

University of Nevada, Reno

Active Insulin Analogs That Do Not Require Refrigeration

A thesis submitted in partial fulfillment
of the requirements for the degree of

**BACHELORS OF SCIENCE IN BIOLOGY/MOLECULAR MICROBIOLOGY AND
IMMUNOLOGY**

by

REBA SURI

Yftah Tal-Gan, Ph.D., Thesis Advisor

May, 2016

**UNIVERSITY
OF NEVADA
RENO**

THE HONORS PROGRAM

We recommend that the thesis
prepared under our supervision by

REBA SURI

entitled

Active Insulin Analogs That Do Not Require Refrigeration

be accepted in partial fulfillment of the
requirements for the degree of

**BACHELORS OF SCIENCE IN BIOLOGY/MOLECULAR MICROBIOLOGY AND
IMMUNOLOGY**

Yftah Tal-Gan, Ph.D., Thesis Advisor

Tamara Valentine, Ph.D., Director, Honors Program

May, 2016

Abstract:

Due primarily to hydrophobic interactions between nonpolar amino acid residues, molecules of insulin tend to interact with each other and undergo fibrillation to form linear aggregates when partially unfolded. Formation of these insulin fibrils significantly decreases insulin stability and activity, and has complicated insulin's role in the therapy of diabetes. This study will investigate techniques to create bioactive insulin analogs that are not susceptible to fibrillation, a temperature sensitive process, and thus do not require refrigeration. Solid phase peptide synthesis techniques and the incorporation of backbone modified amino acids into peptide sequences will be used to design novel insulin analogs.

Acknowledgments:

I would like to thank my thesis mentor, Dr. Yftah Tal-Gan, for guiding me through this project and helping me with my research. Thank you for all of your patience and support. I would also like to thank the Honors Program and Dr. Valentine for guiding me throughout this process.

Table of Contents

Abstract.....	i
Acknowledgements	ii
Table of Contents	iii
List of Figures	iv
Introduction.....	1
Methodology	18
Results	24
Discussion/Significance of Results.....	33
Future Plans	37
References	38

Figure 1	The primary structure of human insulin	4
Figure 2	Monomeric, dimeric, and hexameric forms of insulin	4
Figure 3	Scheme showing the change in conformation that insulin undergoes when binding the insulin receptor	9
Figure 4	Mechanism showing the conversion of an isoacyl peptide to a natural peptide	11
Figure 5	Ester dipeptide formation	11
Figure 6	Synthetic route to insulin via ester dipeptide formation	12
Figure 7	Fmoc solid phase peptide synthesis scheme	14
Figure 8	Chemical modifications compatible with Fmoc chemistry	17
Figure 9	Insulin synthetic scheme portraying the three fragments of insulin that will be synthesized separately, and then combined via native chemical ligation	19
Figure 10	Scheme showing the loading of the Wang and Knorr resins to initiate peptide synthesis	20
Figure 11	Insulin Fragment 1 synthesis	25
Figure 12	HPLC Prep Chromatograms for Insulin Fragment 1 (Knorr)	27
Figure 13	HPLC Prep Chromatograms for Insulin Fragment 1 (Wang)	28
Figure 14	HPLC Analytical Chromatograms for Insulin Fragment 1 (Knorr)	29
Figure 15	HPLC Analytical Chromatograms for Insulin Fragment 1 (Wang)	30
Figure 16	MALDI-TOF MS Results for Insulin Fragment 1 (Knorr)	31
Figure 17	MALDI-TOF MS Results for Insulin Fragment 1 (Wang)	32
Figure 18	DKP (Diketopiperazine) formation in solid phase peptide synthesis	33

Introduction:

Diabetes is a metabolic disease characterized by high blood glucose levels. The human body metabolizes most of the food into glucose, which is used as the primary source of energy. Insulin is a 51-amino acid globular protein that is produced in the pancreas and has a significant role in the regulation of blood glucose levels in the body. Beta cells in the pancreas produce and secrete insulin in response to high blood glucose levels. As a result, insulin increases the rate of glucose uptake and utilization in cells, thus lowering the levels of glucose in the blood (Olefsky, 1976). In patients with diabetes, however, there is either a deficiency in insulin synthesis, or difficulties associated with insulin recognition by the cells, resulting in high levels of glucose accumulating in the blood. The increase in blood sugar levels can lead to severe complications, including cardiovascular disease, blindness, ketoacidosis, and kidney failure (“Basics About Diabetes,” 2014). According to the Centers for Disease Control and Prevention (2014), 29.1 million people across the country have diabetes and it is currently the seventh leading cause of death in the U.S.

The two most prevalent forms of diabetes are Type 1 diabetes, also known as insulin-dependent diabetes mellitus (IDDM), and Type 2 diabetes or non-insulin-dependent diabetes mellitus (NIDDM). Type 1 diabetes, which accounts for ~5-10% of cases in the U.S., can be further subdivided into Type 1A and Type 1B forms (“Basics About Diabetes,” 2014). In Type 1A, insulin is not produced due to the autoimmune destruction of beta cells in the pancreas. Since the pancreas is making little to no insulin, glucose cannot be taken up by cells and utilized as energy to fuel the body. Therefore, patients with Type 1 diabetes need daily insulin injections to allow uptake of glucose by cells. Type 1B is classified as non-autoimmune or idiopathic, in which there is no production of insulin, but the exact cause is unknown. In Type 2 diabetes, which accounts for

~90% of cases in the U.S., cells become insensitive or resistant to the insulin signal due to down-regulation of the insulin receptor, along with a reduction in insulin production (McCarthy & Menzel, 2001). Thus, external administration of insulin is the sole treatment of Type I diabetes and many cases of Type II diabetes.

Clinically, diabetes can appear as hyperglycemia (high blood glucose), hyperlipidemia (increased levels of lipids in the blood), ketonemia (high concentration of ketone bodies in the blood), ketoacidosis (serious condition resulting from the elevated levels of ketone bodies in the blood, making it more acidic), and azoturia (increased levels of nitrogenous compounds in the urine). Risk factors include obesity, hypertension, cardiovascular disease, old age, impaired fasting glucose, impaired glucose tolerance, or being part of a high risk ethnic group (“Basics About Diabetes,” 2014). In order to treat and control diabetes, it is essential for diabetics to maintain a healthy diet and to regularly engage in physical activity and exercise. In addition, Type 1 diabetes requires insulin administration, and Type 2 diabetes requires regular blood glucose testing and may require insulin or other medications to regulate glucose levels in the blood. Medications to regulate cholesterol and blood pressure are often needed as well (“Basics About Diabetes,” 2014).

There are four major types of therapeutic human insulin; rapid-acting, short-acting, intermediate-acting, and long-acting. Rapid-acting insulin reaches the bloodstream and becomes active within five minutes of injection. It reaches its peak after approximately one hour, but continues to work for about two to four hours. Short-acting insulin (also known as regular-acting insulin) becomes active about thirty minutes after injection. It reaches its peak about two to three hours after injection and continues to work for three to six hours. Intermediate-acting insulin does not reach the bloodstream until two to four hours after injection. It reaches its peak between four and twelve hours and remains in effect for about twelve to eighteen hours. Long-acting insulin

takes six to ten hours to become active and reach the bloodstream, and remains in effect for about twenty to twenty-four hours. To reduce the frequency of insulin injections, a rapid-acting insulin analog that is more stable and that can be actively maintained in the body for longer periods of time needs to be created.

It is evident that insulin plays an essential role in regulating blood glucose levels in the body (Olefsky, 1976). Insulin is used therapeutically to treat millions of people who suffer from diabetes worldwide. However, the insulin therapies that are currently available are not the most efficient. Additionally, some patients may have an allergic reaction to insulin additives that are used to keep insulin free from bacteria and to regulate its time of action. It is, therefore, fundamental to understand the structure-function relationship of the insulin molecule in order to develop an efficient, fast-acting insulin analog that is more stable and able to be maintained for longer periods of time in the body.

The structure of insulin is very important to its function. It is a globular peptide hormone composed of two chains—an A chain and a B chain (Figure 1). The A chain is composed of 21 amino acid residues and the B chain is made up of 30 amino acid residues (Liu et al., 2014). Three disulfide bridges are responsible for maintaining the overall tertiary structure of insulin; two bridges cross-link the A and B chain to each other, while a third bridge is formed within the A chain itself (Liu et al., 2014). Insulin can exist as a monomer, dimer, or hexamer coordinated by two Zn^{2+} ions. The monomeric form is biologically active because it undergoes conformational changes in which the C-terminus of the B chain unmask the active core, allowing monomeric insulin to bind to the insulin receptor. Monomeric insulin is also smaller and can diffuse more rapidly (Figure 2). However, the exposed active core makes monomeric insulin more susceptible to fibrillation, and thus less stable. The dimeric and hexameric forms of insulin, on the other hand,

balance between rapid-acting and long-acting insulin is essential. The long-lasting effect is needed to maintain a basal level of insulin at all times, while rapid-acting insulin is needed to create a boost of insulin that facilitates glucose uptake by cells after eating. When insulin is injected, it takes time for the hexameric form to break down into active monomers. To get fast-acting insulin, the hexameric form needs to disassemble more quickly. The aim of this project is to prepare active monomers that are stable so that the hexameric form is not needed. This approach may help establish the balance between long and fast-acting insulin, and could therefore contribute to methods for diabetic therapy.

When trying to administer insulin externally as a therapeutic, diabetic patients are challenged with the fact that insulin is not very stable and is susceptible to inactivation by both degradation and fibrillation (Sluzky et al., 1992). Exposure to heat disrupts insulin's globular structure, causing conformational changes that result in fibril formation. Insulin can form crystals that precipitate out of solution as fibrils, which can clog the injection pumps used to administer insulin. Thus, this susceptibility to fibrillation in the presence of heat is one of the main reasons that insulin has to be refrigerated before use. If not refrigerated, insulin can form linear aggregates, disrupting its function and stability. This fibril formation is currently a significant problem in developing countries because in many impoverished areas, electricity is not available and insulin cannot be refrigerated. Developing an insulin analog that is not susceptible to fibrillation, but is still active, would be beneficial and have an impact on people suffering from diabetes in these areas.

Several studies have been conducted to determine why, and under what conditions, insulin unfolds and changes conformation to form inactive fibrillar structures. Methods for insulin stabilization and interference with hydrophobic surfaces within the insulin molecule to prevent

this aggregation have been noted to be effective. However, the structure-function relationship of insulin is still not fully understood, preventing further breakthroughs on alternative methods for diabetic therapy (Sluzky et al., 1992). Furthermore, existing alternative methods for insulin synthesis result in extremely low yields of the active and functional peptide hormone (Zaykov et al., 2013). After analysis of previous research on this topic, the aim of this project is to build upon the successes of previous experiments and design a method of creating a synthetic insulin analog that is both active and stable in the presence of non-ideal conditions, such as heat and hydrophobic interfaces including air or polymeric material (Brange et al., 1997).

The tendency of insulin to aggregate and form fibrils poses several problems. One of the significant problems of insulin fibrillation and inactivation is that this process is irreversible, unless sizeable changes are made to the solvent environment, which is not feasible (Brange et al., 1997). Another potential problem is the propensity of insulin aggregates to induce an immune response. When the conformation of a protein is disrupted, its immunogenicity can change (Brange et al., 1997). The immunogenicity can, in turn, elicit anti-drug antibodies (ADA) and pose a threat to the safety of biotherapeutic drug development (Ratanji et al., 2014). It was found that several diabetic patients receiving subcutaneous insulin injections developed an allergy to insulin, leading researchers to believe that insulin fibril formation within the injection pumps may be the cause of this allergy (Brange et al., 1997). Insulin aggregation is, therefore, a problem that can have detrimental effects to diabetic therapy, as patients can develop an allergy to their treatment (Ratner et al., 1990). It is thus crucial to develop an analog that is less susceptible to form these aggregates in solution.

Controlled experiments have been conducted to induce insulin fibrillation to study the mechanism of this process. Models studying competitive inhibition and aggregation at

hydrophobic interfaces have been proposed, but they do not address all aspects of the mechanism (Belgi et al., 2011). For example, early models did not determine why the stability of insulin increases, as the concentration of insulin hexamers increases (Sluzky et al., 1992). A later experiment by Sluzky et al. addressed this issue by explaining the stabilization mechanism and by proposing a way to increase the stability of insulin in agitated aqueous solutions. This study used a model to show aggregation of Zn-insulin when exposed to solid hydrophobic interfaces (Sluzky et al., 1992). This model was used to predict the effect of insulin concentration, hydrophobic surfaces, and agitation rates, on insulin aggregation (Sluzky et al., 1992). The research found that insulin monomers are not very stable in agitated solution. Thus, increasing the concentration of these monomers causes aggregation to occur more quickly than when the solution contains larger molecules such as insulin hexamers. The Zn^{2+} ion induces hexamer formation and is thus useful to lower the rate of aggregate formation (Sluzky et al., 1992). The effects of Zn^{2+} were discovered through a Na-Insulin aggregation experiment that contained the Na^+ ion in place of Zn^{2+} ion, and thus contained more insulin monomers that aggregated more quickly than hexamers. In a different experiment testing the effect of pH on insulin aggregation, a similar discovery was made. In an acidic solution with low pH, molecules of insulin tended to be in monomeric form, and thus were more likely to associate with each other and ultimately form fibrils (Brange et al., 1997). Therefore, insulin forms fibrils more rapidly and insulin's function is disrupted, in acidic solutions (Brange et al., 1997).

Insulin plays a significant role in the human body, yet its structure-function relationships are not quite understood, due to its complex structure and thus synthesis (Mayer et al., 2007). From the few studies that were conducted by Brange et al. and Zaykov et al., a region of insulin was identified to control both the activity and the stability of insulin by switching between two protein

conformations, as shown in Figure 3 (Brange et al., 1997). NMR and X-ray diffraction analyses were used to identify this region, located at the *C*-terminus of the B chain, which was found to be very flexible and easily displaced (Brange et al., 1997). In order to prevent insulin fibrillation, this pliable region can be cross-linked to the *N*-terminus of the A chain to stabilize the structure and prevent the *C*-terminus of the B chain from being displaced. Unfortunately, the stabilization of insulin by cross-linking results in complete loss of insulin activity and functionality (Brange et al., 1997). A peptide bond linking the A chain to the B chain was shown to create an insulin crystal structure in which no activity was detected (Derewenda et al., 1991). This inactive crystalline molecule still retained its ability to fold and obtain two different conformations, and was found to be very similar in structure to the native insulin dimer, yet was inactive (Derewenda et al., 1991). X-ray analysis of this structure led to the conclusion that insulin is not active in crystal form, and insulin is not able to bind to the insulin receptor if the *N*-terminus of the A chain is cross-linked to the *C*-terminus of the B-chain (Derewenda et al., 1991). This cross-linking restrains the insulin molecule from adopting the bioactive conformation, which is required for insulin receptor binding.

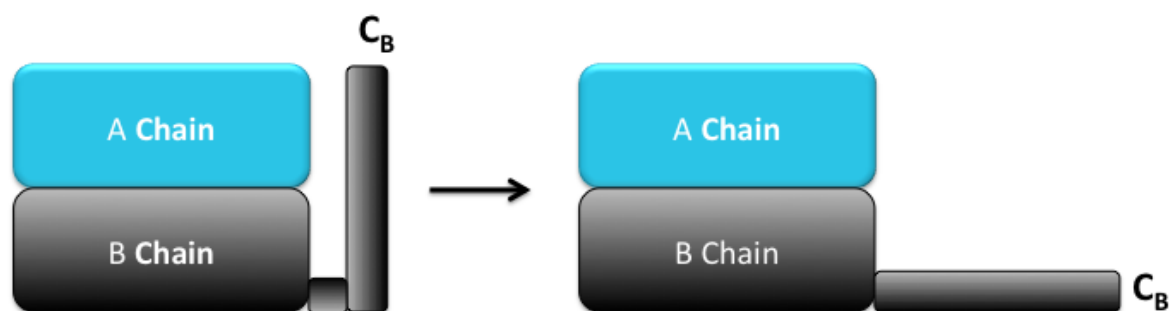


Figure 3. Scheme showing the change in conformation that insulin undergoes when binding the insulin receptor. The *C*-terminus of the B chain (C_B) protects the active core of the insulin molecule. In order for insulin to bind to its receptor, this region of the B chain is displaced to expose the active core. Figure adapted from Avital-Shmilovici et al., 2013.

The active site of insulin is derived from residues of the A and B chains, forming a hydrophobic core that is buried within the insulin molecule and protected from the environment by the *C*-terminus of the B chain. If the B chain is restrained and cannot move to expose this core, insulin will be inactive. By cross-linking the B chain to the A chain, this movement of the B chain is prevented, and the binding surface is not exposed to the insulin receptor (Leyer et al., 1995). On the other hand, removing the *C*-terminus of the B chain exposes the binding core and results in active insulin analogs that fibrillate very fast (Brange et al., 1997). Knowing the structure of insulin and the mechanism by which the *C*-terminus of the B chain controls insulin activity/stability is important in order to create an intermediate conformation that allows the insulin molecule to retain its activity. However, insulin's challenging structure makes it difficult to create analogs efficiently. The aim of this project is to target this region and design insulin analogs that adopt an intermediate conformation where insulin is both active and not susceptible to fibrillation.

Recent experiments were conducted to design a synthetic route to create insulin (Liu et al., 2014). Some of these strategies included using isoacyl peptides, a novel DesDi precursor, or an ester proinsulin surrogate, that can be chemically converted to human insulin. In one of the most recent experiments published in 2014, Liu et al. proposed a model in which they incorporated isoacyl dipeptide fragments into the A and B chains of the insulin molecule. This approach utilized an O-to-N shift at threonine and serine esters of isoacyl peptides (Figure 4). Both the A and B chains of insulin were prepared using an isoacyl substitution. The amino group on the isoacyl peptide was ionized, making each chain more hydrophilic and thus increasing its solubility in water. Once the A and B chains of insulin were synthesized separately, selective disulfide bond formation via orthogonal cysteine protection was used to combine the two chains and correctly fold insulin. Incorporation of the isoacyl peptide fragments was shown to refine the properties of

hydrophobic residues, allowing for higher yields of pure insulin products (Liu et al., 2014). In a similar approach, Zaykov et al. attempted to enhance folding efficiency by using a 49 amino acid insulin scaffold, two amino acids shorter than native insulin, called DesDi. This novel precursor was derived from miniproinsulin and tested for its ability to efficiently fold into the correct conformation and be converted to active insulin (Zaykov et al., 2013). However, unlike the isoacyl peptide analog designed by Liu et al., the DesDi precursor is a linear, single-chain molecule (Zaykov et al., 2013). This single chain folded structure can undergo enzymatic cleavage to form active two-chain insulin, and was shown to create insulin more efficiently and in higher yield than its two-chain alternative structures (Zaykov et al., 2013). Avital-Shmilovici et al. proposed a similar mechanism, using a different precursor. Instead of using DesDi, an ester insulin was used (Avital-Shmilovici et al., 2013). This group studied the crystal structure of insulin, and saw that in the folded stable conformation, the Thr in position 30 on the B chain is in close proximity to the Glu in position 4 on the A chain. The hydroxyl group on Thr can react with the carboxyl group from Glu to form an ester, which can help fold insulin into the correct stable conformation (Figure 5). Ester insulin is inactive, but it helps achieve the stable folded conformation of insulin in higher yields. Once correctly folded, ester insulin can undergo hydrolysis to form active human insulin. To achieve this structure, this lab group used a more robust technique by synthesizing insulin as three separate fragments, and then combining them via native chemical ligation as shown in Figure 6 (Avital-Shmilovici et al., 2013). These three experiments all designed synthetic insulin analogs using different approaches, and each focused on a different aspect or challenge associated with insulin synthesis.

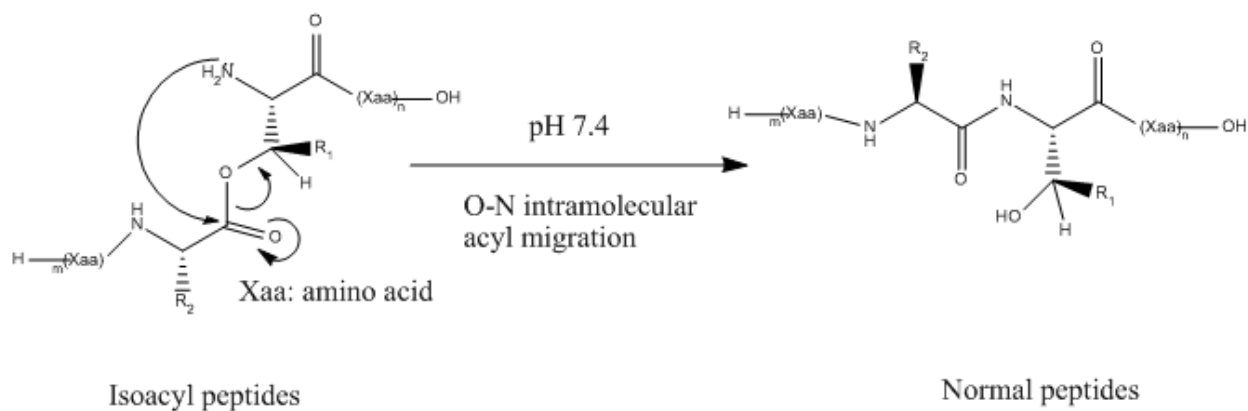


Figure 4. Mechanism showing the conversion of an isoacyl peptide to a natural peptide. The ionized amino group on the isoacyl peptide makes it more hydrophilic and thus increases its solubility in water. Figure adapted from Liu et al., 2014.

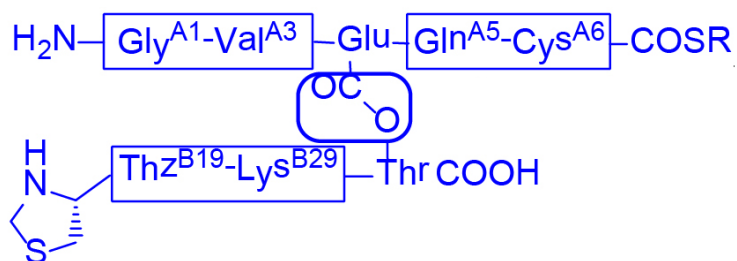


Figure 5. Ester dipeptide formation. The ester formed from the reaction between the carboxyl group of Glu in position 4 on the A chain and the hydroxyl group of Thr in position 30 on the B chain.

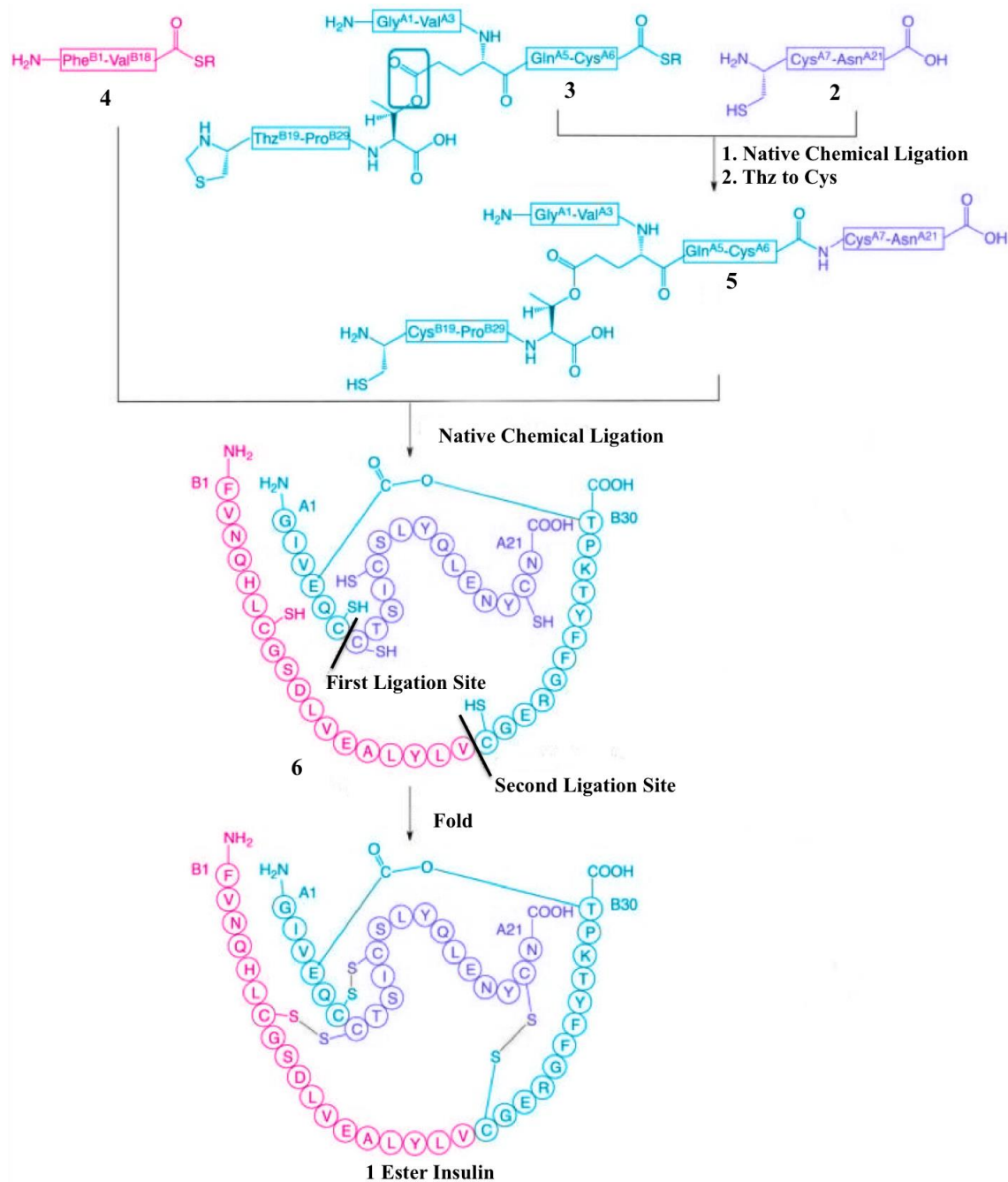


Figure 6. Synthetic route to insulin via ester dipeptide formation. Insulin synthetic scheme portraying the three fragments of insulin that were synthesized separately, and then combined via native chemical ligation. Fragment 1 is indicated in purple, fragment 2 in pink, and fragment 3 in turquoise. The bottom structure portraying inactive ester insulin can undergo ester hydrolysis to form active human insulin. Figure adapted from Avital-Shmilovici et al., 2013.

However, although these methods of insulin synthesis seem promising, they typically result in yields of less than 10% of the functional insulin protein (Zaykov et al., 2013). The materials and

equipment needed to run these experiments are very costly, and obtaining such a low yield is wasteful. Furthermore, the more costly the experiment, the more expensive the product once it hits the market. This would almost be counterintuitive, as these insulin analogs should be made available to underserved areas where electricity is often not available to store native insulin, but the high costs would prevent them from affording this new product. The focus of this project is to create an active and stable insulin analog, in high yield, that is not susceptible to fibrillation, and does not require refrigeration. Creation of this analog would significantly improve methods for diabetic therapy and insulin administration, while providing a more stable structure that is available to underserved and impoverished populations.

To move towards this goal, a native insulin analog will be made first to verify that the synthetic route developed results in insulin that is as active as its natural/commercial counterpart. Insulin analogs will then be made synthetically using solid phase peptide synthesis to incorporate different non-natural modifications that can rigidify the C-terminus of the B chain that protects the susceptible region of insulin. Solid phase peptide synthesis is a technique that utilizes a solid support, polystyrene beads in this case, to initiate peptide synthesis. To insure the proper peptide sequence is achieved, protecting groups, such as Fmoc and Boc, are used to protect the amino group of the amino acid. Amino acid side chains are protected with orthogonal protecting groups to prevent side reactions of functional groups during synthesis. A coupling reagent is added to form a covalent bond between the first amino acid of the peptide, and the resin. A labile protecting group is then removed from the amino group of this amino acid, to allow it to bind to the next amino acid of the peptide. This process is repeated until all amino acids needed to synthesize the peptide are added. The newly made peptide chain is then cleaved from the resin (Figure 7).

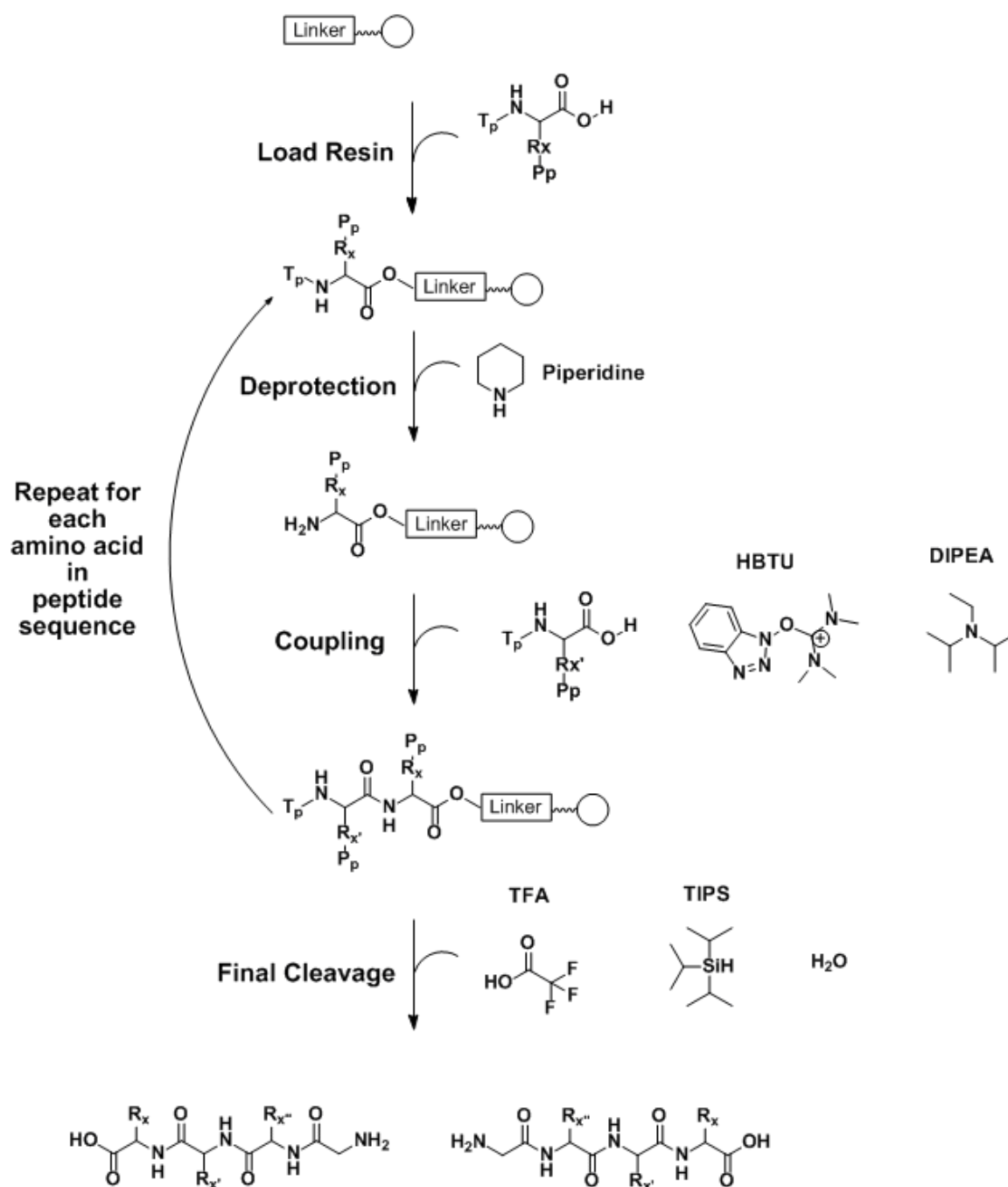


Figure 7. Fmoc Solid Phase Peptide Synthesis Scheme. SPPS is a technique that utilizes a solid support, polystyrene beads in this case, to initiate peptide synthesis. To insure the proper peptide sequence was achieved, an Fmoc protecting group was used to protect the amino group of each amino acid. Fmoc was then removed from the amino acid to allow it to bind to the next amino acid of the peptide. This process was repeated until all amino acids needed to synthesize the peptide were added. The newly made peptide chain was then cleaved from the resin. Amino acid side chains were protected with orthogonal protecting groups to prevent side reactions of functional groups during synthesis, and were only removed during final cleavage

Many of the previous approaches to create insulin analogs do not address insulin's structure-activity relationships, and as previously mentioned, result in very low yields. Additionally, most of these approaches are not robust, and require synthesis of the entire insulin sequence every time, which is time consuming and wasteful. The approach used by Avital-Shmilovici et al., on the other hand, is a more robust method that allows synthesis and modification of individual peptide fragments within the insulin molecule, rather than requiring the synthesis of the entire molecule each time. This technique is advantageous when trying to synthesize and test several different analogs, because synthesizing the peptide fragments individually allows modification of a single fragment that can then be combined with fragments that have already been made. The problem with this approach, however, is that it was designed for chemistry that does not permit the modifications that are desirable for this research project.

For this project, the goal is to maintain the activity of native insulin by keeping the same side chain residues, but modifying the backbone of the peptide to create a more stable conformation. Therefore, a novel approach is used by modifying the chemistry used by Avital-Shmilovici et al. Instead of using Boc chemistry, Fmoc chemistry synthesis will be used, which is more compatible with the incorporation of additional modifications. For one, Boc chemistry uses harsh conditions and chemicals, such as HF (hydrogen fluoride), which are not ideal. Additionally, Boc chemistry is stable under acidic conditions, but is not compatible with basic conditions. Boc chemistry is, therefore, not compatible with some of the backbone modifications needed for this project because these modifications require basic conditions during synthesis. Fmoc chemistry, on the other hand, is designed for basic conditions and, therefore, allows for the incorporation of a broad spectrum of modifications. One such modification is the formation of peptoids, also known as poly-N-substituted glycines. In a naturally occurring amino acid such as alanine, the side chain

is attached to the α -carbon. In peptoids, on the other hand, the side chain is instead attached to the amino group of the peptide backbone (Figure 8). Peptoids can be used to modify the conformation of insulin, and strong basic conditions are required for their formation. Another modification is N-methylation, in which the hydrogen atom from the amino group of the original amino acid is replaced with a more bulky methyl group, forming a tertiary amide when incorporated into a peptide sequence (Figure 8). Coupling between two naturally occurring amino acids results in the formation of a secondary amide that can act as a hydrogen bond donor. However, due to the absence of the hydrogen atom normally present in a secondary amide, the tertiary amide can no longer act as a hydrogen bond donor, thus affecting the peptide's structure. The addition of the bulky methyl group also causes steric hindrance, which can restrict the conformation of the insulin molecule to some extent, and will hopefully increase its stability. Both peptoids and N-methyl modifications can affect the structure and rigidity of the peptide. Therefore, the incorporation of backbone modified amino acids to form peptide mimics may allow for the synthesis of insulin analogs with rigidified or more flexible structure. These analogs will then be used to study the structure-function of insulin and identify stable insulin analogs that do not require refrigeration. Increasing the stability of insulin in solution would decrease risks associated with high blood glucose levels and improve the quality of life for diabetics.

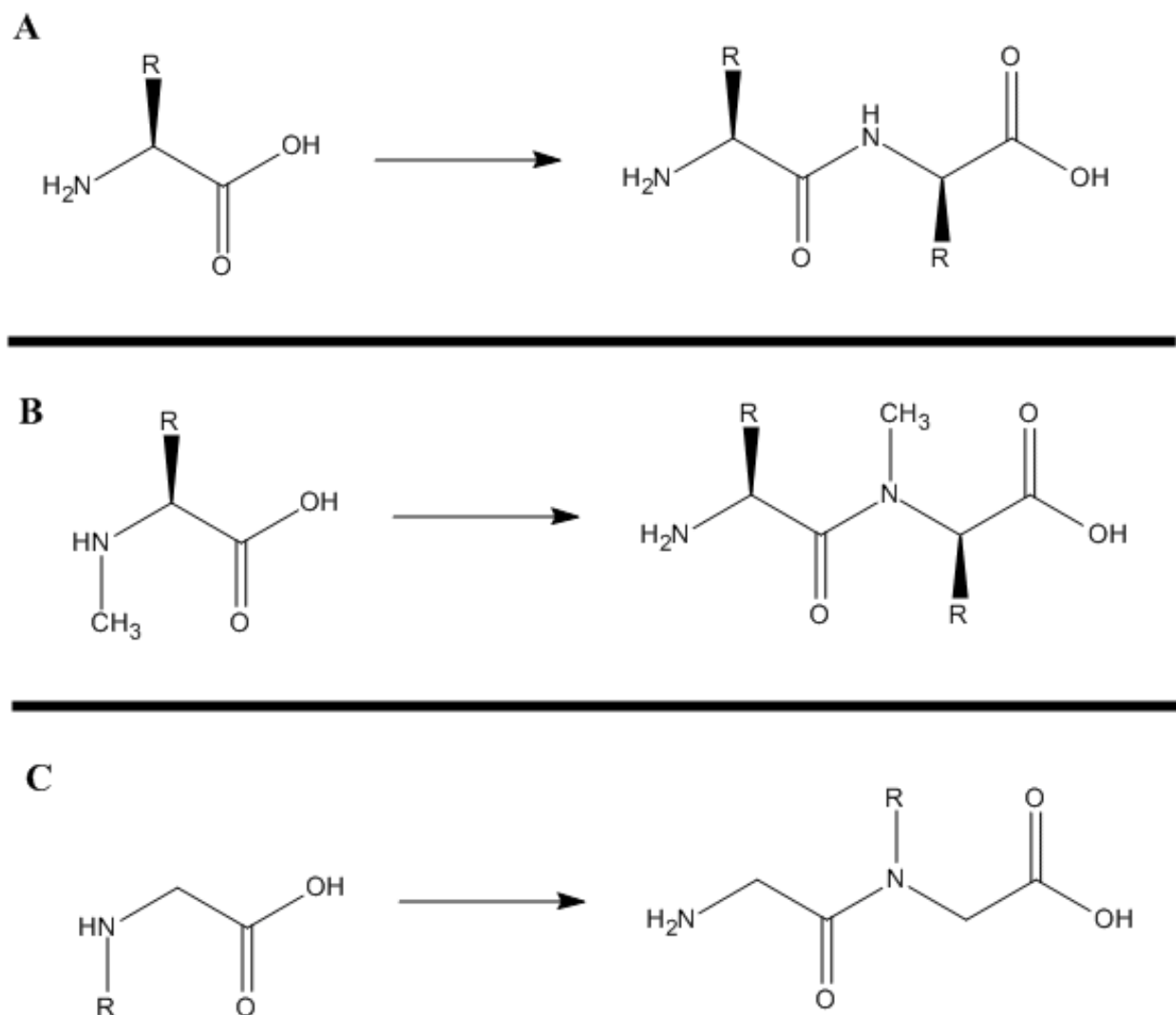


Figure 8. Chemical modifications compatible with Fmoc chemistry. (A) Naturally occurring amino acids are linked via a peptide bond to form a secondary amide. The hydrogen atom on the secondary amide allows it to act as a hydrogen bond donor. (B) N-methyl modified amino acids. Methylation at the N-terminus of the amino acid causes the formation of a tertiary amide between two amino acid residues, in contrast to the secondary amide formed between two naturally occurring amino acids. Due to the absence of the hydrogen atom normally present in a secondary amide, the tertiary amide can no longer act as a hydrogen bond donor. Additionally, the addition of the methyl group adds bulkiness which restricts the conformation of the amide bond due to steric hindrance. (C) Peptoids (poly-N-substituted glycines). In peptoids, the side chain is attached to the amino group of the peptide backbone instead of to the α -carbon as it is in naturally occurring amino acids.

Methodology:

Solid Phase Peptide Synthesis

Fmoc-based Solid Phase Peptide Synthesis (SPPS) was used to synthesize Insulin Fragment 1 (Figure 9). SPPS consists of four primary steps: Loading of the resin, synthesis involving a series of deprotection and coupling reactions, and final cleavage (Figure 7). Insulin fragment 1 was synthesized using two different resins, Knorr (Rink Amide-MBHA) and pre-loaded Wang resin, to determine which method would result in higher yields of the peptide. The loading of the Wang resin is a difficult process and therefore, in most cases, it already comes pre-loaded. In this case, the Wang Resin was purchased pre-loaded with asparagine (Asn), the first amino acid from the C-terminus of the Insulin Fragment 1 sequence, via an ester bond. The ester bridge is formed between the linker of the resin and the carboxyl group at the C-terminus of Asn (Figure 10). In order to load the Knorr Resin with Asn via an alternate route, Fmoc-L-Asp-OtBu was coupled to the Knorr Resin to form a stronger amide bond, which in turn formed Asn. Instead of attaching Asp to the resin via the carboxyl group on its C-terminus, in order to convert Asp to Asn, Asp was attached to the resin via the carboxyl group on its side chain (Figure 10). The side chain was connected to the linker of the Knorr resin via an amide bridge. The carboxyl group at the C-terminus of Asp was protected with Boc to prevent it from attaching to the resin. After peptide elongation, final cleavage yielded a peptide with Asn at the C-terminus. The peptide was initially synthesized manually and was later synthesized with the AAPPTec Focus-XC peptide synthesizer.

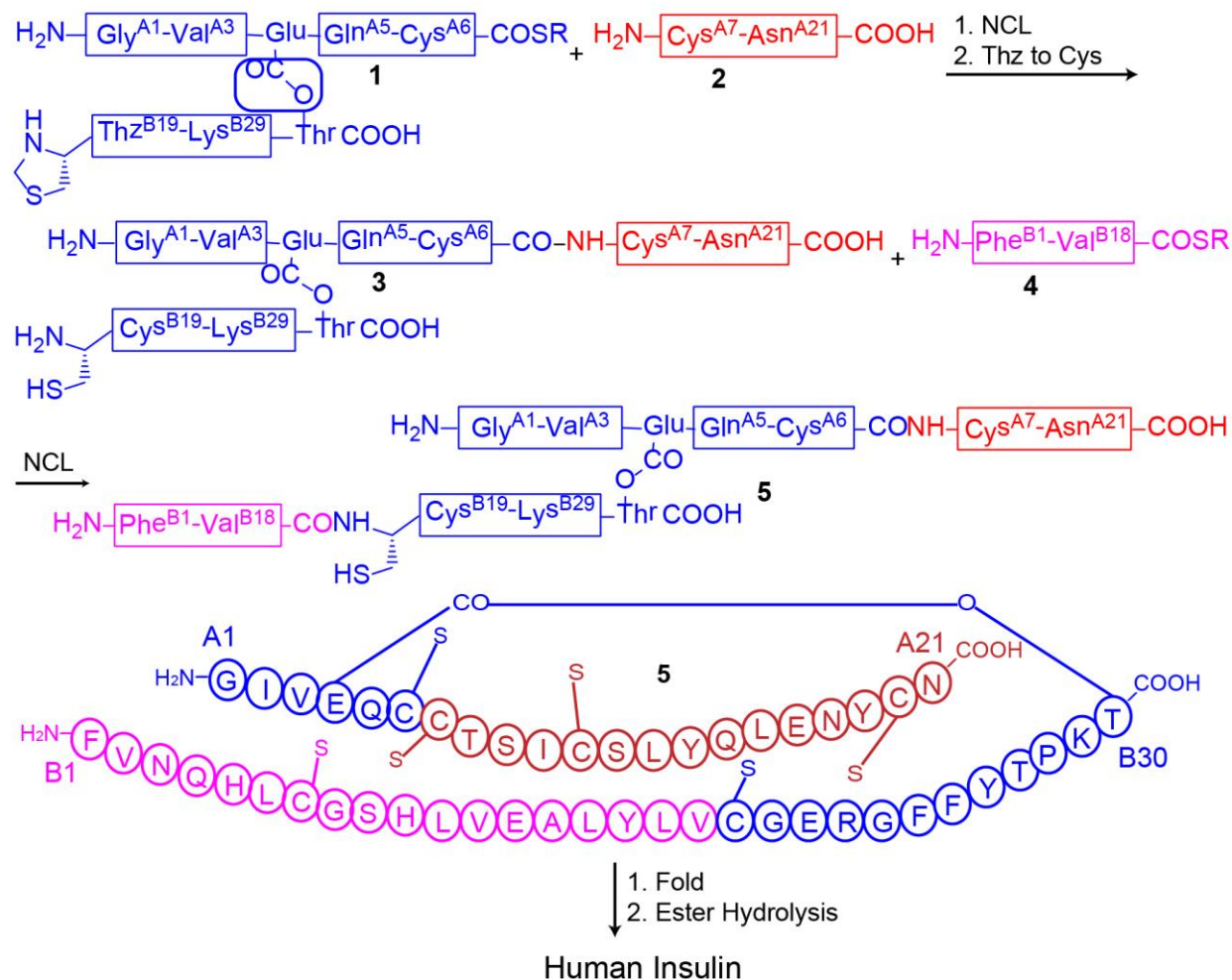


Figure 9. Insulin synthetic scheme portraying the three fragments of insulin that will be synthesized separately, and then combined via native chemical ligation. Fragment 1 is indicated in red, fragment 2 in pink, and fragment 3 in blue.

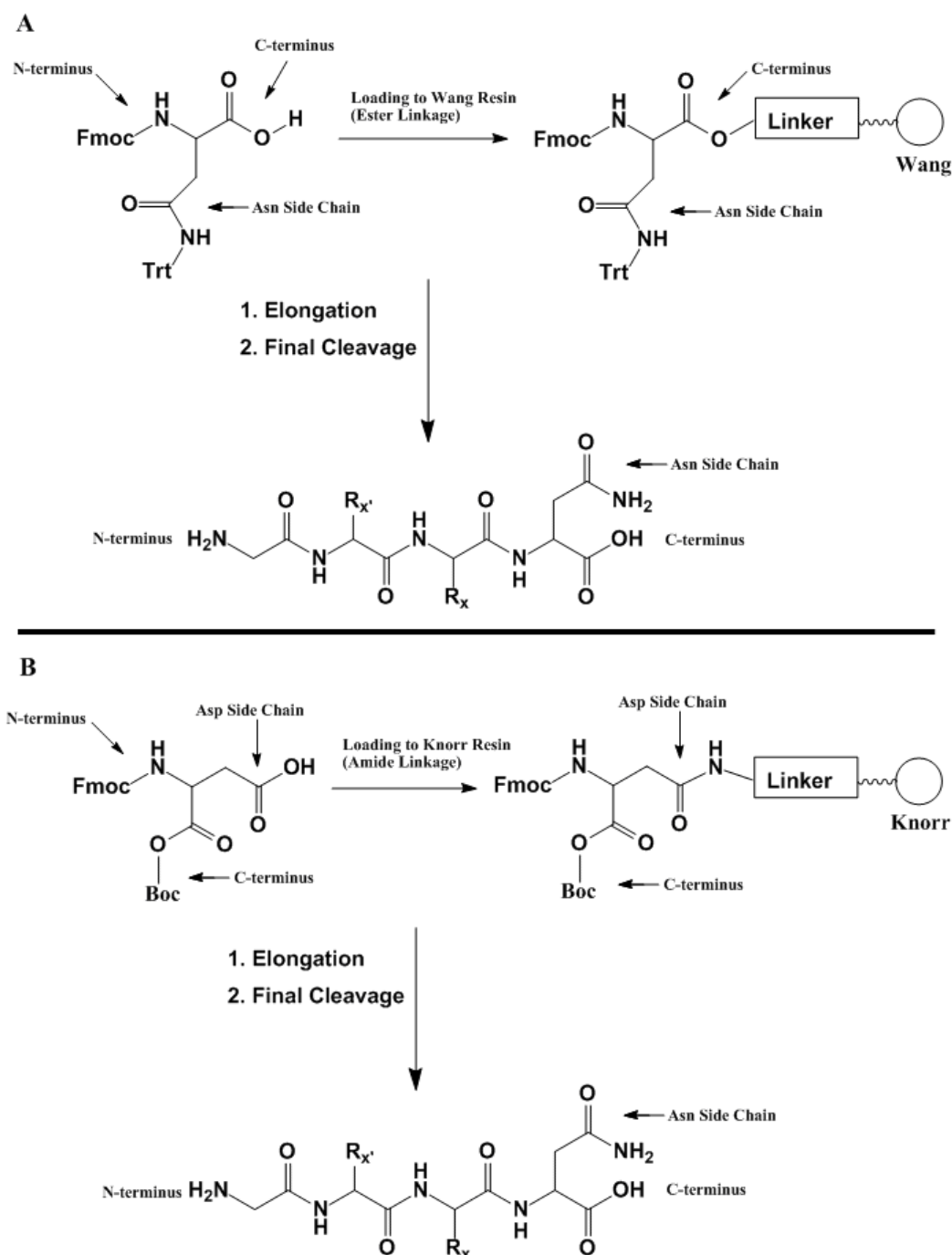


Figure 10. Scheme showing the loading of the Wang and Knorr resins to initiate peptide synthesis. (A) Mechanism showing the loading of Asn onto the Wang resin. An ester bridge is formed between the linker of the resin and the carboxyl group of Asn. The side chain of Asn is protected with a Trt group to prevent it from attaching to the resin. (B) Mechanism showing the loading of Asp onto the Knorr resin. Asp was attached to the linker of the resin via the carboxyl group on its side chain forming an amide bridge. The carboxyl group at the C-terminus of Asp was protected with Boc to prevent it from attaching to the resin. After peptide elongation, final cleavage yielded a peptide with Asn at the C-terminus.

Synthesis. In two reaction vessels, 0.1 grams of each resin was measured out. The Wang (Fmoc-Asn (Trt)) resin has a loading of 0.70 mmol/g and the loading of the Knorr resin is 0.89 mmol/g. The following procedure was then followed for each resin. The resin was swelled with 2-3 mL of DCM (Dichloromethane) for 30 min and then drained. DMF (3 x 1 min) was then used to wash the resin. The resin was then deprotected (Fmoc removal) with 2 mL of 20% Piperidine in DMF solution on the shaker (2 x 7.5 min). After Fmoc removal, the resin was washed with DMF (3 x 1 min). For each coupling reaction, 2 equivalents (with respect to the loading of the resin) of Fmoc-protected amino acid and 2 equivalents of a coupling reagent, HBTU (*N,N,N',N'*-Tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate), were dissolved in 3-4 mL of DMF. DIPEA (diisopropylethylamine, 2 equiv) was added and the mixture was vortexed for 1 min to allow for pre-activation. The solution was then added to the resin and set to shake for 1 hour to allow for the amino acid to couple to the resin. The solvent was flushed and the resin was washed with DMF (3 x 1 min) after coupling. Before the next amino acid could be coupled to the resin, the Fmoc protecting group was removed using 20% Piperidine in DMF as mentioned above. After deprotection, the resin was washed with DMF once again (3 x 1 min) before the next amino acid could be added. This process was repeated until synthesis of the peptide was complete. Fmoc was removed from the final amino acid using 2 mL of 20% Piperidine in DMF solution on the shaker (2 x 7.5 min). After Fmoc removal, the resin was washed with DMF (3 x 1 min) and DCM (2 x 1 min). Diethyl ether was then used to dry the resin (1 x 3 min) and it was transferred to a 15 mL falcon tube.

Cleavage. To the dried resin, a solution of 95% TFA (trifluoroacetic acid), 2.5% 18 Ω M H₂O, and 2.5% TIPS (triisopropylsilane), was added and the mixture was left to shake for 3 hours at room temperature to cleave the peptide from the resin. The resulting cleavage solution was

separated from the resin by filtration and the resin was rinsed with TFA to ensure that all of the peptide was collected in solution. To precipitate the peptide from the filtrate, 45 mL of a cooled solution of 1:1 ether: hexane was added and the mixture was set to cool in -20 °C freezer for 20 min to allow precipitation to occur. The mixture was then centrifuged at 3000 RPM for 5 min to allow for the precipitate to sediment. The supernatant was removed and the remaining pellet was dissolved in 1:1 18 Ω M H₂O: ACN solution. The solution was frozen in a dry ice/acetone bath and lyophilized.

RP-HPLC Separation and Purification

After lyophilization, the resulting crude peptide was purified using RP-HPLC (Reversed-Phase High Performance Liquid Chromatography). Mobile Phase A was a solution composed of 18 Ω M H₂O + 0.1% TFA and Mobile Phase B was a solution composed of ACN + 0.1% TFA. The peptide was dissolved in 3:1 18 Ω M H₂O: ACN solution. The solution was filtered and 3-5 mL was injected for purification at a time. RP-HPLC was performed using a Shimadzu system equipped with a CBM-20A communications bus module, two LC-20AT pumps, an SIL-20A auto sampler, an SPD-20A UV/VIS detector, a CTO-20A column oven, and an FRC-10A fraction collector.

Preparative Runs. HPLC prep runs were used to purify the crude peptide by separating it into different components based on structure and polarity using a linear gradient (first prep 5% B \rightarrow 45% B over 40 min and second prep 20% B \rightarrow 30% B over 45 min). Preparative work was done using a semi-preparative Phenomenex Kinetex C18 column (5 μ m, 10 mm \times 250 mm, 110 Å) and a semi-preparative Phenomenex Gemini C18 column (10 μ m, 10 mm \times 250 mm, 110 Å), with the flow rate set to 5 mL/min. After an initial prep run was done, a second or sometimes third

prep run was used to further purify the fractions containing the desired peptide until they reached a purity of >90% (as determined by analytical runs).

Analytical Runs. HPLC analytical runs were used to determine the purity of the fractions containing the desired peptide that were collected during the preparative work. The purity was measured using a linear gradient (5% B → 95% B over 27 min). Analytical work was done using an analytical Phenomenex Kinetex C18 column (5 μm , 4.6 mm \times 250 mm, 110 \AA) and a flow rate set to 1 mL/min.

Mass Detection

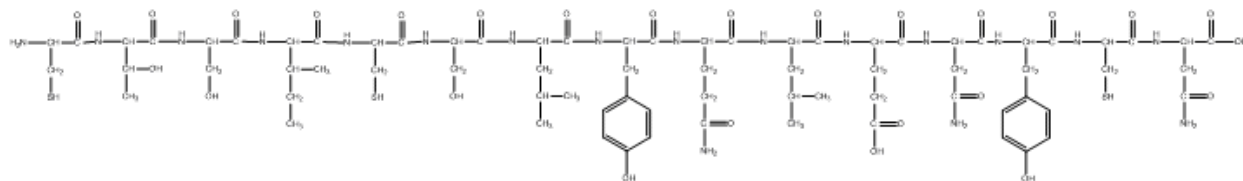
After purification, to determine whether each fraction collected contained the desired peptide, the MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization Time-Of-Flight Mass Spectrometry) was used to detect the peptide's mass. A matrix solution (0.7 μL) made of alpha-Cyano-4-hydroxycinnamic acid dissolved in 1:1 18 ΩM H_2O : ACN solution, and 0.7 μL of each sample were plated on a target and analyzed. Peptide masses were detected based on ionization with either a proton H^+ (+1 charge), a sodium ion Na^+ (+23 charge), or a potassium ion K^+ (+39 charge). MALDI-TOF MS data were obtained on a Bruker Microflex spectrometer equipped with a 60 Hz nitrogen laser and a reflectron. In positive ion mode, the acceleration voltage on Ion Source 1 was 19.01 kV.

Results:

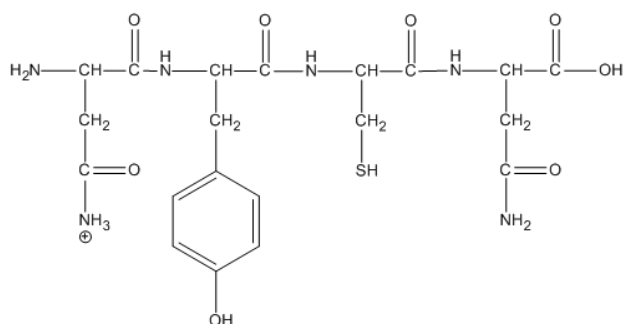
Insulin fragment 1 was synthesized using two different resins, Knorr (Rink Amide-MBHA) and pre-loaded Wang resin, to determine which method would result in higher yields of the desired peptide. The Wang resin was initially used to synthesize insulin fragment 1. A solution of 2% DBU/2% Piperidine in DMF was used as the base for deprotection. After coupling approximately four to eight amino acids, a small cleavage was performed to determine whether or not the synthesis was effective. After small cleavage, the sample was tested on the MALDI-TOF MS in order to determine whether it contained the desired mass. A small cleavage was performed after coupling six amino acids, and was repeated after eight amino acids were coupled to the Wang resin, and in both cases the mass was not detected on the MALDI.

It was hypothesized that the ester bond between the Asn and the Wang resin was too labile. Therefore, the Knorr resin was then used to synthesize insulin fragment 1. A solution of 2% DBU/2% Piperidine in DMF was again used as the base for deprotection. However, after performing several small cleavages to test whether the synthesis was effective, the mass was still not detected on the MALDI. In an attempt to troubleshoot, the base used for deprotection was changed from a solution of 2% DBU/2% Piperidine in DMF to 20% Piperidine. After coupling four amino acids (Asn-Tyr-Cys-Asn) to the Knorr resin, Fmoc was removed and a small cleavage was performed (Figure 11). This four amino acid sequence had a calculated mass of 513 m/z, and the approximate mass plus Na⁺ was detected on the MALDI (Figure 11). The peptide was resynthesized with the pre-loaded Wang resin using 20% Piperidine as the base for deprotection, and the mass was observed. Therefore, changing the base used for deprotection from a solution of 2% DBU/2% Piperidine in DMF to 20% Piperidine allowed synthesis to be optimized.

A. Insulin Fragment 1 Sequence



B. Knorr Resin Small Cleavage: Asn-Tyr-Cys-Asn (513 m/z)



C. Knorr Positive Ion Mode: 533 m/z (+23)

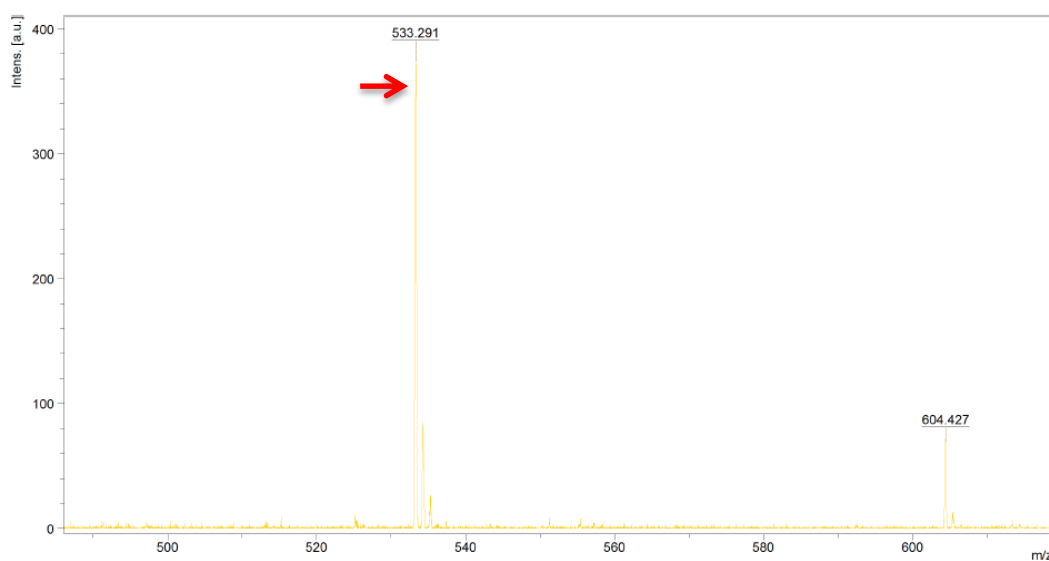
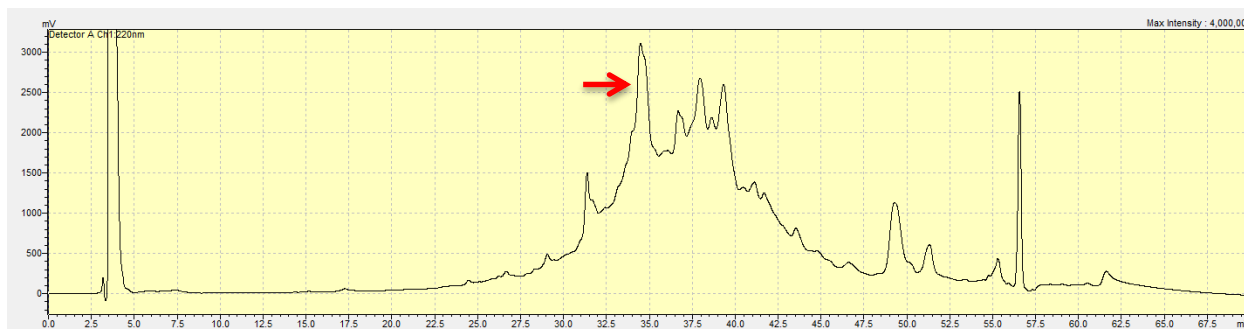
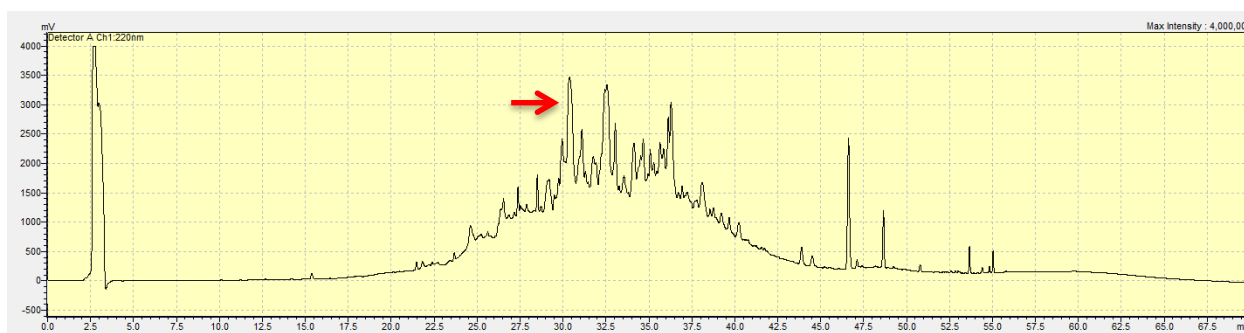


Figure 11. Insulin Fragment 1 synthesis. (A) Insulin Fragment 1 sequence. (B) A small cleavage was performed after four amino acids, Asn-Tyr-Cys-Asn, with a mass of 513 m/z, were coupled to the Knorr resin. (C) In positive ion mode, the mass was detected with a +20 charge indicating the potential ionization with Na⁺.

After optimizing the synthesis of Insulin Fragment 1, Reversed-Phase High Performance Liquid Chromatography (RP-HPLC) was used to purify the peptide. Initially, it was difficult to dissolve the peptide in 3:1 18 Ω M H₂O: ACN solution when preparing the sample for purification. To facilitate dissolution, different concentrations of 18 Ω M H₂O: ACN, sonication, and addition of a small amount of base (NaOH), were experimented with. A 10 μ m column was used to separate the crude peptide during the first prep run. Despite several attempts to increase the yield and purity, in both methods the yields were still low and it remained difficult to purify the peptide. Therefore a new preparative column with higher resolution (5 μ m) was purchased. The 5 μ m column allowed for better separation of the peptide than the 10 μ m column, and although still low, the pure peptide was obtained in higher yield when the 5 μ m column was used for separation. The percent yield of the peptide with >90% purity that was synthesized from the Knorr resin was 2.86%, and the yield from that synthesized using the Wang resin was 1.45%.

Figures 12 and 13 show the HPLC preparative work done for the crude and semi-pure samples. The chromatograms for the crude peptide, synthesized from both the Knorr and Wang resins, indicate better separation when the 5 μ m column was used in comparison to when the 10 μ m column was used. Figures 14 and 15 show the analytical work done on the fractions collected during preparative work that contained the desired peptide with a purity of >90%. Figures 16 and 17 show the MALDI-TOF MS results of the desired peptide synthesized from both the Knorr and Wang resins. The desired peptide, Insulin Fragment 1, has a mass of 1754 m/z which was detected in negative ion mode. In positive ion mode, the mass was detected with a +23 charge indicating ionization with Na⁺.

A. Knorr Crude First Prep (10 μm Column)B. Knorr Crude First Prep (5 μm Column)

C. Knorr Second Prep

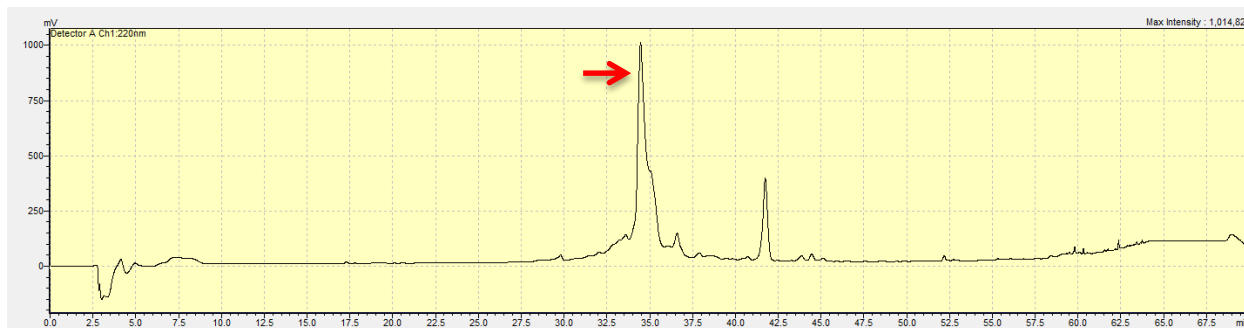
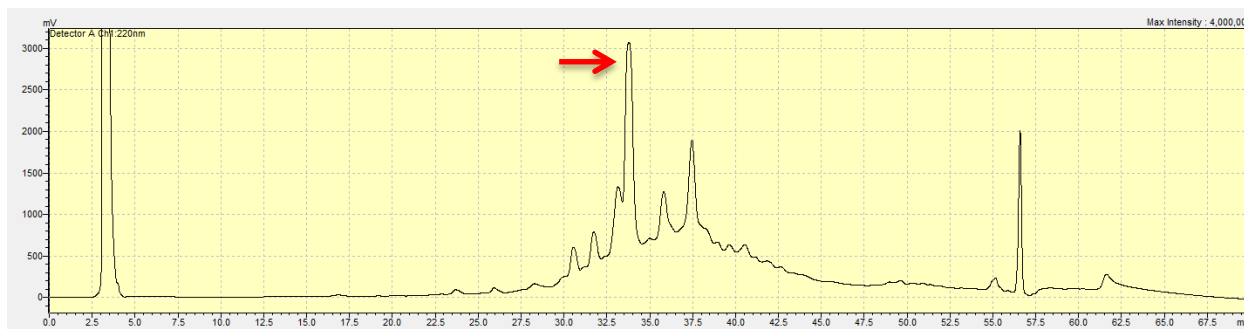
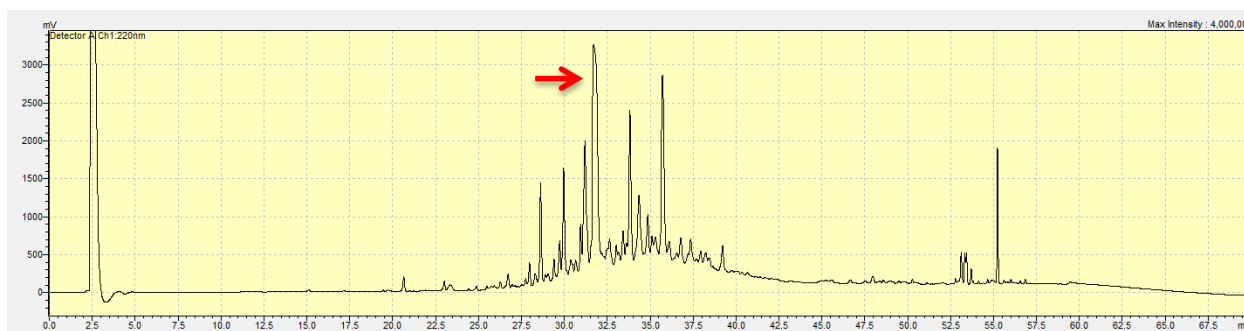


Figure 12. HPLC Prep Chromatograms for Insulin Fragment 1 (Knorr). Red arrows in the chromatograms indicate the fractions that contained the desired peptide. The crude peptide samples purified using the 10 μm (A) and 5 μm (B) columns were synthesized from the same batch. The chromatograms for the crude peptide show more peaks (higher resolution) and indicate better separation when the 5 μm column was used, in comparison to when the 10 μm column was used. (C) Insulin fragment 1 after a second prep run.

A. Wang Crude First Prep (10 μm Column)B. Wang Crude First Prep (5 μm Column)

C. Wang Second Prep

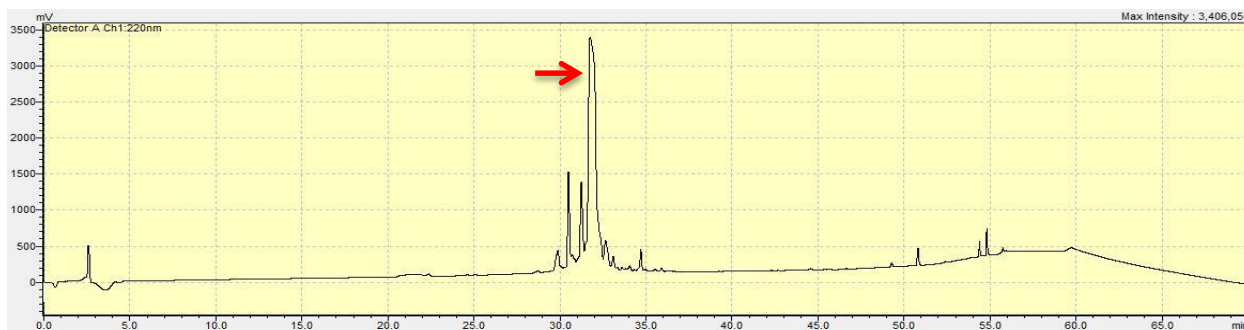
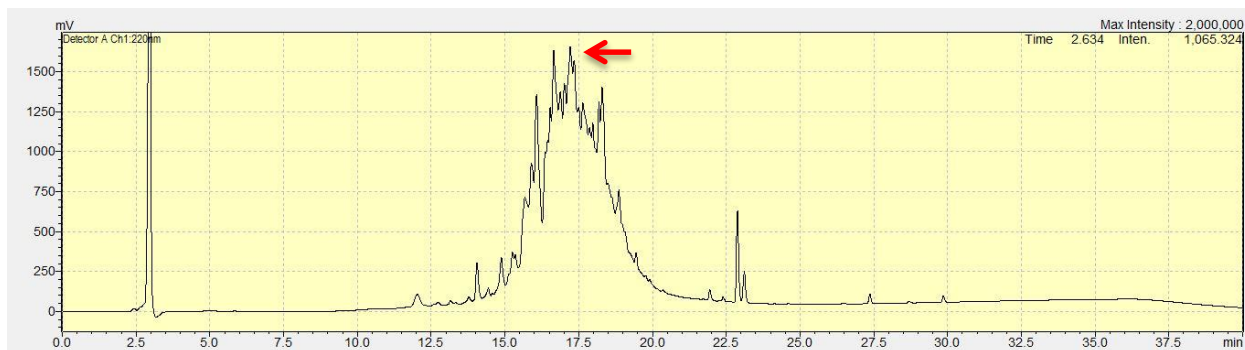


Figure 13. HPLC Prep Chromatograms for Insulin Fragment 1 (Wang). Red arrows in the chromatograms indicate the fractions that contained the desired peptide. The crude peptide samples purified using the 10 μm (A) and 5 μm (B) columns were synthesized from the same batch. The chromatograms for the crude peptide show more peaks (higher resolution) and indicate better separation when the 5 μm column was used, in comparison to when the 10 μm column was used. (C) Insulin fragment 1 after a second prep run.

Knorr Crude Analytical



Knorr 92% Pure Analytical

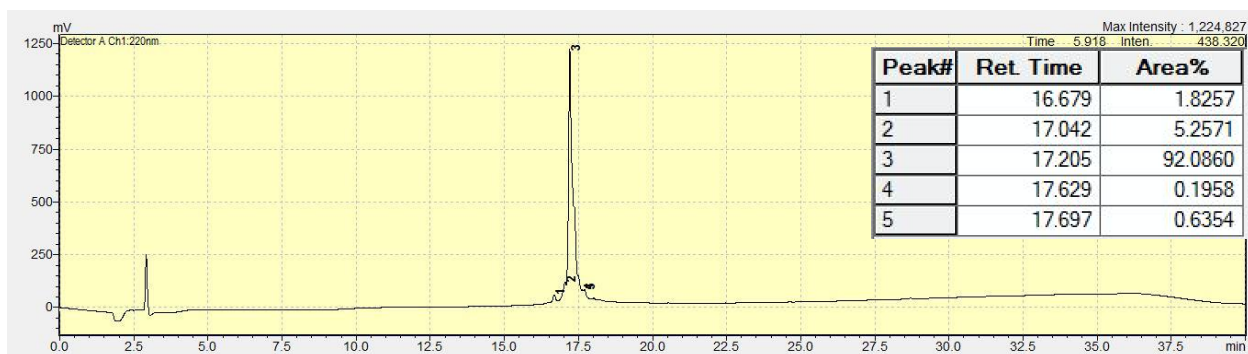
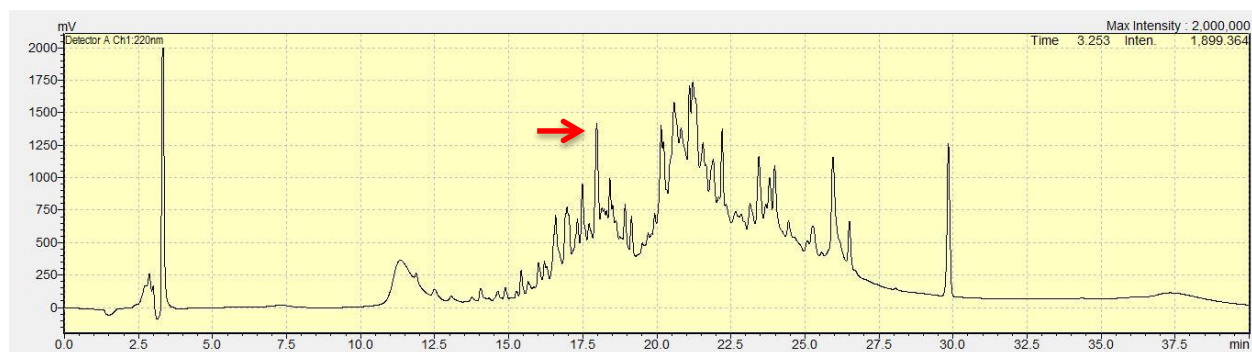


Figure 14. HPLC Analytical Chromatograms for Insulin Fragment 1 (Knorr). Tables on top right corner of each chromatogram indicate the purity of each peak shown.

Wang Crude Analytical



Wang 96% Pure Analytical

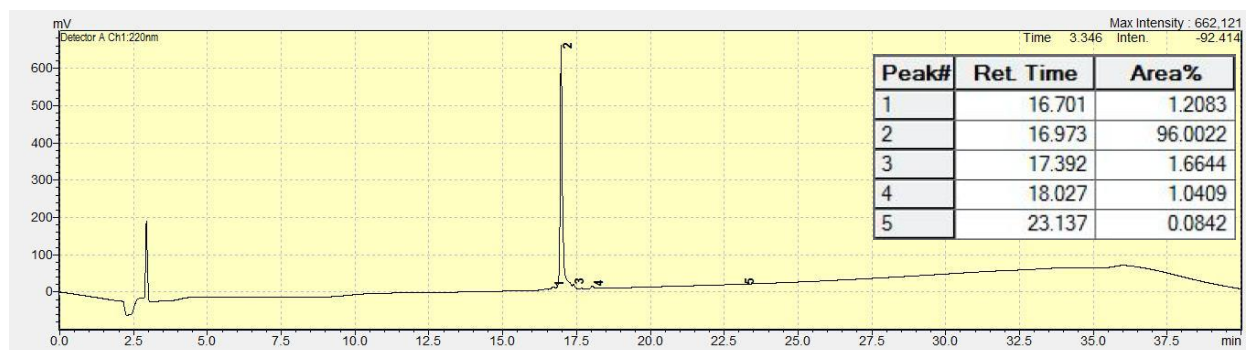
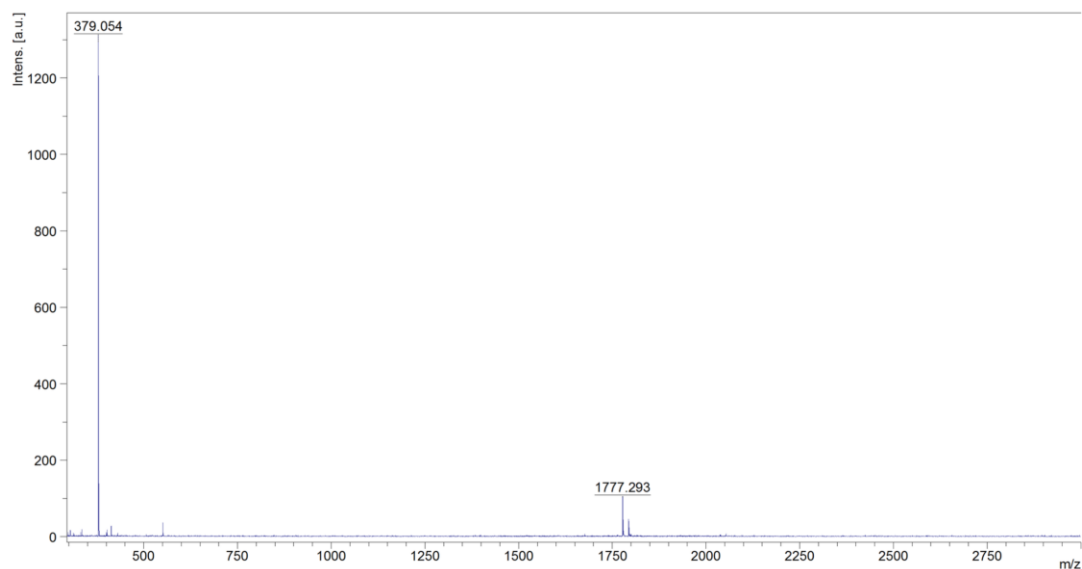


Figure 15. HPLC Analytical Chromatograms for Insulin Fragment 1 (Wang). Tables on top right corner of each chromatogram indicate the purity of each peak shown.

Knorr Positive Ion Mode: 1777 m/z (+23)



Knorr Negative Reflection: 1754 m/z

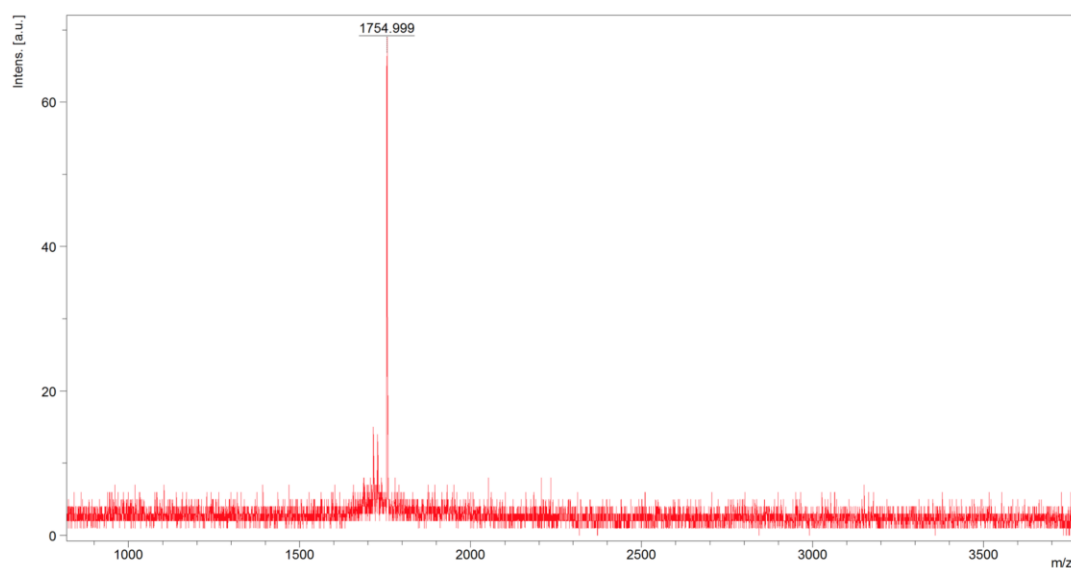
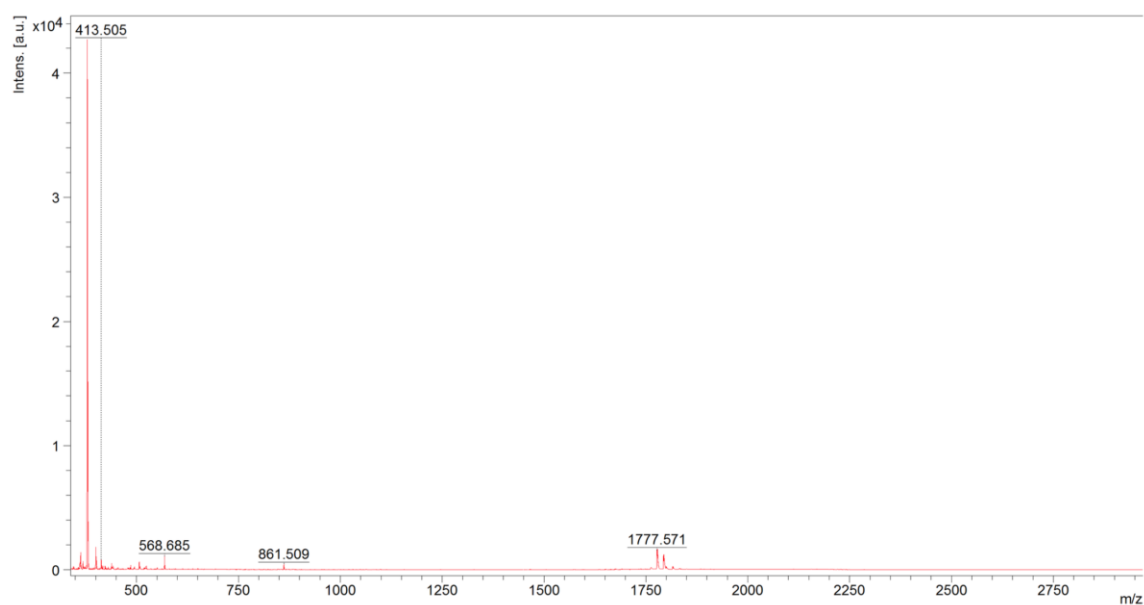


Figure 16. MALDI-TOF MS Results for Insulin Fragment 1 (Knorr). The negative reflection method detected the mass of Insulin Fragment 1 (1754 m/z). Negative mode represents the mass of the peptide with a -1 m/z compared to the actual mass because the negative mode detects the mass and ionization of the peptide after a proton is removed from the carboxyl group at the C-terminus, thus decreasing the mass by ~1 mass unit. When in positive ion mode, the mass of Insulin Fragment 1 plus Na⁺ was detected 1777 m/z (+23). The detection values indicate that the sample contained the desired peptide.

Wang Positive Ion Mode: 1777 m/z (+23)



Wang Negative Reflection: 1754 m/z

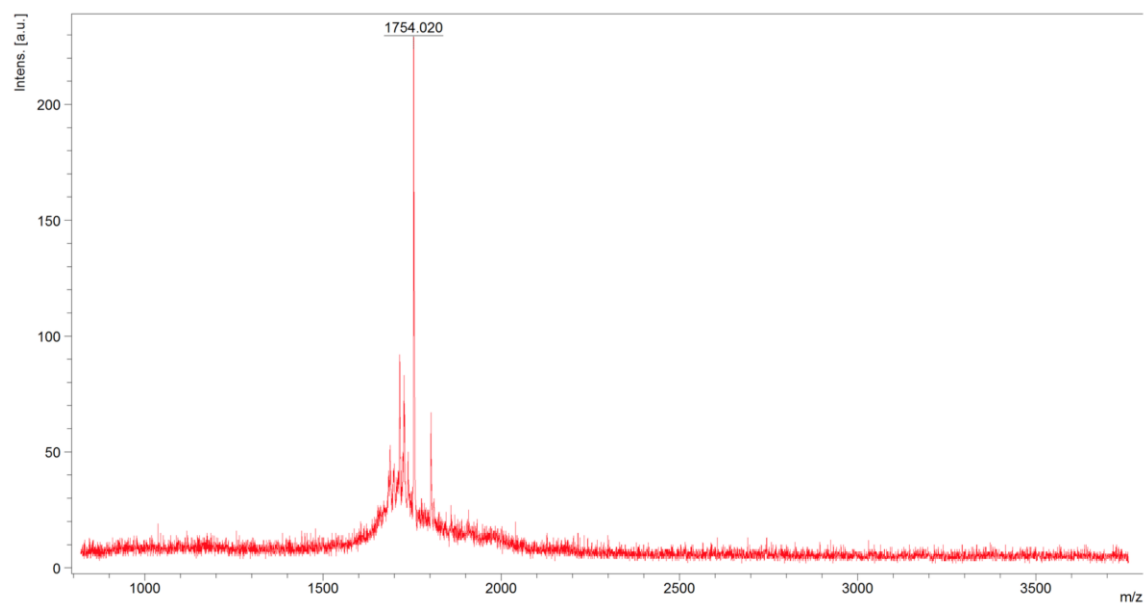


Figure 17. MALDI-TOF MS Results for Insulin Fragment 1 (Wang). The negative reflection method detected the mass of Insulin Fragment 1 (1754 m/z). Negative mode represents the mass of the peptide with a -1 m/z compared to the actual mass because the negative mode detects the mass and ionization of the peptide after a proton is removed from the carboxyl group at the C-terminus, thus decreasing the mass by ~1 mass unit. When in positive ion mode, the mass of Insulin Fragment 1 plus Na⁺ was detected 1777 m/z (+23). The detection values indicate that the sample contained the desired peptide.

Discussion:

The aim of this project was to optimize the synthesis of each of the three fragments of insulin, as shown in Figure 9, in order to verify that the synthetic route developed results in active insulin. Several challenges involving synthesis, purification, and percent yield were encountered along the way. However, steps were taken to troubleshoot and fix these problems in order to systematically progress and optimize the conditions for synthesis.

During manual synthesis of Insulin Fragment 1, the pre-loaded Wang resin was initially used to initiate peptide synthesis via an ester bond. A solution of 2% DBU/2% Piperidine in DMF was used as the base to deprotect each amino acid during coupling. However, this method resulted in low yields of the desired peptide. It was found that 2% DBU/2% Piperidine in DMF solution was too strong of a base, and in turn catalyzed an unfavorable side reaction, DKP formation, resulting in lower yields of the desired peptide (Figure 18). In an attempt to fix this problem, a weaker base, 20% Piperidine solution, was used in the deprotection step, and higher yields were obtained.

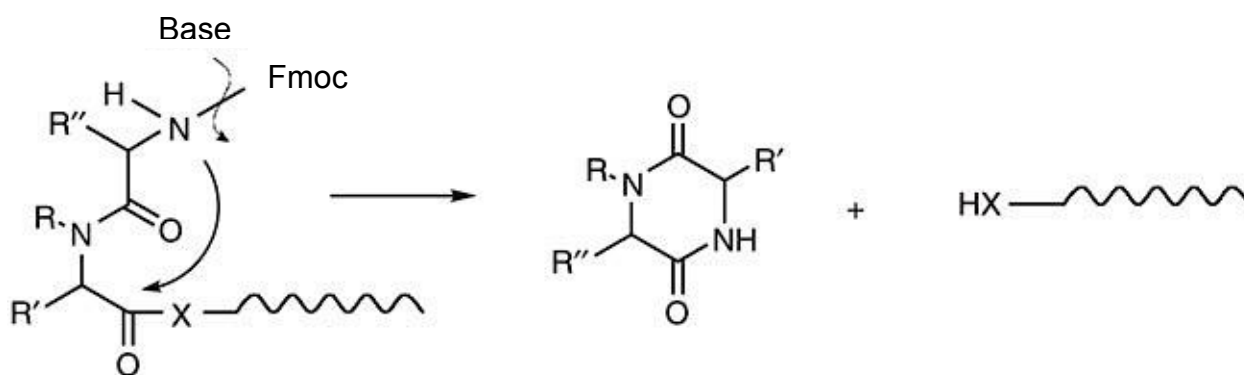


Figure 18. DKP (Diketopiperazine) Formation in solid phase peptide synthesis. DKP formation is an unfavorable side reaction catalyzed by a strong base during deprotection.

Periodically during manual synthesis, after coupling approximately four to seven amino acids, a small cleavage was performed to determine whether or not the synthesis was effective. After small cleavage, the sample was tested on the MALDI-TOF MS in order to determine whether the sample contained the desired mass. In some cases, none of the side chains of the amino acids in the peptide sequence could be easily ionized, and therefore the peptide's mass could not be detected. As a result, Fmoc had to be removed to expose the amino group at the N-terminus of the peptide sequence, because the amino group was easily ionizable for the MALDI to detect. However, after several attempts, the mass of the desired peptide sequence was still not detected.

In an attempt to troubleshoot, Insulin Fragment 1 was re-synthesized using the Knorr (Rink-Amide MBHA) resin instead of the pre-loaded Wang resin. The Knorr resin initiates peptide synthesis via an amide bond as opposed to the ester bond associated with the pre-loaded Wang resin. The pre-loaded Wang resin already has asparagine, the first amino acid in the Insulin Fragment 1 sequence, coupled to the resin. In order to load the Knorr resin with asparagine via an alternate route, Fmoc-L-Asp-OtBu was coupled to the Knorr resin to form an amide bond, which in turn formed asparagine. Changing the base used for deprotection from a solution of 2% DBU/2% Piperidine in DMF to 20% Piperidine, along with switching the resin from Wang to Knorr, allowed manual synthesis of Insulin Fragment 1 to be optimized. Synthesis on the Knorr resin ultimately led to a higher percent yield of the pure peptide, as discussed below. To increase efficiency, an AAPPTec Focus-XC peptide synthesizer was then used to optimize the synthesis of each insulin fragment.

After optimization of the synthesis of Insulin Fragment 1, Reversed-Phase High Performance Liquid Chromatography (RP-HPLC) was used to purify the peptide. However, purification was often difficult and several challenges were encountered. Initially, it was difficult

to dissolve the peptide in 3:1 18 Ω M H₂O: ACN solution when preparing the sample for purification. To facilitate dissolution, different concentrations of 18 Ω M H₂O: ACN were experimented with. Additionally, sonication was used to help promote the dissolution of the peptide. Often times, by the time the peptide dissolved, the solution was too dilute. In an attempt to decrease the total volume of 18 Ω M H₂O: ACN solution needed for dissolution, a small amount of base, NaOH, was used to help dissolve the peptide. When base was added, the pH of the solution was monitored to make sure that the pH fell between a pH range of 5 to 8. If too much base was added, acid was used to neutralize the solution before purification. The pH had to be adjusted because if the pH gets too high, the solution would become a strong base, causing the proton on each cysteine residue to be deprotonated, which in turn causes the formation of unwanted disulfide bridges and ruins the column.

Once the method for dissolution was optimized, the sample was injected for purification. Initially, a 10 μ m column was used to separate the crude peptide during the first prep run. The 10 μ m value indicates the size of the particles within the column, which act as the stationary phase for separation. These particles have different affinities for different components in the sample based on their structure and polarity. Solvents are pumped through the column, creating high pressure to force the mobile phase through the column, allowing for separation of the sample. Smaller particles within the column allow for a more effective separation, and yield peaks on the chromatogram with a higher resolution. In this case, the 10 μ m column did not separate the peptide very well, yielding fractions with low purity. Therefore several prep runs were needed in order to obtain the peptide with a purity of >90%. However, after several prep runs, despite obtaining a peptide with high purity, the yield of the desired peptide was very low. There is some material loss, ~5-10% of the sample, with each HPLC run. Since several prep runs were needed to purify

the sample, a lot of material was lost during the purification process, contributing to the low yield. In an attempt to obtain a higher overall yield of the pure peptide, a 5 μm column was ordered for the HPLC. The 5 μm column allowed for better separation of the peptide than the 10 μm column. Because of this better separation, fewer prep runs were needed to yield a peptide with >90% purity. The percent yield of the peptide with >90% purity that was synthesized from the Knorr resin was 2.86%, and the yield from that synthesized using the Wang resin was 1.45%. These yields are still clearly low, so the conditions for peptide synthesis and purification are still being optimized.

Developing an active and stable insulin analog will significantly improve means of administering insulin to diabetic patients. Currently, insulin needs to be kept cold during transport and storage to prevent irreversible fibrillation and thus, inactivity. However, the need for refrigeration serves as a problem in many developing countries where electricity is simply not available nor affordable in numerous communities. This poses a risk to individuals with diabetes in communities who are suffering from a disease that they either cannot afford, do not have proper access to, or do not have means to store the treatment for. If an active analog can be made that remains stable under various temperatures and non-ideal conditions, it could benefit countless people in underserved and impoverished areas. From an economic standpoint, it would also reduce the cost of insulin in industrial areas by eliminating the cost of refrigeration for insulin storage and transport.

Future Plans:

After designing and constructing an insulin analog, its activity will be assessed using cellular assays used to test insulin activation of the insulin receptor signaling cascade. These assays will be used to observe the trends in activity based on the incorporation of different modifications. The insulin analog's stability will then be compared to that of native insulin using various stability assays. Structural analysis will then be done on selected, unique analogs to align and assess the relationship between structure, function, and stability.

References:

- Avital-Shmilovici, M., Mandal, K. Kent, SB., et al. (2013). Fully convergent chemical synthesis of ester insulin: determination of the high resolution x-ray structure by racemic protein crystallography. *Journal of the American Chemical Society*. 135: 3173-3185.
- “Basics About Diabetes.” (2014, October 21). Retrieved February 19, 2015, from <http://www.cdc.gov/diabetes/basics/diabetes.html>
- Belgi, A., Hossain, MA., Tregear, GW., Wade, JD. (2011). The chemical synthesis of insulin from past to the present. *Immunology, Endocrine, & Metabolic Agents in Medicinal Chemistry*. 11: 40-47.
- Brange, J., Andersen, L., Laursen, ED., et al. (1997). Toward understanding insulin fibrillation. *Journal of Pharmaceutical Sciences*. 86(5): 1-9.
- Derewenda, U., Derewenda, Z., Dodson, EJ., et al. (1991). X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. *Journal of Molecular Biology*. 220: 425-433.
- Leyer, S., Gattner, HG., LeithaUser, M. (1995). The role of the C-terminus of the insulin B-chain in modulating structural and functional properties of the hormone. *International Journal of Peptide and Protein Research*.. 46: 397-407.
- Liu, F., Luo, EY., Flora, DB., et al. (2014). A synthetic route to human insulin using isoacyl peptides. *Angewandte Chemie*. 53: 3983-3987.
- Liu, F., Luo, EY., Flora, DB, Mayer, JP. (2013). Concise synthetic routes to human insulin. *Organic Letters*. 15: 960-963.

- Mayer, J., Zhang, F., DiMarchi, RD. (2007). Insulin structure and function. *Journal of Peptide Science*. 88: 687-713.
- McCarthy M., Menzel, S. (2001). The genetics of type 2 diabetes. *British Journal of Clinical Pharmacology*. 51(3): 195-199.
- Olefsky, JM. (1976). The insulin receptor: its role in insulin resistance of obesity and diabetes. *Diabetes*. 25(12): 1154-1161.
- Ratanji, KD., Derrick, JP., Kimber, I. (2014). Immunogenicity of therapeutic proteins: influence of aggregation. *Journal of Immunotoxicology*. 11(2): 99-109.
- Ratner, RE., Phillips, TM., Steiner, M. (1990). Persistent cutaneous insulin allergy resulting from high-molecular-weight insulin aggregates. *American Diabetes Association*. 39(6): 728-733.
- Sluzky, V., Klibanov, AM., Langer, R. (1992). Mechanism of insulin aggregation and stabilization in agitated aqueous solutions. *Biotechnology and Bioengineering*. 40: 895-903.
- Zaykov, AN., Mayer, JP., Gelfanov, VM., et al. (2013). Chemical synthesis of insulin analogs through a novel precursor. *ACS Chemical Biology*. 9(3): 683-691.