phage was done. This phage was titered. Bacteria from the plaques from the titer were cultured at 37°C for 5 hours and the phage was obtained through centrifugation and PEG/NaCl precipitation. The phage DNA

was isolated through centrifugation and sent for sequencing. The following peptide sequences were obtained:

**ZSF32** Q-T-H-N-T-E-V  
**ZSF33** S-L-S-D-W-P-Q  
**ZSF34** F-T-P-S-S-D-V  
**ZSF35** T-W-F-R-M-L-S  
**ZSF36** H-L-S-L-T-H-H  
**ZSF37** H-H-Q-F-S-F-N  
**ZSF38** G-S-G-N-L-K-T  
**ZSF39** E-P-W-N-P-I-P

No consensus sequence is seen in these peptides. However, this is not surprising, as the ligands were being screened for general binding affinity for Hb S and not for one particular site of Hb S. These molecules were then synthesized and tested for any effect on polymerization. One molecule, ZSF39, showed to have an inhibitory effect on polymerization.



## Chapter III: Peptide Combinatorial Library Synthesis

### *A. Overview of Peptides as Therapeutics*

Peptides are defined as short chains of less than 50 amino acids. Naturally occurring peptides have a diverse range of applications in the body, including roles as signaling and regulatory molecules, neurotransmitters, and antibiotics.<sup>24</sup> They have also been used as therapeutic agents and have great potential in the upcoming pharmaceutical market. Currently, there are over 60 FDA approved peptide drugs and over 125 peptides in clinical and preclinical trials.<sup>25</sup> The market for peptide drugs has increased over the last decade and is expected to continue expanding in the next few years, growing from a market of \$14.1 billion in 2011 to a predicted \$17 billion in 2018.<sup>26</sup> This growth in the market is due to peptide therapeutics occupying the niche between small molecule drugs and large protein and antibody pharmaceuticals. As the drug market evolves, an increase in the number of peptide drugs is expected.

Peptides as drugs have a number of advantages. Compared to antibodies and proteins, they trigger less of an immune response and have better efficacy because they penetrate farther in tissues due to their smaller size.<sup>27</sup> They also are cheaper to make. Compared to small organic molecules, peptides can have better specificity and are less toxic because the products of their catabolism are amino acids, which are not harmful to the body. Recent updates in technology for peptide screening and synthesis as well as an increased push for safety in pharmaceuticals have attracted more attention for use of peptides as drugs.

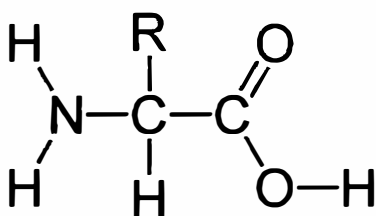
Peptides do have limitations as pharmaceuticals. They have low bioavailability because they are easily metabolized before they can reach their target. One main issue is that they are subject to proteases and cannot be taken orally or they will be easily degraded in the

gastrointestinal tract. They also have difficulty crossing the GI epithelial cells. Consequently, most of the peptide-based drugs on the market must be delivered intravenously.<sup>26</sup> This is not ideal from a patient perspective and can make regular administration difficult. Another issue is that peptides have a short half-life in the body, and are cleared from the plasma rapidly through hepatic and renal elimination. They also have a hard time crossing cellular membranes because of their relative hydrophilicity.

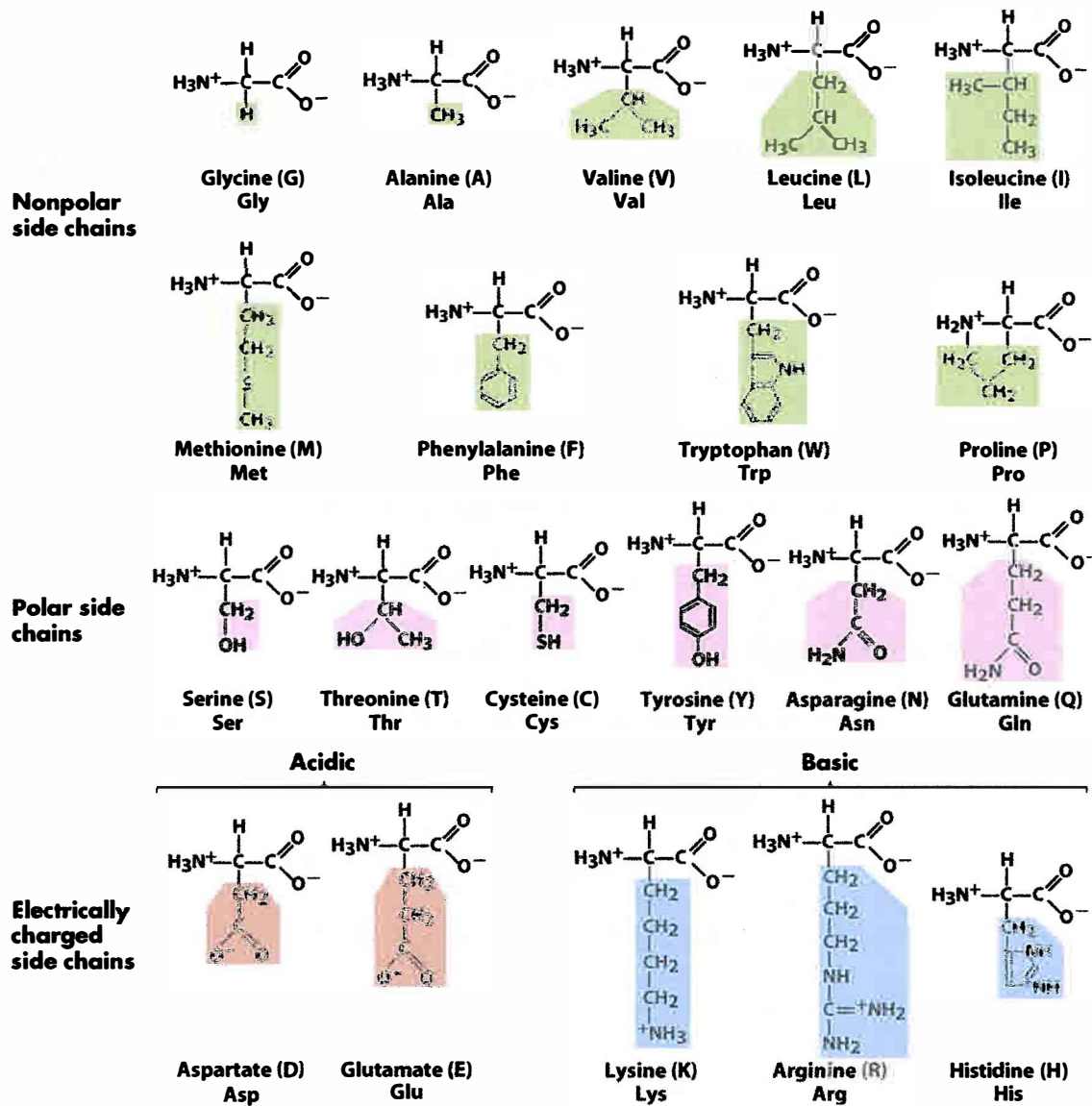
Strategies have been developed to overcome these limitations of peptide pharmaceutical agents. One way is to use unnatural amino acids or other modifications to the basic structure of the peptide to create a molecule with similar properties that will not be targeted by proteases. For example, the drug desmopressin, an analogue of the nine-amino acid long hormone vasopressin, was created by replacing the L-arginine at the eighth position with a D-arginine. This allows desmopressin to be taken orally because it is more resistant to protease activity and can cross the epithelial cells of the gastrointestinal tract more easily, while vasopressin has very low bioavailability when taken orally.<sup>28</sup> Another strategy is to co-administer the peptide with a protease inhibitor to protect the peptides.<sup>29</sup> A defense against plasma clearance is to attach groups to the peptide that target serum albumin, a major blood protein, which will then carry the compound through the blood. A second strategy against elimination is to conjugate the peptide with polyethylene glycol in a process known as PEGylation and increase the size of the overall compound so it is not renally excreted. PEGylation also draws in a larger solvation shell around the peptide which can protect it against proteases and make it less immunogenic.<sup>24</sup> The peptides can also be delivered in a packaging technique, such as liposomes or nanoparticles, that encapsulate the peptide and help shield it against potential degradation.<sup>28</sup> Utilizing these techniques can greatly increase the potential of peptides as therapeutic agents.

## B. The Structure of Peptides

Peptides are comprised of amino acid building blocks. The general structure of an amino acid is shown in Figure III.1. An amino acid consists of a compound containing an amine group, a carboxylic acid functional group, and a side chain, all attached to a central or  $\alpha$ -carbon. Generally, amino acids exist as zwitterions in solution, as the amine functional groups are usually protonated and the carboxylic acid functional groups are usually deprotonated. The specific properties of the amino acid are determined by the side chain, also known as the R group. These side chains range from nonpolar functional groups to polar groups or acidic and basic groups and are unique to each amino acid. There are twenty standard amino acids that make up the majority of the residues found in the body, which are shown in Figure III.2. Almost all of the standard amino acids are chiral and are found in nature almost exclusively as the L enantiomer. The amino acids glycine and proline are exceptions to the typical amino acid structure. Because glycine's side chain is a hydrogen atom, it is achiral and has a more flexible structure than most amino acids. The structure of proline is also unusual because its side chain is a ring, making it a secondary amine.<sup>30</sup>



**Figure III.1** The structure of a general amino acid. Usually, the carboxylic acid moiety is deprotonated and the amine group is protonated, and the structure exists as a zwitterion. The “R” is the side chain, which varies between each amino acid.

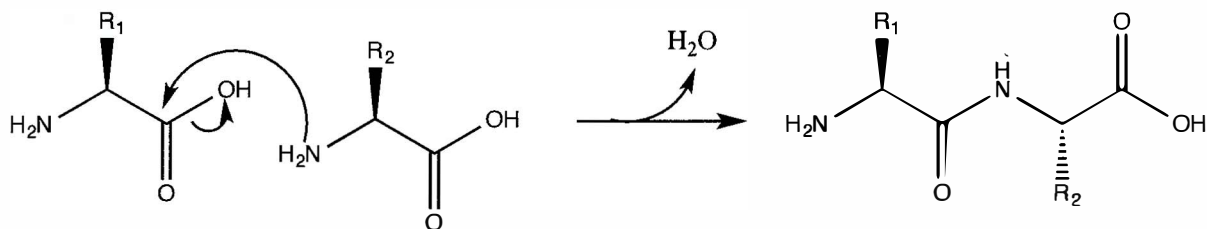


**Figure III.2** The structures of the twenty natural amino acids. Depending on their side chains, they have different properties that give them different functions in the body.<sup>24</sup>

### C. The Synthesis of Peptides

The coupling of amino acids results in the formation of peptides. This is done by a condensation reaction between the carboxylic acid group of one amino acid and the amine of another amino acid, forming water and a peptide bond connecting the two residues, as displayed in Figure III.2. This can be repeated to form a polypeptide chain. In nature, this process is known as translation and is carried out in the ribosome. Translation is used to create longer peptide

chains that eventually adopt tertiary structures and become proteins. Peptides can also be synthesized via solution phase synthesis or solid phase synthesis.



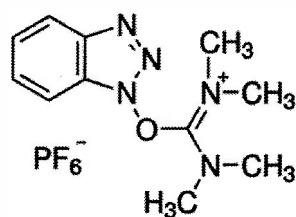
**Figure III.3.** The condensation reaction between two amino acids that results in the formation of a dipeptide.

In solution phase synthesis, all of the reactants are dissolved in a solvent and reacted leading to product. A major drawback of this technique is that the peptide intermediates need to be purified through conventional methods after each step. This can lead to a very time consuming synthesis if longer chains are desired. A heterogeneous phase synthesis of peptides linked to a solid support known as solid phase peptide synthesis (SPPS) is usually utilized because of its ease and utility.<sup>24</sup> In SPPS, the peptide is connected to a solid resin bead via a linker compound and the reagents for synthesis are dissolved in a solvent. The reactions all occur in one reaction vessel fitted with a filter that allows for easy removal of liquids. Because of this facile removal of the excess reagents and impurities through filtration coupled with washing steps, SPPS has extremely simplified purification steps compared to solution phase synthesis. Another advantage of SPPS is that a large excess of reagents can be used, usually between 2-4 equivalents, so the reactions are driven almost all the way to completion. The final yield of the peptide is effectively quantitative, usually above 99.5%.<sup>30</sup> SPPS is also useful because unnatural amino acids or other modifications can be easily introduced, and it can be scaled up to produce short peptides in high numbers.<sup>27</sup> Because of these advantages, most peptide chemists use SPPS to synthesize their compounds.

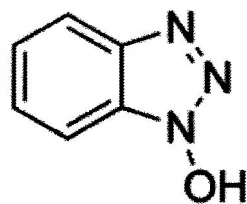
In SPPS, the growing peptide chain begins with a solid resin with linker molecules. The resin, which is comprised of many beads, is generally made of polystyrene crosslinked with divinylbenzene. The resin utilized in this project was TentaGel resin, which is comprised of polyethylene (PEG)-functionalized polystyrene<sup>24</sup> and has a diameter of about 100  $\mu\text{M}$  per bead. The resin used must meet certain criteria to ensure successful peptide synthesis. It must be physically stable and not be damaged during the filtering process. It must also be chemically inert and not participate in the reactions during the peptide synthesis. It must be able to swell in solvent so all of the reactive sites on the bead are accessible.<sup>31</sup> The beads each have about  $10^{13}$  reactive sites and can swell to about six times their dried volume to reveal all of the sites. These reactive sites have linkers attached that will react with the first amino acid of the peptide chain and ensure that proper spacing of the growing peptide chains is maintained in order to prevent aggregation.

After the C-terminus of the first residue is attached to the linker, a series of repeating steps is used to sequentially add amino acids onto the growing peptide chain until the desired molecule is obtained. Unlike peptide synthesis in the body, peptides are synthesized from the C-terminus to the N-terminus in SPPS. First, the carboxylic acid functional group of the amino acid that is being added must be activated, which is often achieved using HBTU. HBTU, the structure of which is shown in Figure III.4, is an uronium reagent that reacts with the carboxylic acid and produces a reactive ester species which the N-terminus of the existing peptide chain can attack.<sup>30</sup> HOBt, or 1-hydroxybenzotriazole, is also used in amino acid coupling as a racemization suppressant. It is a trapping agent that ensures that the reaction does not result in racemization. Its structure is seen in Figure III.5. These two reagents are added with the incoming amino acid in a basic environment to facilitate the coupling of the amino acid onto the existing peptide

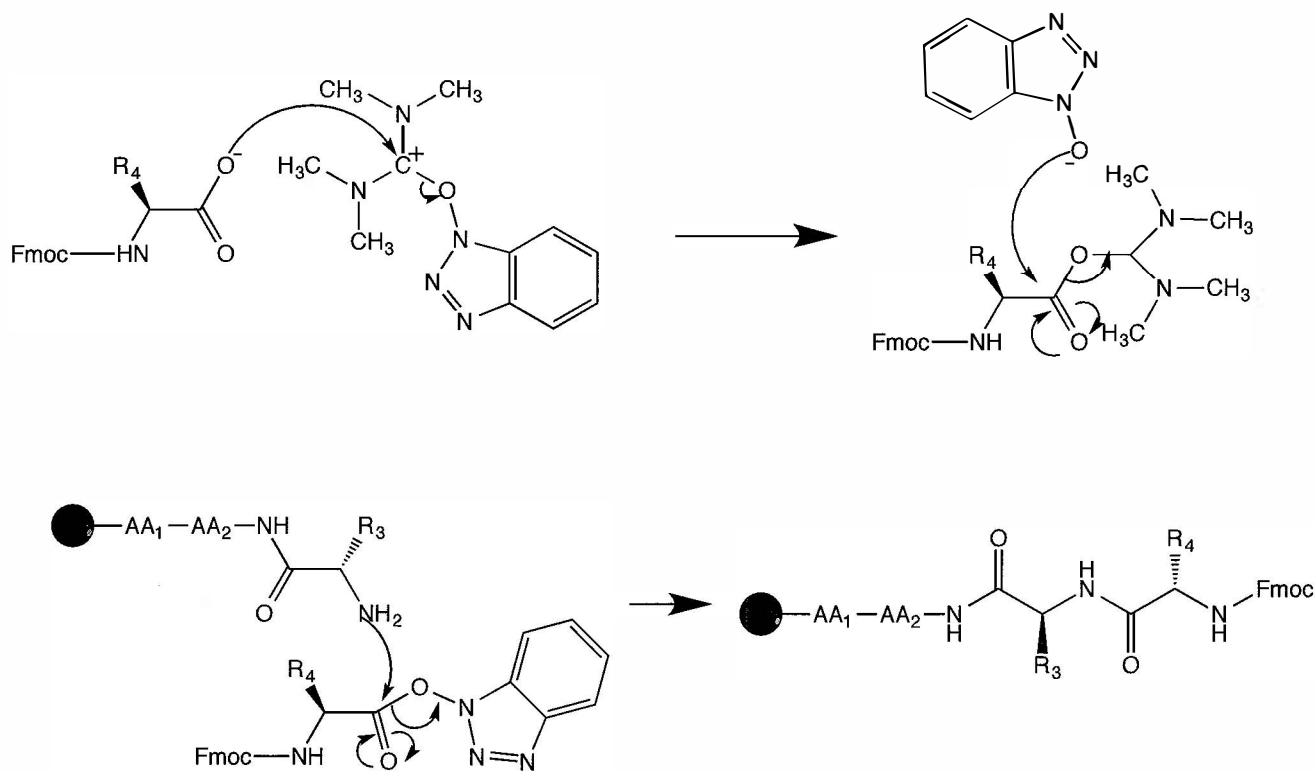
chain. The reaction of the HOBt and HBTU with the amino acid that is being added to the chain is depicted in Figure III.6.



**Figure III.4** The structure of HBTU, the coupling reagent used to activate the carboxylic acid of the incoming amino acid.

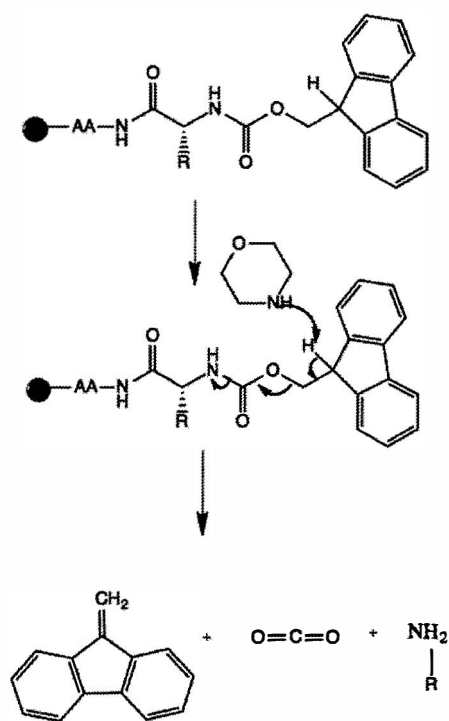


**Figure III.5** The structure of HOBt, the trapping agent used to prevent racemization of the growing peptide chain.



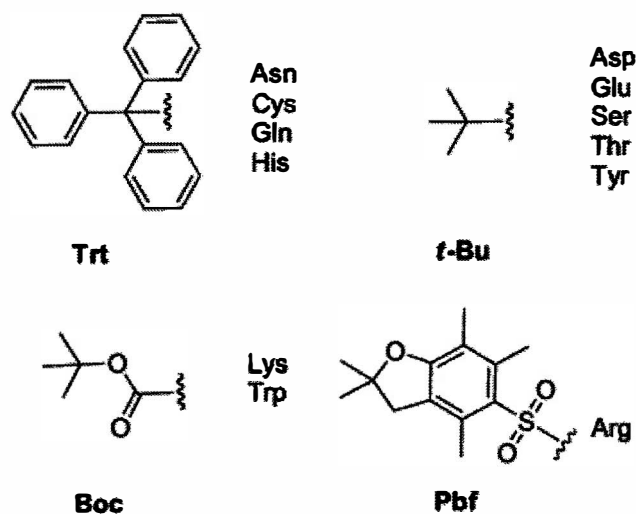
**Figure III.6** The mechanism of amino acid coupling using HOBt and HBTU in a basic environment. In the first step, the amino acid to be added attacks the HBTU, forming an activated ester. Next, HOBt attacks the amino acid, forming another activated ester that cannot be racemized. Finally, the amine of the peptide chain attacks the incoming amino acid, forming the peptide bond and completing the coupling.

The amino ends of the incoming amino acids, known as the  $N^\alpha$  amino groups, are protected to ensure that only one residue is added to each chain and that there are no undesired side reactions. Fmoc (9- fluorenylmethoxycarbonyl) groups are commonly used to protect these ends. The Fmoc protecting groups are base labile and are removed after every amino acid addition. The removal of an Fmoc group by morpholine, a weak base, is depicted in Figure III.7. Reactive side chains of amino acids also need to be protected to ensure they do not participate in side reactions; however, these protecting groups must be orthogonal to the Fmoc groups because they are not removed until all of the amino acids have been added, and therefore are usually acid labile and removed using trifluoroacetic acid (TFA).<sup>24</sup> The structure of the common side chain protecting groups can be seen in Figure III.8.



**Figure III.7** The structure of a peptide chain on a resin with an Fmoc protected  $N^\alpha$  amino group, and the removal of the Fmoc compound by morpholine.

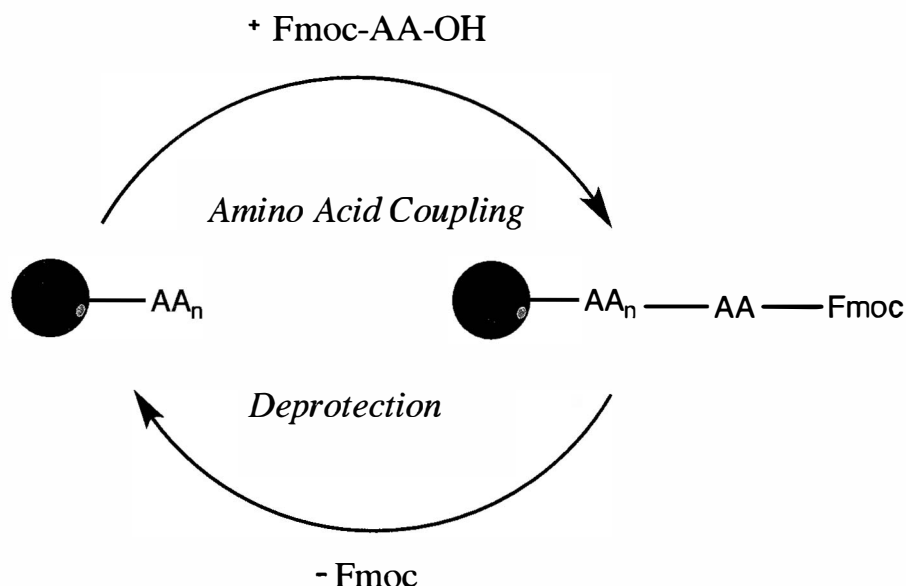




without SPG: Ala, Gly, Ile, Leu, Met, Phe, Pro, Val

**Figure III.8** Typical amino acid side chain protecting groups for Fmoc/*t*-Bu coupling. “Trt” stands for trityl and “Pbf” stands for pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl. These are used to ensure that reactive side chains do not undergo undesirable reactions and are removed by TFA at the end of the synthesis.<sup>24</sup>

After the first amino acid is added to the resin, a cycle of deprotection of the Fmoc group and coupling of the next amino acid is repeated until the desired sequence is reached. A depiction of this cycle can be seen in Figure III.9. The process can be tracked using a ninhydrin test, or a Kaiser test, to monitor whether primary or secondary amines are present in order to determine whether reactions have gone to completion. This simple two-step method makes peptide synthesis facile and allows for modifications such as the addition of an unnatural amino acid. It also can be automated easily, so many automated peptide synthesizers have now been made and fine tuned to synthesize many different types of peptides.<sup>32</sup> Using this method, peptide chains with a maximum of about 50 residues can be synthesized easily and in high yield. After the peptide chain is synthesized, it is cleaved from the resin using TFA, which also cleaves the side chain protecting groups. After cleaving, the peptide can be purified through standard analytical methods such as chromatography and identified using mass spectrometry.



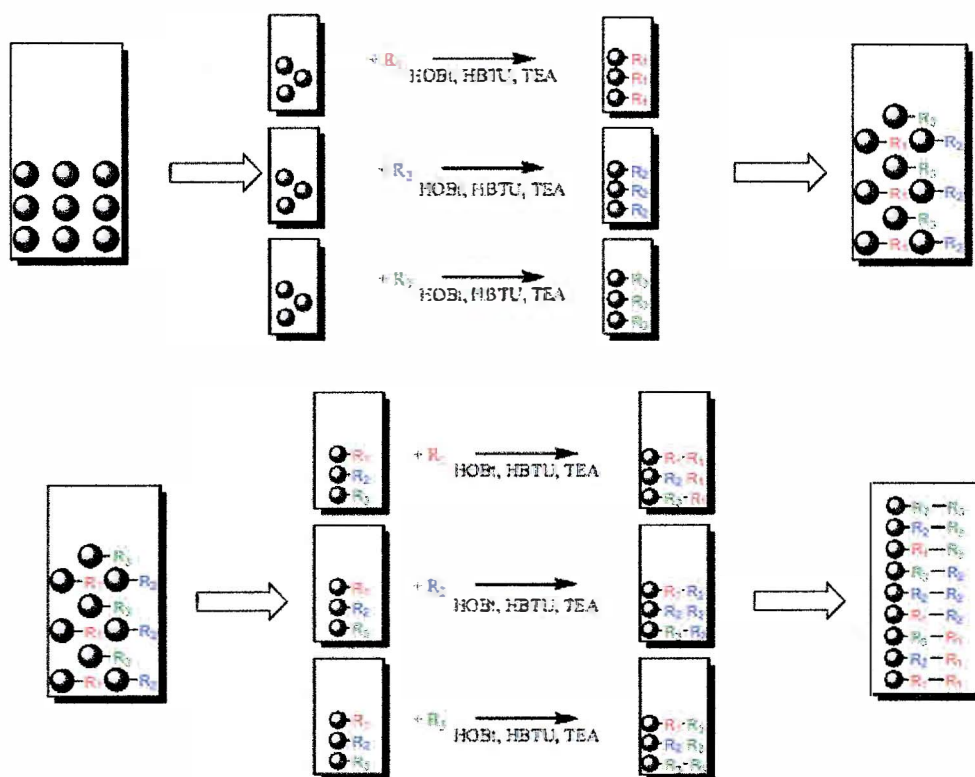
**Figure III.9** A depiction of the cycle of SPPS. After the first amino acid is initially bound to the resin, it is washed and deprotected by a weak base. It is washed again and another amino acid is coupled using coupling reagents. This two step cycle can be repeated until the desired peptide is synthesized. It is then cleaved from the resin. In this figure, “P” represents Fmoc protecting groups.

#### *D. Combinatorial Peptide Libraries*

Another advantage of SPPS is that it can easily be adapted to combinatorial chemistry and used to make an incredibly diverse library. These libraries can be used in screens to identify lead compounds with a desired effect from a huge number of possibilities. Combinatorial chemistry has a few advantages over traditional chemical libraries screening. It can be used to generate huge libraries quickly, instead of synthesizing each compound individually. Also, because a larger number of compounds are being screened, it is more likely that a substantial hit will be identified.<sup>31</sup>

To synthesize a combinatorial library a one-bead one-compound technique is used. In this method, which is also known as a split-and-mix synthesis, the positions of the amino acids are randomized. To do this, beads are split into equal groups and each individual group of split beads receives a different amino acid. Then the groups are all re-pooled and thoroughly mixed. The

splitting process is then repeated so each group has an equal amount of beads from the previous groups, and another randomized amino acid is added. This method is repeated until the chain of the desired length is synthesized. This method ensures that every possible peptide is made, and it should give an equal amount of each compound.<sup>32</sup> For a peptide library, the number of compounds generated is equal to  $A^n$ , where  $A$  is the number of amino acids used in synthesis (usually 20) and  $n$  is the number of amino acids in the final sequences. This library synthesis is known as one-bead one-compound because each bead only contains peptides at its reactive sites with a single sequence. This allows for screening based on bead and easy determination of the sequence of the peptides on that bead.



**Figure III.10.** An outline of the synthesis of a combinatorial library. First, the resin beads are split into equal fractions and a different amino acid is added to each fraction. The beads are pooled again, mixed equally, then split into fractions again. Each fraction contains an equal number of beads from each fraction in the previous step. Another randomized amino acid is added to each fraction. This process can be repeated until the chain of desired length of randomized positions is synthesized.

The sequences of unknown peptides can be identified through processes such as Edman degradation. In Edman degradation, the amino acid at the N-terminus of the peptide is reacted with phenylisothiocyanate and cleaved from the chain. The resulting compound produced can then be identified via HPLC or some other analytical method. This cycle is repeated sequentially until the entire peptide has been sequenced. The reaction and removal of the N-terminus in Edman degradation can be seen in Figure III.11.

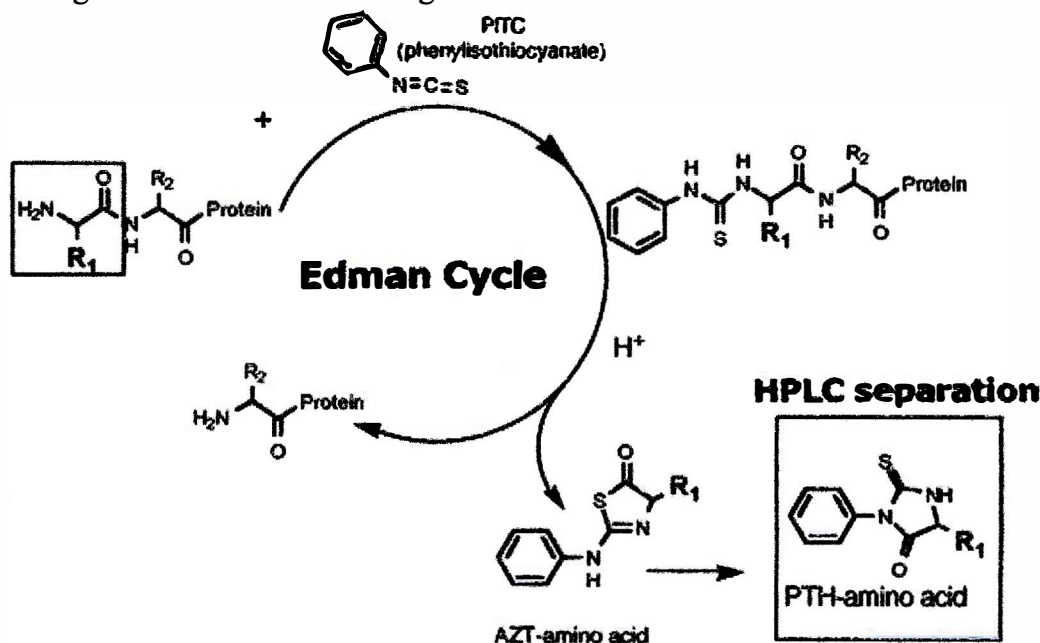


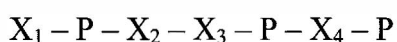
Figure  
III.11 A  
general

overview of Edman degradation. The N-terminus of the peptide to be identified is reacted with phenylisothiocyanate and forms a PTH-amino acid compound that can be identified by a characteristic HPLC peak. This cycle can be repeated until all of the amino acids comprising the peptide have been identified.<sup>33</sup>

#### *E. Goal of the Combinatorial Proline-Rich Peptide Library*

Combinatorial peptide chemistry was utilized in this project to generate a peptide library based on the results from the phage display. The inhibition of polymerization by ZSF39 is intriguing, as ZSF39 contains three prolines, the amino acid with a ring structure as a side chain. Proline is unique because its constrained conformation produces a kink in the backbone of the

peptide. Three prolines in a seven-mer peptide suggests that the proline have a significant effect on the properties and structure of the peptide. Therefore, we wanted to further investigate how the structure of this peptide affects polymerization and if the non-proline positions of ZSF39 could be optimized to produce a stronger effect on polymerization. To do this, a combinatorial library was synthesized in which the three proline positions were kept constant and the other positions of the peptide were randomized. The final compounds had the sequence seen in Figure III.10.



**Figure III.10** The sequence of the peptides synthesized in the combinatorial library, where “X” is a randomized amino acid.

#### *F. Experimental*

The solid support used for the synthesis was 200 mg (0.0518 meq) of TentaGel R-NH<sub>2</sub> resin, a non-cleavable resin. It was swelled in a reaction vessel with about 4 mL of 50% morpholine in dimethylformamide (DMF) for an hour. It was washed three times with DMF. A coupling solution of 3 molar equivalents each of Fmoc-L-Pro-OH, triethylamine (TEA), HOBt, and HBTU dissolved in about 4 mL DMF was added to the resin and allowed to react for 1.5 hours. The resin was washed three times with DMF. A Kaiser test was performed to ensure that the coupling reaction had gone to completion. The solution was removed by filtration and the resin was washed three times for five minutes each with 4 mL of 50% morpholine in DMF to remove the Fmoc protecting group. Another Kaiser test was performed to ensure that all of the protecting groups had been removed. The resin was washed with DMF three more times, then split equally into sixteen different reaction vessels. Separate coupling solutions containing 0.2 molar equivalents to the initial amount of resin were made of HOBt, HBTU, TEA, and one Fmoc-protected amino acid dissolved in 0.75 mL of DMF. All of the standard amino acids were

used except tryptophan, methionine, and cysteine, as they are susceptible to oxidation, and proline, to ensure that the sequences that were synthesized contained only the three prolines in the positions desired. The coupling reaction was allowed to run overnight. The beads were repooled, then deprotected using morpholine washes and checked with a Kaiser test. These processes were repeated to produce a peptide with prolines at positions 2, 5, and 7, and randomized amino acids at the other positions. The total number of possible sequences synthesized was  $16^4$ , or  $6.5 \times 10^4$ . After the synthesis of the chain was completed, the side chains were deprotected using a 4 mL solution of 95% trifluoroacetic acid (TFA), 2.5% triisopropylsilane (TIS), and 2.5% deionized water.

The success of the synthesis was assessed with Edman degradation. Two beads at random were selected to undergo sequencing. The sequence of the first bead was Arg-Pro-Glu-Gly-Pro-Ser-Pro, and the sequence of the second bead was Val-Pro-Phe-Ser-Pro-Gly-Pro. They both contain proline at the 2, 5, and 7 positions and a mixture of the other amino acids used at the randomized positions. Therefore, the synthesis of the randomized library was determined to be a success, and the library was ready to be screened for binding affinity.

## Chapter IV: ELISA Screen and Polymerization Assays

### *A. ELISA Overview*

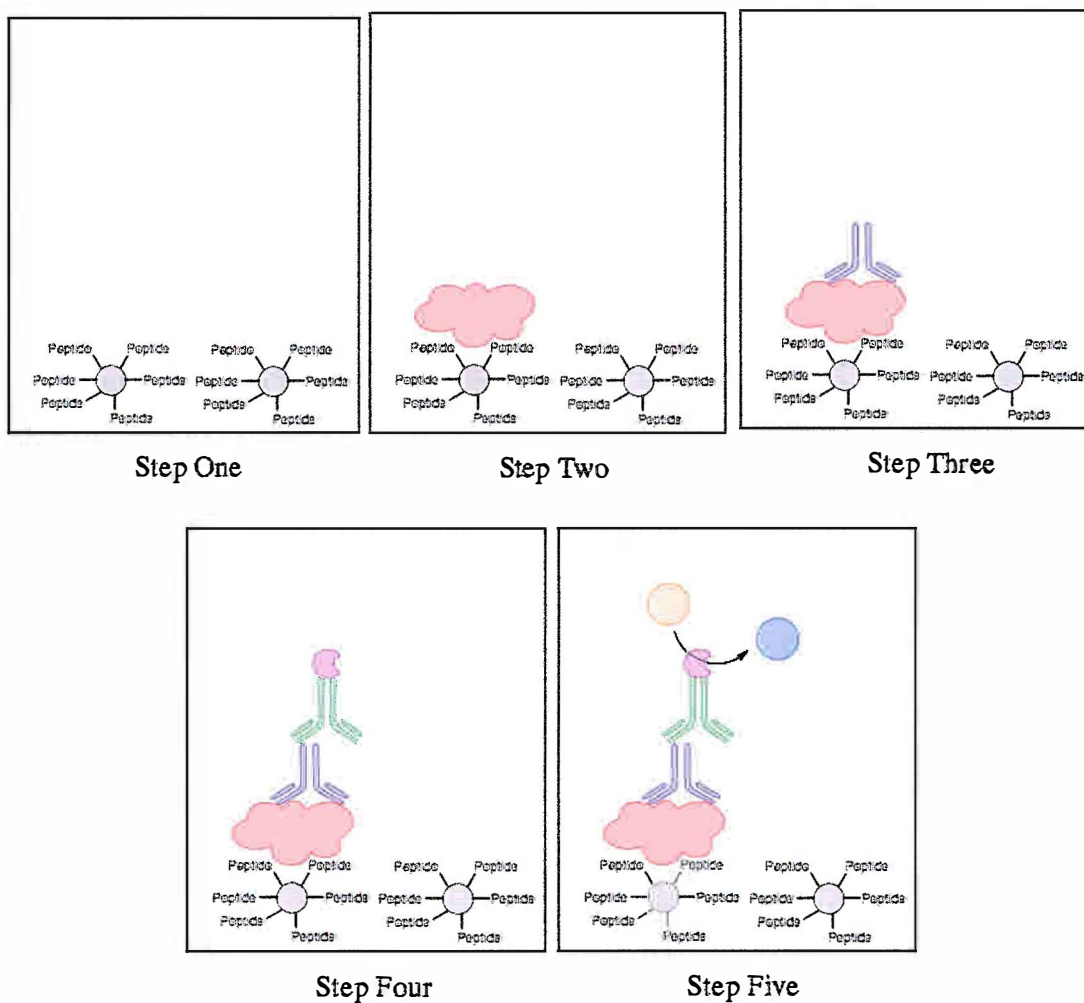
Three different methods can be used to screen for a desired property in a peptide combinatorial library. The first is to use a biologic peptide method, such as a phage display.<sup>34</sup> The second is to synthesize the library on a solid support, cleave the compounds from the resin, then run the screen. The third method is to synthesize the library on the resin in a one-bead one-compound fashion, run the screen with the peptides remaining on the bead, and then select the beads with the desired properties to analyze. We chose to analyze the library through the third method to ensure easy identification of hit compounds, and used an ELISA, or enzyme-linked immunosorbent assay, to select for the peptides in the combinatorial library that had the highest binding affinity to hemoglobin. An ELISA is a simple technique where enzyme chemistry and immunology are linked to identify the presence of an antigen with high sensitivity.<sup>35</sup> It is a common technique utilized in many laboratories.

An ELISA can be performed as heterogeneous assays or homogenous assays. As the name suggests, homogenous ELISAs take place in one phase, generally solution phase. These assays require all of the reagents to be added simultaneously as the intermediates cannot be purified.<sup>36</sup> Because of this, these ELISAs have reduced efficiency and are less popular. More often, heterogeneous ELISAs are performed, which involve adhering one of the reagents, generally the antigen, to a solid phase. Because this solid phase can be easily separated from solvents, washing steps are simple; thus, reagents can be added sequentially with washing between each step.

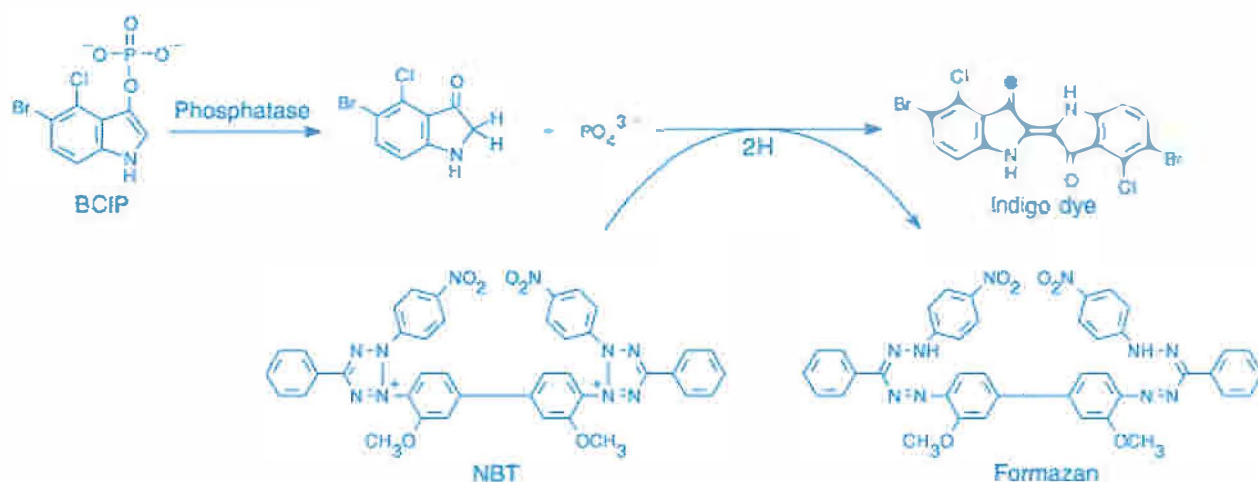
An ELISA can be utilized with a one-bead one-compound peptide library to select for the peptides from a combinatorial library that bind the most tightly to an antigen. A depiction of this

process can be seen in Figure IV.1. First, the antigen, which is deoxygenated Hb S in this experiment, is diluted in phosphate-buffered saline (PBS) and incubated in a plate with the combinatorial library, allowing it to interact with the peptides on the beads. The beads containing peptides that bind to the antigen more tightly will bind to a lot of the antigen, while the beads containing weakly-binding peptides will only bind a small amount of antigen and the beads containing nonbinding peptides will bind a negligible amount of the antigen. After this differential binding process takes places, the excess antigen is washed off and a solution of primary antibodies, which bind specifically to the antigen, is added. This is added in a blocking buffer that has detergents or other compounds that block nonspecific interactions to prevent the antibodies from binding with the parts of the plate instead of the antigen. After the antibodies bind, the excess primary antibodies are washed off, and a solution of secondary antibodies is added. Secondary antibodies bind specifically to the primary antibody, and are linked to an enzyme such as alkaline phosphatase that catalyses a colorimetric reaction. After these are bound and the excess is washed off, a reagent such as BCIP (5-bromo-4-chloro-3-indyl phosphate) is added to react with the enzyme and produce a colored product as shown in Figure IV.2 that precipitates on the bead,<sup>37</sup> which allows for identification and quantification of the antigen. The beads that have the most tightly binding peptides interact with the highest amount of antigen and therefore will have the greatest color change to a deep purple. Beads that contain weakly binding and nonbinding peptides have a small or no color change.





**Figure IV.1** The overview of the ELISA screen with a one-bead one-compound combinatorial library. The antigen (Hb S) interacts with the peptides. Next, the primary antibodies are added and bind to the antigen. In the next step, the secondary antibodies are added and bind to the primary antibodies. Finally, the substrate for the conjugated enzyme is added and the enzymes catalyzes a colorimetric reaction. The end result is a color change where antigen is present.



**Figure IV.2.** The reaction of BCIP and NBT to produce two colored products, indigo dye and formazan.

These types of ELISAs have distinct advantages as screening methods. The first is that they are simple and widely used. The materials needed are relatively inexpensive and safe, and the screening process is rapid. Millions of beads can be tested in a short amount of time.<sup>37</sup> Also, the process is easily automated.<sup>38</sup> The results can be quantified by measuring the color intensity. The process is sensitive, and levels of 0.01 to 1  $\mu\text{g/mL}$  of antigen can be identified.<sup>36</sup> Because of these advantages, peptide screening using ELISA has become a widely used tool. For example, ELISAs are used frequently in diagnosing bacterial and viral infections because of their high sensitivity and simplicity.<sup>39</sup> Therefore, we utilized an ELISA to screen the combinatorial library to find peptides that bind tightly to sickle cell hemoglobin.

### *B. ELISA Experimental*

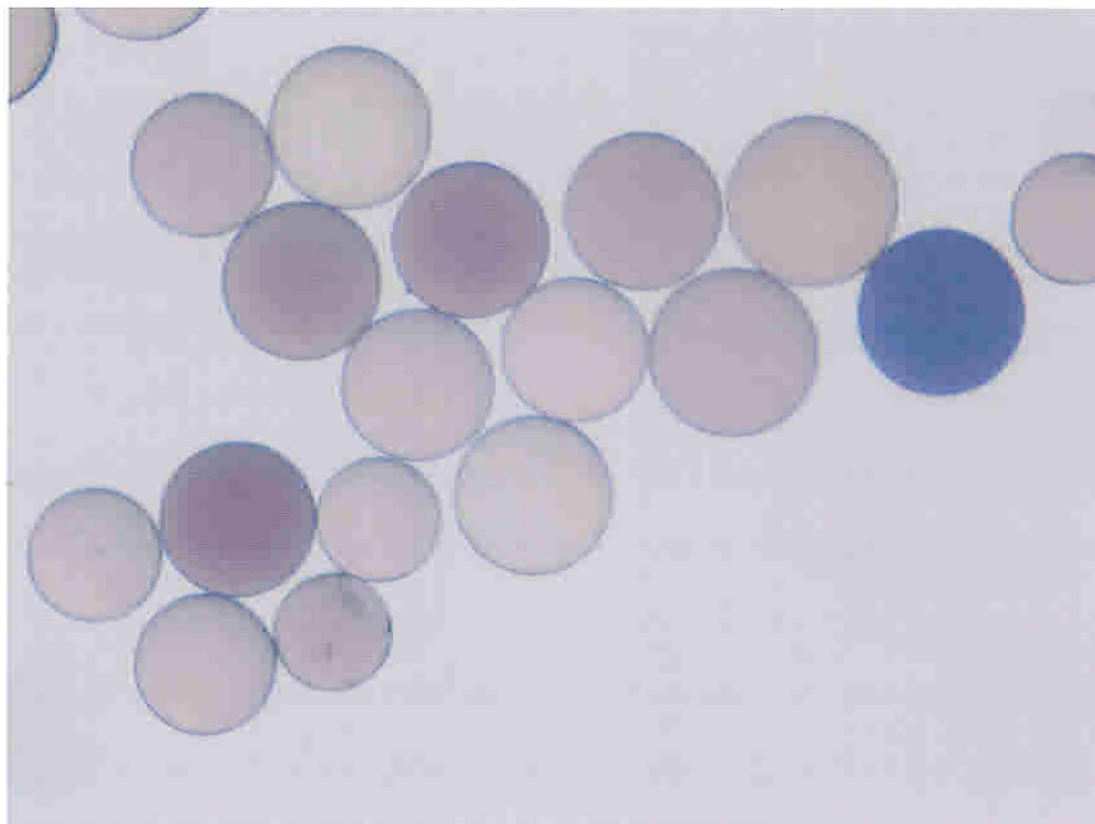
Before the ELISA was run, a Western blot was performed to ensure that the stock solutions of primary and secondary antibodies were functional. Solutions of sickle cell hemoglobin with concentrations of 0.2 mg/mL, 0.04 mg/mL, and 0.008 mg/mL, loading buffer that contained 50 mM Tris-HCl, 2.5% SDS, 0.02% Bromophenol Blue, 5%  $\beta$ -mercaptoethanol, and 10% glycerol, and 50 mM PBS were heated at 100°C for ten minutes. The solutions were run

on a SDS polyacrylamide gel which was then stained with Gel CodeBlue. The proteins were transferred onto an Immobilon-P membrane. The membrane was shaken with a blocking buffer of 0.1% TBST with 5% evaporated milk for 90 minutes. A primary antibody solution with a 1:200 dilution of the goat anti-hemoglobin stock solution in blocking buffer was shaken with the membrane overnight at 4°C. The membrane was rinsed four times with 0.1% TBST and then incubated for 90 minutes with a solution with a 1:5000 dilution of the rabbit anti-goat stock solution in blocking buffer. The membrane was washed four times with 0.1% TBST and alkaline phosphate (AP) buffer. A solution with a 1:100 dilution of NBT/BCIP stock solution in AP buffer was shaken with the membrane until color developed. The blot developed purple regions, showing that the hemoglobin in the solution was successfully identified with the reagents used and the stock solutions were functioning correctly.

Next, the library was screened using an ELISA. All of the steps were performed in the anaerobic chamber to test deoxygenated hemoglobin. The library was placed in a 35 by 10 mm plate and washed thoroughly with methanol then ultrapure water to ensure purity. It was then shaken with blocking buffer for three hours. It was rinsed three times with PBST and one time with TBS. A solution was prepared of 2 mg/mL of Hb S, 5 mg/mL of sodium dithionite to deoxygenate the hemoglobin, and 0.1% TBST. This was shaken with the resin for three hours. The supernatant was collected and analyzed by ultraviolet visible spectroscopy to ensure that the hemoglobin remained deoxygenated. The spectrum showed a single peak at 554 nm, which is characteristic of deoxygenated hemoglobin. The resin was then washed three times with PBST and once with TBS. A primary antibody solution with a 1:200 dilution of the goat anti-hemoglobin stock solution in blocking buffer was shaken with the resin for three hours. The supernatant was removed and another washing cycle was performed. A secondary antibody

solution with a 1:5000 dilution of the rabbit anti-goat stock solution in blocking buffer was shaken with the resin for three hours. The resin was washed three times with PBST, once with TBS, and once with AP buffer. A solution with a 1:100 dilution of NBT/BCIP stock solution in AP buffer was shaken with the resin until a purple color was visible on some of the beads. Finally, the colorimetric reaction was terminated with a few drops of hydrochloric acid.

The appearance of the beads after the screen was completed can be seen in Figure IV.3. A gradient of colors from colorless to deep purple was seen, indicating that the screen was successful in separating strongly binding peptides from weakly binding and non-binding peptides. The beads with the peptides that bound the most strongly to the sickle cell hemoglobin had the strongest response to the ELISA and had the darkest color.



**Figure IV.3** The appearance of the beads after the ELISA screen. A gradient of coloration was seen, indicating that the screen was able to effectively distinguish tightly binding peptides from loosely binding or nonbinding peptides. Each bead is about 100  $\mu\text{M}$  in diameter.

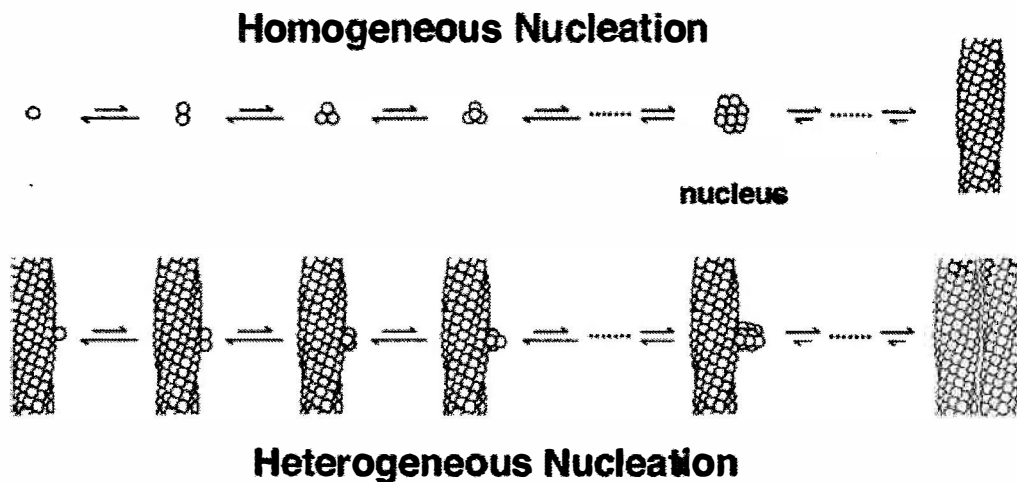
The beads that were colored the darkest were isolated and sequenced by Edman degradation. Unfortunately due to sequencing errors due to improper washing of the final bead, the majority of the peptides that were isolated could not be identified. However, the sequence of one of the peptides was obtained, Lys-Pro-Asp-Val-Pro-Phe-Pro, named LHS1. To compare, the sequence of ZSF39 is Glu-Pro-Trp-Asn-Pro-Ile-Pro. While the position of the prolines are obviously kept constant, the other positions do not show any similarity. For example, at position 1, LHS1 has a positively charged lysine residue, while ZSF39 has a negatively charged glutamic acid residue. If these two peptides have similar mechanisms of action, it is probably due to the effect of the prolines on the structure of the backbone and not interactions from the other four positions. It is also possible that LHS1 is binding to different area of hemoglobin than ZSF39 and will either have a different mechanism of inhibiting polymerization or will not affect polymerization.

### *C. Polymerization Assay Overview*

To determine if LHS1 has any activity against polymerization, polymerization assays were performed. In a polymerization assay, the hemoglobin is deoxygenated with sodium dithionite and incubated with the peptide of interest in a high salt buffer of PBS at 30 °C. The status of polymerization is tracked using ultraviolet visible spectroscopy. As the hemoglobin polymerizes, the solution becomes more turbid, causing a change in absorption. Hemoglobin polymerization is characterized by a delay time between deoxygenation and when polymers start to form. This delay time is very dependent upon concentration, and is reproducible in macroscopic volumes but not at very small concentrations. It has been the subject of a lot of study, as the ramifications could have clinical significance. This delay time is often longer than the transit time for a red blood cell in circulation, meaning usually hemoglobin in erythrocytes is

reoxygenated before polymerization can occur in the body.<sup>40</sup> This explains why the symptoms of sickle cell disease only occur in patients when the deoxygenation process is exceptional, for example in times of exertion when more oxygen is released to the muscle cells.

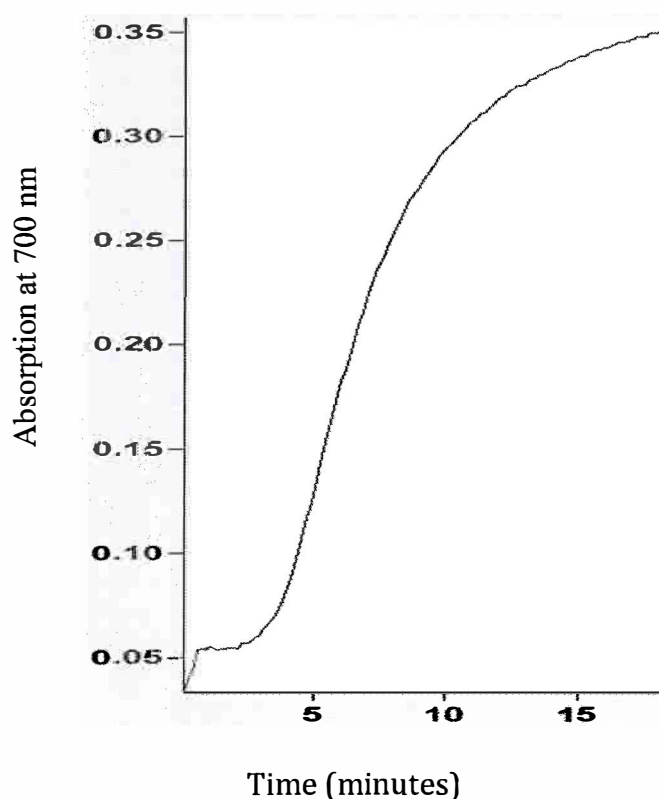
The kinetics of the delay time have been attributed to two types of nucleation, homogenous and heterogenous.<sup>41</sup> When the hemoglobin is initially deoxygenated, the monomers randomly aggregate to form nuclei in the solution. These nuclei are not very stable and tend to fall apart easily. Therefore, many nuclei must form before one is stable enough to allow for further aggregation and the formation of polymers.<sup>9</sup> This is known as homogenous nucleation, and is a fairly slow process. After homogenous nucleation has occurred and polymers have formed, new nuclei can form on the surface of the polymers. These nuclei are much more stable than the ones formed by homogeneous nucleation, so polymerization occurs more rapidly, leading to an exponential growth of polymers.<sup>10</sup> The two nucleation processes can be seen in Figure IV.4.



**Figure IV.4** The two models of nucleation. In homogenous nucleation, nuclei must come together and form a stable aggregate in order to build a polymer. This process is slow because the aggregate is often unstable. In heterogenous nucleation, in comparison, nuclei grow on the surface of the polymers already created and are much more stable. This happens at a quicker rate and generates exponential polymer growth.<sup>10</sup>

#### *D. Polymerization Assays Experimental*

Polymerization assays were done on samples with hemoglobin with no peptide added as a control, hemoglobin with LHS1 added, and hemoglobin with ZSF39 added. To run an assay, solutions were made of 5 mg/mL sodium dithionite and 1.87 M PBS and incubated at 0 °C for ten minutes. Other solutions of 0.69 mg/mL of hemoglobin, 100  $\mu$ M of the peptide, and enough 50 mM PBS to bring the total volume to 500 mL were made and incubated at 4 °C for ten minutes. After incubation, the two solutions were mixed and their temperatures were raised to 30 °C. The absorptions of the solutions at 700 nm were measured until the maximum polymerization had been reached. An example of a typical polymerization curve can be seen in Figure IV.5.



**Figure IV.5** A typical output from a polymerization assay. As can be seen, after a short delay time, the hemoglobin polymerizes in an exponential fashion, causing the solution to become more turbid and allowing less light to be transmitted. The delay time can be calculated to compare the efficacy of different polymerization inhibition methods.

Polymerization assays of 100  $\mu$ M of LHS1, ZSF39, and control solutions were performed simultaneously and the fold change of each delay times as compared to the control was calculated. The results from the polymerization assays can be seen in Tables IV.1 and IV.2.

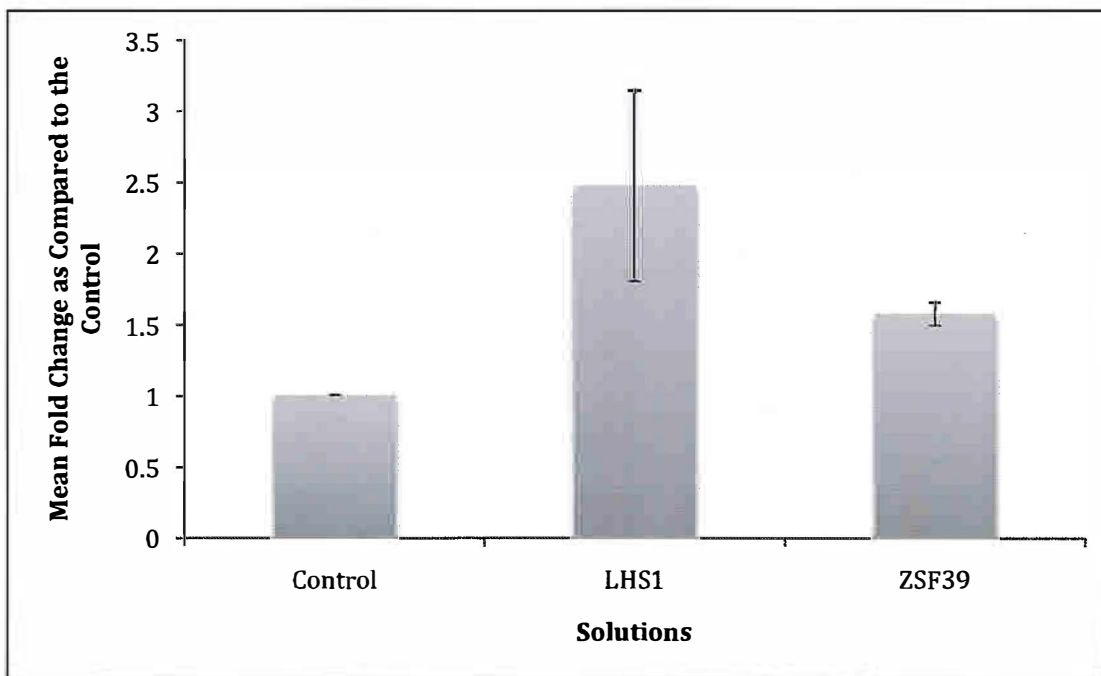
**Table IV.2** The delay times of the control solutions, LHS1 solutions, and ZSF39 solutions as compared to the control.

<b>Solution</b>	<b>Control</b>	<b>LHS1</b>	<b>ZSF39</b>
Delay Time for Trial 1 (sec)	2.57	7.52	4.20
Delay Time for Trial 2 (sec)	3.24	9.04	4.78
Delay Time for Trial 3 (sec)	7.03	11.86	11.12
Mean of Delay Times	4.28	9.47	6.70
Standard Deviation of Delay Times	2.41	2.20	3.84

**Table IV.2** The fold changes for the delay times of the control solutions, LHS1 solutions, and ZSF39 solutions as compared to the control.

<b>Solution</b>	<b>Control</b>	<b>LHS1</b>	<b>ZSF39</b>
Delay Time Fold Change for Trial 1	1	2.92	1.64
Delay Time Fold Change for Trial 2	1	2.80	1.48
Delay Time Fold Change for Trial 3	1	1.70	1.60
Mean of Fold Changes	1	2.47	1.57
Standard Deviation of Fold Changes	0.00	0.67	0.08





**Figure IV.6** The mean fold changes for the delay times as compared to the control for the control solutions, the solutions with LHS1, and the solutions with ZSF39, with error bars corresponding to the standard deviation.

### *Conclusions and Future Work*

These exciting results show that LHS1 is a peptide that shows promise in inhibiting sickle cell hemoglobin polymerization. It is able to extend the delay time significantly longer than both Hb S without a peptide ligand and Hb S with ZSF39. This result opens up many avenues for future exploration. One factor to explore is whether LHS1 and ZSF39 are interacting with hemoglobin in the same area of the protein, or if they have different mechanisms of affecting polymerization. Because they have a similar backbone structure but the other four positions are very different between the two peptides, a possible hypothesis would be that the secondary structure allows it to interact with Hb S in a particular fashion and that the other four residues of LHS1 were optimized for higher affinity to Hb S than ZSF39. It would be insightful to take other routes, such as better preparation of the samples before sequencing, to determine the structure of

the other peptides identified from the screen and see if any similarities are present between them and LHS1, and what types of residues contribute to a larger effect on polymerization delay times. Another interesting direction would be to establish where the peptides are interacting with the hemoglobin. To have an inhibitory effect on polymerization, the peptides could either be interacting with the oxygenated form of Hb S in a stabilizing manner or with the deoxygenated form in a way that blocks polymerization. Because the peptides were both identified from deoxygenated screens, the mechanism of inhibition is probably with the deoxygenated form. Therefore, the peptides are most likely either interacting with the hydrophobic pocket or the mutated valine and directly blocking their interaction, or interacting with another portion of deoxygenated sickle cell hemoglobin and causing a slight conformational change that inhibits polymerization. Future work in this area should be conducted to elucidate the nature of the structure-function relationship of this peptide and how it can be further optimized to understand how it is functioning and how it can be made better.

- <sup>1</sup> Abboud, Miguel R., and Khaled M. Musallam. "Sickle Cell Disease at the Dawn of the Molecular Era." *Hemoglobin* 33.S1 (2009): S93-106.
- <sup>2</sup> Williams., Thomas N., and Stephen K. Obaro. "Sickle Cell Disease and Malaria Morbidity: A Tale with Two Tails." *Trends in Parasitology* 27.7 (2011): 315-20.
- <sup>3</sup> Bender, MA, and Gabrielle Douthitt Seibel. "Sickle Cell Disease." (1993).
- <sup>4</sup> Parmar, Dinesh. "Sickle Cell Anemia: Review And Remedial Hope." *Egyptian Journal Of Medical Human Genetics* 10.2 (2009): 130-143.
- <sup>5</sup> Casiday, Rachel and Regina Frey. "Hemoglobin and the Heme Group: Metal Complexes in the Blood for Oxygen Transport". *Washington University in St. Louis Department of Chemistry*. 5 Sept 2008  
<<http://www.chemistry.wustl.edu/~edudev/LabTutorials/Hemoglobin/MetalComplexinBlood.html#References>>.
- <sup>6</sup> Mathews, Christopher, Kensal Van Holde and Kevin Ahern. 2000. *Biochemistry* 3 rd edition.  
<http://www.aw-bc.com/mathews/ch07/c07emhp.htm>
- <sup>7</sup> Damanhour, Ghazi, Jummanah Jarullah, Samy Marouf, S.i. Hindawi, Gohar Mushtaq, and Mohammad A. Kamal. "Clinical Biomarkers in Sickle Cell Disease." *Saudi Journal of Biological Sciences* (2014)
- <sup>8</sup> Harrington, Daniel J., Kazuhiko Adachi, and William E. Royer. "The High Resolution Crystal Structure of Deoxyhemoglobin S." *Journal of Molecular Biology* 272.3 (1997): 398-407
- <sup>9</sup> Akbar, MGK, et al. "Inhibition Of Hemoglobin S Polymerization In Vitro By A Novel 15-Mer EF-Helix, Beta 73 Histidine-Containing Peptide." *Biochemistry* 45.27 (2006): 8358-8367
- <sup>10</sup> Ferrone, Frank A. "The Delay Time In Sickle Cell Disease After 40 Years: A Paradigm Assessed." *American Journal Of Hematology* (2015).
- <sup>11</sup> Green, NS, and S Barral. "Genetic Modifiers Of Hbf And Response To Hydroxyurea In Sickle Cell Disease." *Pediatric Blood & Cancer* 56.2 (2011): 177-181.
- <sup>12</sup> Abdulmalik, Osheiza, et al. "5-Hydroxymethyl-2-Furfural Modifies Intracellular Sickle Haemoglobin And Inhibits Sickling Of Red Blood Cells." *British Journal Of Haematology* 128.4 (2005): 552-561.
- <sup>13</sup> Conran, Nicola, and Fernando F. Costa. "Review: Hemoglobin Disorders And Endothelial Cell Interactions." *Clinical Biochemistry* 42.Hemoglobin Disorders (2009): 1824-1838.
- <sup>14</sup> Perumbeti, A, et al. "A Novel Human Gamma-Globin Gene Vector For Genetic Correction Of Sickle Cell Anemia In A Humanized Sickle Mouse Model: Critical Determinants For Successful Correction." *Blood* 114.6 (n.d.): 1174-1185.
- <sup>15</sup> Tantawy, Azza A. G. "The Scope Of Clinical Morbidity In Sickle Cell Trait." *Egyptian Journal Of Medical Human Genetics* 15.4 (2014): 319-326.
- <sup>16</sup> Hamzeh-Mivehroud, Maryam, et al. "Review: Phage Display As A Technology Delivering On The Promise Of Peptide Drug Discovery." *Drug Discovery Today* 18.(2013): 1144-1157.
- <sup>17</sup> Pande, J, MM Szewczyk, and AK Grover. "Phage Display: Concept, Innovations, Applications And Future." *Biotechnology Advances* 28.6 (2010): 849-858.
- <sup>18</sup> Ullman, Christopher G., Laura Frigotto, and R. Neil Cooley. "In Vitro Methods For Peptide Display And Their Applications." *Briefings In Functional Genomics* 10.3 (2011): 125-134
- <sup>19</sup> "Ph.D. Phage Display Libraries Instruction Manual". Version 1.2 Ipswich, MA: New England Biolabs Inc, February 2014.
- <sup>20</sup> Ladner, Robert C, et al. "Review: Phage Display-Derived Peptides As Therapeutic Alternatives To Antibodies." *Drug Discovery Today* 9.(2004): 525-529.

- <sup>21</sup> Babickova, J, et al. "In Vivo Phage Display - A Discovery Tool In Molecular Biomedicine." *Biotechnology Advances* 31.8 (2013): 1247-1259.
- <sup>22</sup> Uchiyama, F, et al. "Designing Scaffolds Of Peptides For Phage Display Libraries." *Journal Of Bioscience And Bioengineering* 99.5 (2005): 448-456.
- <sup>23</sup> Nixon, AE, DJ Sexton, and RC Ladner. "Drugs Derived From Phage Display From Candidate Identification To Clinical Practice." *Mabs* 6.1 (2014): 73-85.
- <sup>24</sup> Made, V, S Els-Heindl, and AG Beck-Sickinger. "Automated Solid-Phase Peptide Synthesis To Obtain Therapeutic Peptides." *Beilstein Journal Of Organic Chemistry* 10.(2014): 1197-1212.
- <sup>25</sup> Kaspar, Allan A., and Janice M. Reichert. "Review: Future Directions For Peptide Therapeutics Development." *Drug Discovery Today* 18.(2013): 807-817.
- <sup>26</sup> Fosgerau, Keld, and Torsten Hoffmann. "Review: Peptide Therapeutics: Current Status And Future Directions." *Drug Discovery Today* 20.(2015): 122-128.
- <sup>27</sup> Vlieghe, Patrick, et al. "Synthetic Therapeutic Peptides: Science And Market." *Drug Discovery Today* 15.1-2 (2010): 40-56.
- <sup>28</sup> Renukuntla, J, et al. "Approaches For Enhancing Oral Bioavailability Of Peptides And Proteins." *International Journal Of Pharmaceutics* 447.1-2 (2013): 75-93.
- <sup>29</sup> Uhlig, Thomas, et al. "The Emergence Of Peptides In The Pharmaceutical Business: From Exploration To Exploitation." *Eupa Open Proteomics* 4.(2014): 58-69.
- <sup>30</sup> Hecht, Sidney M. *Bioorganic Chemistry: Peptides and Proteins*. New York: Oxford UP, 1998.
- <sup>31</sup> Jha, Arvid Kumar1, et al. "A Short Review On Concept And Types Of Combinatorial Chemistry." *Asian Journal Of Research In Chemistry* 7.2 (2014): 243-248.
- <sup>32</sup> Mishra, A. K.1 et al. "Solid Phase Synthesis And Their Screening System - Review." *Asian Journal Of Research In Chemistry* 4.3 (2011): 362-369.
- <sup>33</sup> "Edman Degradation". *Genosphere Biotechnologies*. <<http://www.genosphere-biotech.com/Edman-Degradation.html>>.
- <sup>34</sup> Lam, KS, et al. "Application Of A Dual-Color Detection Scheme In The Screening Of A Random Combinatorial Peptide Library." *Journal Of Immunological Methods* 180.2 (1995.): 219-223.
- <sup>35</sup> Devlin, Thomas M. *Biochemistry: With Clinical Correlations*. Hoboken, NJ: Wiley, 2010.
- <sup>36</sup> Crowther, J. R. *The ELISA Guidebook*. Totowa, NJ: Humana, 2001.
- <sup>37</sup> Lam, KS, et al. "Applications Of One-Bead One-Compound Combinatorial Libraries And Chemical Microarrays In Signal Transduction Research." *Accounts Of Chemical Research* 36.6 (n.d.): 370-377.
- <sup>38</sup> Redman, JE, et al. "Discovery Of G-Quadruplex Stabilizing Ligands Through Direct ELISA Of A One-Bead-One-Compound Library." *Organic & Biomolecular Chemistry* 4.23 (2006): 4364-4369.
- <sup>39</sup> Lang, Annemarie, et al. "Development Of A Peptide ELISA For Discrimination Between Serological Responses To Equine Herpesvirus Type 1 And 4." *Journal Of Virological Methods* 193.(2013): 667-673.
- <sup>40</sup> Ferrone, F A, J Hofrichter, and W A Eaton. "Kinetics Of Sickle Hemoglobin Polymerization. II. A Double Nucleation Mechanism." *Journal Of Molecular Biology* 183.4 (1985): 611-631.
- <sup>41</sup> Galkin, Oleg, et al. "Two-Step Mechanism Of Homogeneous Nucleation Of Sickle Cell Hemoglobin Polymers." *Biophysical Journal* 93.(2007): 902-913.