

University at Albany, State University of New York

Scholars Archive

Biological Sciences

Honors College

5-2023

Progress Towards Peptide and Protein Synthesis Involving Thiol Moieties

Ryan Lo Chen

The University at Albany community has made this article openly available.

Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/honorscollege_biology



Part of the [Biology Commons](#)

Rights Statement



License

Progress towards Peptide and Protein Synthesis involving Thiol Moieties

An honors thesis presented to the
Department of Biological Sciences,
University at Albany, State University of New York
in partial fulfillment of the requirements
for graduation with Honors in Biological Sciences
and
graduation from The Honors College

Ryan Lo Chen

Research Mentor: Baotong Tian
Research Advisor: Dr. Qiang Zhang, Ph.D.
Committee Member: Dr. Andrew Berglund, Ph.D

May 2023

Abstract

Polypeptides and proteins play crucial roles in biological systems, including enzymatic activity regulation, molecule transport, and cell and tissue structural support. Amino acids are the building blocks of peptides and proteins, and peptides can act as hormones, neurotransmitters, and enzymes. Proteins are made up of one or more polypeptide chains and have diverse functions such as catalyzing chemical reactions, transporting molecules, and providing structural support. Protein preparation can be done biologically by expression systems or isolation and purification from natural sources. While these methods have their advantages, they also have their shortcomings. Chemical synthesis is an alternative method that allows for the introduction of post-translational modifications and non-canonical amino acids with precise control over their location and frequency. Solid-phase peptide synthesis (SPPS) is a widely used chemical synthesis method that involves the use of a solid support, typically a resin, to which the first amino acid of the peptide is attached. Subsequent amino acids are then added one at a time, in a stepwise fashion, with each amino acid protected by a temporary functional group that prevents unwanted side reactions. The use of a solid support allows for easy purification of the desired peptide and allows for the synthesis of longer and more complex peptides. Chemical synthesis methods have advantages such as precise control over PTMs, introducing non-canonical amino acids, and the ability to produce proteins that are difficult to obtain through biological methods. In this thesis, we propose two primary methods of peptide synthesis; 1.) intermolecular bonding of disulfide bond using and 2.) synthesis of protecting group-free sugar-containing oligopeptides. The study aims to synthesize a disulfide-rich peptide using solid-phase peptide synthesis and reacting it with a phosphate buffer to investigate the formation of mixed disulfide bonds. The experiments involve various analytical techniques, such as HPLC and mass spectrometry. Overall, these studies enriched our understanding of thiol functional group reactivities in the polypeptide settings and complemented the existing thiol oxidation tactics.

Acknowledgements

I am writing this to express my sincere gratitude and appreciation to Professor Zhang, Qiang and Professor Berglund, Andrew for your invaluable guidance, support, and encouragement throughout my undergraduate research thesis journey. Without your unwavering commitment, it would have been impossible for me to accomplish this important milestone in my academic career. Your profound knowledge, expertise, and dedication to your work have been a constant source of inspiration and motivation for me. You have challenged me to think critically and creatively, and your constructive feedback and suggestions have helped me to refine my research and writing skills.

I also want to extend my thanks to the rest of the faculty and staff, Bao Tong, Huan, and Wayne who have contributed to my academic and personal growth during my time at the university. Their unwavering support, guidance, and encouragement have been invaluable to me. I would like to express my deepest gratitude to my family and friends for their unwavering support and encouragement throughout my academic journey. Their love, guidance, and encouragement have kept me motivated and focused during the most challenging times. Especially, my parents, who have supported and encouraged me for over 23 years. Once again, thank you for everything. Your support and guidance have been instrumental in helping me to achieve this significant accomplishment.

List of Figures (if any)

Figure 1. General approach of Solid-phase peptide synthesis.....	4
Figure 2. Hypothesize steps of intermolecular bonding of disulfide bond.....	7
Figure 3. Hypothesize steps of Synthesis of Protecting group-free Thiol-containing OIpeptides	9
Figure 4. Solid-phase peptide synthesis of GC(trt)GPen(trt)Gresin.....	10
Figure 5. Side chain deprotection and resin cleavage of GC(trt)GPen(trt)G-TGT resin.....	11
Figure 6. Formation of intermolecular disulfide bond of GCGPenG.....	11
Figure 7. Preparation of Boc- β -Thiolactone.....	12
Figure 8. Peptide synthesis in the basic aqueous buffer of acetylhomocysteinylglycine.....	13
Figure 9. Peptide synthesis in the basic aqueous buffer of acetylhomocysteinylalanine.....	13
Figure 10. Peptide synthesis in the basic aqueous buffer of Boc-cysteinylglycine.....	13
Figure 11. Preparation of Fmoc-Ala-TGT Resin.....	14
Figure 12. SPPS & side chain deprotection and resin cleavage of Fmoc-CGA.....	14
Figure 13. SPPS & side chain deprotection and resin cleavage of CGA.....	15
Figure 14. SPPS & side chain deprotection and resin cleavage of Fmoc-CGFA.....	15
Figure 15. SPPS & side chain deprotection and resin cleavage of CGFA.....	15
Figure 16. SPPS & side chain deprotection and resin cleavage of Fmoc-PenGFA.....	15
Figure 17. Peptide cyclization of acetylhomocysteinylglycine by tert-Butyl isocyanide.....	16
Figure 18. Peptide cyclization of Boc-CG by tert-Butyl isocyanide.....	17
Figure 19. Peptide cyclization of Fmoc-CGFA by tert-Butyl isocyanide.....	17
Figure 20. Peptide cyclization of CGA by tert-Butyl isocyanide.....	17
Figure 21. Peptide cyclization of Fmoc-CGFA by tert-Butyl isocyanide.....	17
Figure 22. Peptide cyclization of CGFA by tert-Butyl isocyanide.....	18
Figure 23. Peptide cyclization of Fmoc-PenGFA by tert-Butyl isocyanide.....	18
Figure 24. Mass spectrometry of peptide (GCGPenG).....	19
Figure 25. Mass spectrometry of intermolecular disulfide bond in two GCGPenG.....	19
Figure 26. Mass spectrometry of intermolecular disulfide bond in two GCGPenG.....	20
Figure 27. Mass spectrometry of peptide (acetylhomocysteinylglycine).....	21
Figure 28. Mass spectrometry of peptide (acetylhomocysteinylalanine).....	21
Figure 29. Mass spectrometry of peptide (Boc-cysteinylglycine).....	22
Figure 30. Mass spectrometry of peptide (Fmoc-CGA).....	22
Figure 31. Mass spectrometry of peptide (CGA).....	23
Figure 32. Mass spectrometry of peptide (Fmoc-CGFA).....	23
Figure 33. Mass spectrometry of peptide (CGFA).....	24
Figure 34. Mass spectrometry of peptide Fmoc-PenGFA.....	24
Figure 35. Mass spectrometry of cyclized acetylhomocysteinylglycine.....	25
Figure 36. Mass spectrometry of cyclized Boc-CG.....	25
Figure 37. Mass spectrometry of cyclized Fmoc-CGA.....	26
Figure 38. Mass spectrometry of cyclized CGA.....	26
Figure 39. Mass spectrometry of cyclized Fmoc-CGFA.....	27
Figure 40. Mass spectrometry of cyclized CGFA.....	27
Figure 41. Mass spectrometry of cyclized Fmoc-PenGFA.....	28

Table of Contents

Abstract	ii
Acknowledgements	iii
List of Figures	iv
Table of Contents	v
Introduction	1
1.1. Synthesis of Disulfide Bond.....	6
1.2. Synthesis of Protecting group-free Thiol-containing Oligopeptides.....	7
Materials and Methods	10
2.1. Synthesis of Disulfide Bond.....	10
2.2. Synthesis of Protecting group-free Thiol-containing Oligopeptides.....	12
Results and Discussions	19
3.1. Synthesis of Disulfide Bond.....	19
3.2. Synthesis of Protecting group-free Thiol-containing Oligopeptides.....	21
Conclusions	29
References	32

1. Introduction

Amino acids, peptides, and proteins are essential components of biological systems and play a crucial role in various physiological processes. These biomolecules are involved in a wide range of functions, including the transport of molecules, the regulation of enzymatic activity, and the formation of structural components of cells and tissues. Amino acids are the building blocks of peptides and proteins. There are 20 naturally occurring amino acids, each with a unique side chain that imparts specific chemical properties. These amino acids are linked together by peptide bonds to form peptides and proteins. Peptides are a class of biomolecules that are essential for the functioning of living organisms. [1] They are composed of amino acids, which are the building blocks of proteins. Peptides can have a wide range of functions in the body, including acting as hormones, neurotransmitters, and enzymes. Peptides are formed by the condensation of amino acids, in which the carboxyl group of one amino acid reacts with the amino group of another amino acid, resulting in the formation of a peptide bond. The sequence of amino acids in a peptide determines its three-dimensional structure and its function.[2] In addition to being used to build proteins, amino acids also serve as precursors for a variety of important biological molecules, such as nucleotides. Peptides also play important roles in the immune system, by acting as antigens and eliciting an immune response. Furthermore, peptides can act as cell signaling molecules, by binding to specific receptors on the surface of cells and initiating a signaling cascade.[1] Proteins are large biomolecules that are made up of one or more polypeptide chains. The functions of proteins are diverse and include catalyzing chemical reactions, transporting molecules, and providing structural support. The unique properties of proteins, such as their ability to fold into complex three-dimensional structures, are critical for their function. The folding and unfolding of proteins can be influenced by a variety of factors, including changes in temperature, pH, and the presence of other molecules.[1]

There are several methods used for the preparation of protein biologically, one of the most common methods used for the preparation of proteins biologically is with expression systems. These systems involve

the transfer of genetic material encoding the desired protein into a host cell, which then produces the protein. There are several types of expression systems, including bacterial, yeast, insect, and mammalian systems. Another method is through isolation and purification from natural sources. This involves the extraction of the protein from its natural source, followed by purification to remove impurities and obtain a pure protein. While these methods have their advantages, such as high specificity and purity, they also have their shortcomings, biological preparation of proteins can be complex and time-consuming, involving multiple steps such as gene cloning, cell culture, protein expression, and purification. The yield of protein obtained through biological methods can be low, especially for complex proteins that require extensive post-translational modifications. Also, its inability to provide post-translationally modified (PTM) motifs, such as glycosylation, phosphorylation, and nitration. PTMs are essential for the proper functioning of many proteins, and their absence can limit the functionality of produced proteins. Furthermore, PTMs lack template control, making them difficult to replicate through biological methods. Moreover, introducing non-canonical amino acids can be challenging through biological methods.[3] Chemical synthesis is an alternative method for preparing proteins that involves the stepwise assembly of amino acids in a controlled manner. This approach allows for the introduction of PTMs and non-canonical amino acids with precise control over their location and frequency.[4] Moreover, chemical synthesis can be applied to produce proteins that are difficult to obtain through biological methods. The ability to synthesize peptides in a controlled and efficient manner is essential for advancing the field of peptide research and its applications. Chemical peptide synthesis involves the chemical assembly of amino acids in a specific order to form a peptide chain. Chemical peptide synthesis can be carried out using a variety of methods, including solution-phase peptide synthesis, solid-phase peptide synthesis, native chemical ligation, and recombinant DNA technology.

In solution-phase synthesis, the amino acids are dissolved in a solvent and the condensation reaction is carried out in solution. Solution-phase synthesis is relatively straightforward and can be used to synthesize peptides up to a certain length. However, as the peptide length increases, the purification of the

desired peptide becomes more challenging and time-consuming. Solid-phase peptide synthesis (SPPS) was developed as a solution to the challenges of solution-phase synthesis. SPPS is a widely used method for peptide synthesis that was developed by R.B. Merrifield in the 1960s.[5] SPPS involves the use of a solid support, typically a resin, to which the first amino acid of the peptide is attached. Subsequent amino acids are then added to the growing peptide chain one at a time, in a stepwise fashion, with each amino acid protected by a temporary functional group that prevents unwanted side reactions. The use of a solid support allows for easy purification of the desired peptide and allows for the synthesis of longer and more complex peptides.

The chemical reactions involved in SPPS include the activation of carboxylic acids and the coupling of amino acids. The first amino acid is attached to the solid support via a linker, and the amino group is protected by a temporary functional group, such as the Fmoc (9-fluorenylmethyloxycarbonyl) group or the Boc (tert-butyloxycarbonyl) group. The functional group is deprotected by a deblocking reagent, such as piperidine and DBU (1,8-Diazabicyclo). The carboxylic acid group of the first amino acid is activated by a coupling reagent, such as HATU (Tetramethyl Uronium), DIPEA, (Ethlydiisopropylamine N,N- Diisopropylethylamine), HOBt (1-hydroxybenzotriazole), or DIC (diisopropylcarbodiimide), which then reacts with the amino group of a second temporarily protected amino acid residue, which is functional group.[5] The process is repeated until the desired peptide sequence is obtained. After completion of the peptide chain, the peptide is cleaved from the resin and the protecting groups are removed, resulting in the final peptide product.

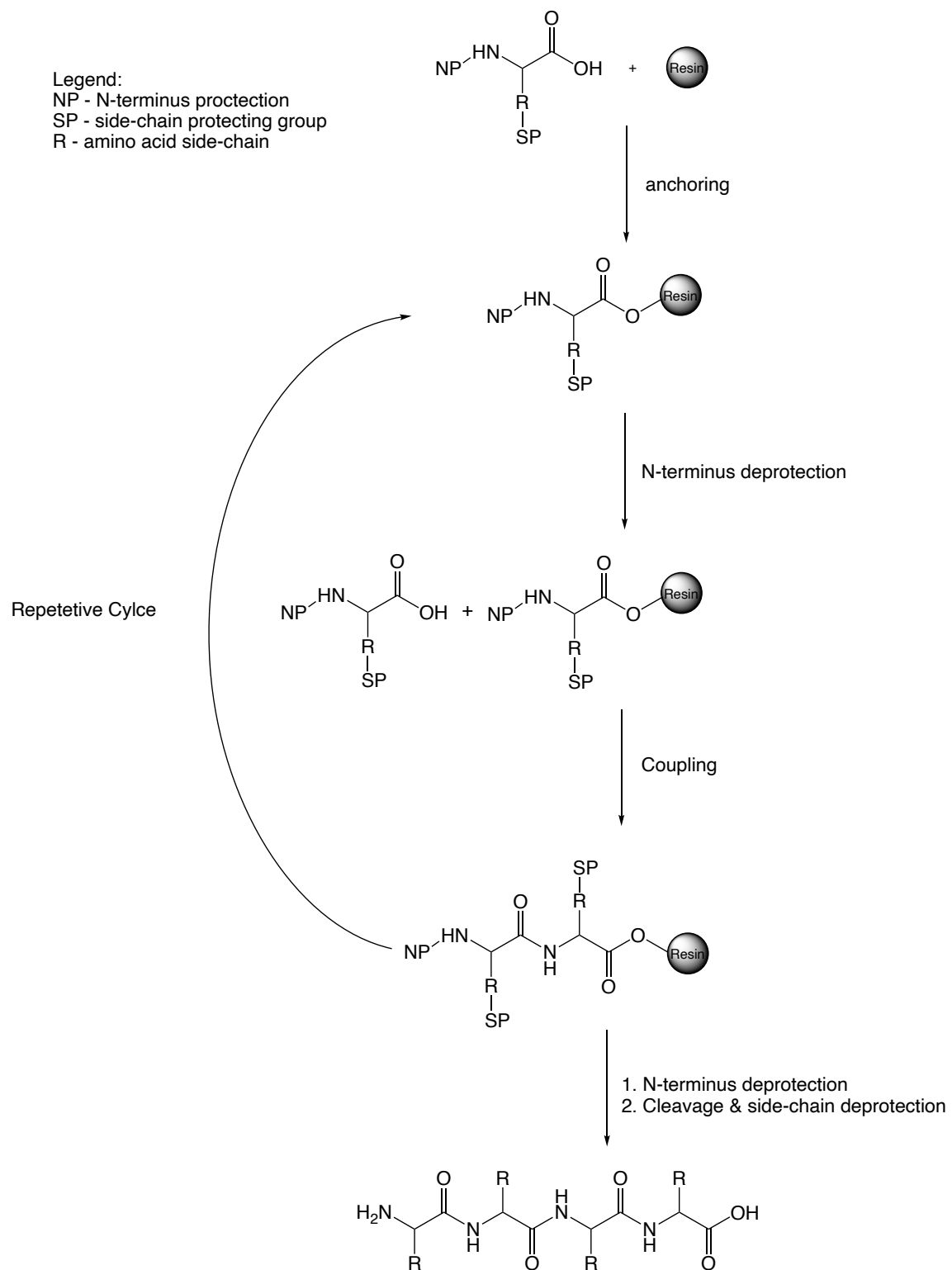


Figure 1: General approach of Solid-phase peptide synthesis.

SPPS has several advantages over other methods for peptide synthesis. It allows for the synthesis of longer and more complex peptides with greater efficiency and purity than ever before. The use of solid supports also allows for the synthesis of peptide libraries, in which large numbers of peptides can be synthesized simultaneously for use in drug discovery and materials science. But there are some limitations for SPPS, such as difficulty in synthesizing peptides that contain certain amino acid residues, which can be overcome with side chain protecting group, for example Trt (triphenylmethane) protecting group, the potential for racemization, which can occur when coupling amino acids during the synthesis. Racemization results in the production of a mixture of both D- and L- amino acid residues, which can negatively impact the biological activity of the synthesized peptides, and the length of the protein that can be synthesized using SPPS. Besides two peptidyl fragments cannot be chemically connected easily using SPPS.[6] However, a solution to this problem is provided by native chemical ligation (NCL). It is a powerful technique that allows for the chemo selective connection of two unprotected synthetic peptides under physiological conditions. This means that two peptidyl fragments can be connected using NCL without the need for protecting groups. The technique of NCL has revolutionized peptide synthesis by enabling the efficient synthesis of longer peptides and proteins that were previously not possible with SPPS alone. By combining the power of SPPS with NCL, it is now possible to synthesize proteins with complex structures and biological activities.[7] Despite these limitations, SPPS remains a valuable tool for synthesizing peptides, especially for peptides that are relatively short and contain common amino acids. Researchers are continually developing new strategies to overcome the limitations of SPPS, such as using alternative coupling reagents or modifying the resin used in the synthesis process.

In this thesis, we propose a central hypothesis that involves two main methods of peptide synthesis. The first method is 1.) intermolecular bonding of disulfide bond forming disulfide-rich peptides and the second method is 2.) the synthesis of protecting group-free thiol-containing oligopeptides via activation of γ -thiolactones and isocyanide cyclization. To test our hypothesis, we will conduct a series of experiments that will utilize various analytical techniques such as HPLC and mass spectrometry. Using solid-phase

peptide synthesis as the base of our peptide related experiments. Our goal is to contribute to the development of new and efficient methods for peptide synthesis with further potential.

1.1 Synthesis of Disulfide Bond

Disulfide-rich peptides are peptides that contain two or more disulfide bonds. Disulfide-rich peptides are widespread and occur naturally in nature, carrying different biological functions. Disulfide-rich peptides is used in many biochemistries and medicinal research as a potential curing molecule. Ribosomal expression of polypeptide chains and the subsequent oxidative folding is obligatory in natural production of disulfide-rich peptides.[8] Precise pairing of cysteine residues to form native disulfides, which is a process administered by molecular folding chaperones and enzymes catalyzing the disulfide formation/isomerization, determines the accuracy of folding of disulfide-rich peptides. Moreover, disulfide-rich peptides can be synthesized artificially and to further aid the preciseness of folded disulfide-rich peptides, both spontaneous oxidation and orthogonal protecting group strategies have been used.[9]

These strategies involve the use of specific chemical compounds that can selectively protect certain amino acid residues during the peptide synthesis process, allowing for precise control of the formation of disulfide bonds. Additionally, advances in peptide synthesis techniques, such as solid-phase peptide synthesis, have made it easier to produce large quantities of disulfide-rich peptides for use in various applications. One potential application of disulfide-rich peptides is in the development of new therapeutics. Due to their unique structural features and biological activities, disulfide-rich peptides have shown promise as drugs for the treatment of a variety of diseases, including cancer, diabetes, and microbial infections.[8] Their ability to target specific cellular receptors and signaling pathways has made them attractive candidates for the development of targeted therapies.

Another area where disulfide-rich peptides have shown potential is in the development of biomaterials. By incorporating disulfide-rich peptides into various materials, researchers can create novel

materials with unique mechanical and biological properties. For example, disulfide-rich peptides have been used to create hydrogels for tissue engineering applications, as well as to improve the mechanical properties of existing biomaterials.[8] Overall, the versatility and potential applications of disulfide-rich peptides make them an exciting area of research in both biochemistry and medicine.

The primary objective of this study is to become familiar with the techniques involved in solid-phase peptide synthesis. Specifically, the focus will be on synthesizing a disulfide-rich peptide, which requires the formation of intermolecular disulfide bonds. The peptide will be synthesized using solid-phase peptide synthesis, which is a well-established method in the field of peptide chemistry. Once the peptide has been synthesized, the intermolecular disulfide bonds will be formed by reacting the peptide with a phosphate buffer.[10] This approach will enable the investigation of the formation of mixed disulfide bonds as an initial study. Through this work, it is hoped that a deeper understanding of the chemistry and biology of disulfide-rich peptides can be gained, leading to the development of new and effective therapeutic agents for various diseases with another research group. Upon completion of the first project, I engaged more challenging protein synthesis.

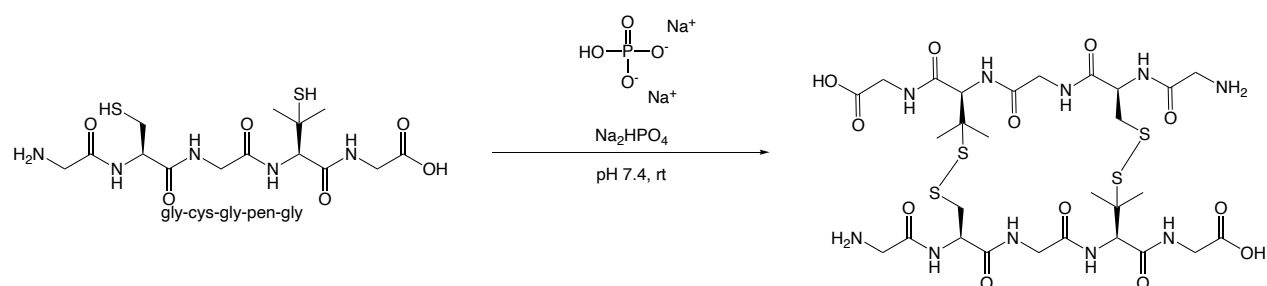


Figure 2: Hypothesize steps of intermolecular bonding of disulfide bond

1.2 Synthesis of Protecting group-free Thiol-containing Oligopeptides

Peptides have emerged as novel therapeutic agents with good pharmacological properties such as high target selectivity, low toxicity, a low tendency to develop resistance, and a high potential for targeting

diseases associated with protein-protein interactions, which small molecule drugs struggle to achieve. Which makes peptides attractive targets for synthetic and medicinal chemists. Since amide bonds are the primary chemical linkages in peptides and proteins, efforts to create new, efficient methods for amide bond formation continue to be crucial chemistry tasks.[11]

There are various methods for synthesizing polypeptides, proteins, and glycoproteins, such as solid-phase peptide synthesis (SPPS), a process of linking amino acids to resin, and insoluble polymeric support, to form peptides. The peptide is expanded by a sequence of addition cycles. The nature of the solid-phase approach is that the excess soluble reagents in the reaction can be removed easily, by filtration and washing without any losses of the product. Once the peptide chain has been completed, the peptide chain will be released from the resin by deprotection.[12] Yet, it is atom uneconomical because it requires protected amino acid and loads of coupling agent. For example, thiols, have a strong nucleophilic activity that allows for a variety of chemical reactions to take place under mild circumstances. Thiols have thus been used in the chemical alteration of biomolecules such as DNA, peptides, and proteins. Using cysteine or homocysteine residues, traditional peptide synthesis can provide thiol-containing peptides. However, due to the reactivity of the sulfhydryl functionality, protection as tert-butylsulfenyl (StBu), S-trityl (STrt), or S-acetomidomethyl (SAcm) groups are frequently used, and the corresponding deprotection steps either involve harsh conditions using strong acids, toxic mercury(II) reagents as milder alternatives, the use of valuable metals, or initial transformation to the corresponding disulfide followed by reduction, all of which make these synthetic strategies environmentally unfriendly and not atom economical. [11]

In this work, we hypothesize that by taking the advantage of γ -thiolactone and its derivatives, we can develop a protocol for protecting group-free thiol-containing peptide synthesis. Using Isocyanide chemistry to cyclize thiol-containing peptide, then activate the cyclized peptide by free amino acid in a slightly basic aqueous buffer. The oligopeptide is formed by repeating the cycle.

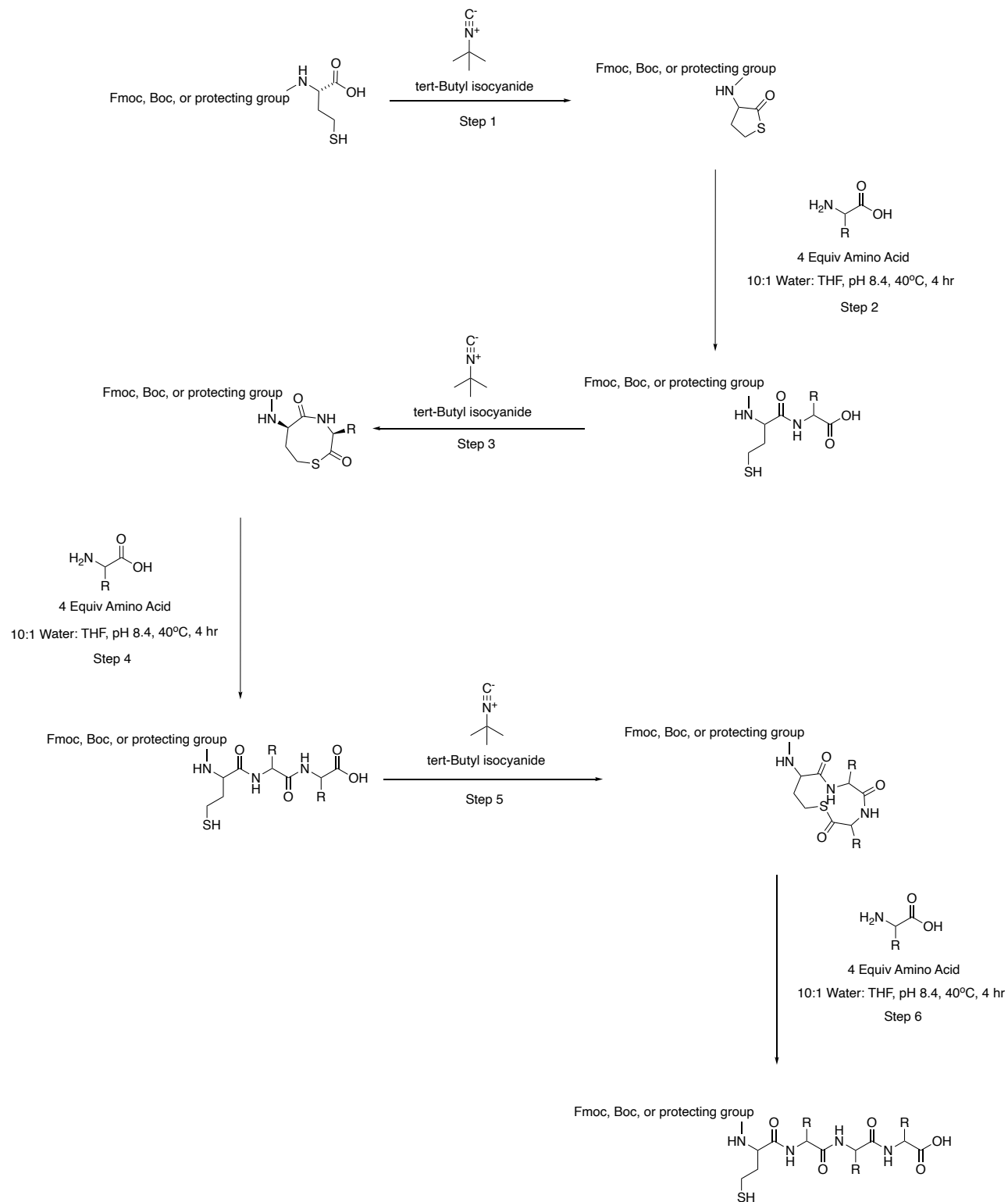


Figure 3: Hypothesize steps of Synthesis of Protecting group-free Thiol-containing Oligopeptides

2. Materials and Methods

2.1. Synthesis of Disulfide bond

2.1.1 Materials

Fmoc-Gly-TGT Resin, Amino Acids (Fmoc-Gly-OH, Fmoc-Cys(Trt)-OH, and Fmoc-Pen(Trt)-OH), Sodium Phosphate Dibasic (HN_2PO_4), Piperidine, 1,8-Diazabicyclo (DBU), Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium (HATU), Ethlydiisopropylamine N,N- Diisopropylethylamine (DIPEA), Dimethylformamide (DMF), Trifluoroacetic Acid (TFA), Triisopropyl silane (TIPS), and Water (H_2O).

2.1.2 Methods

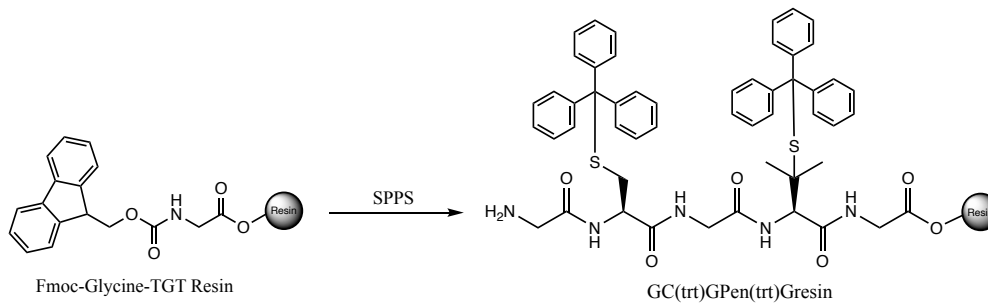


Figure 4: Solid-phase peptide synthesis of GC(trt)GPen(trt)Gresin

Solid-phase peptide synthesis: In figure 4, the solid-phase peptide synthesis (machine) started with deprotection of Fmoc from Fmoc-gly-TGT resin by 2% of piperidine and 2% of DBU in DMF for 7 minutes at room temperature, then an amino acid is coupled with gly-TGT resin by HATU and DIPEA with the equivalent of 1:4:8:8 resin, amino acids, HATU, and DIPEA for 9 minutes at 40°C. The product is then wash with DMF x3 for 30 minutes. This process is repeated with Fmoc-pen-OH, Fmoc-gly-OH, and Fmoc-cys-OH. Lastly, deprotect the final product with 2% of piperidine and 2% of DBU in DMF for 7 minutes at room temperature to remove the Fmoc from the C-terminus.

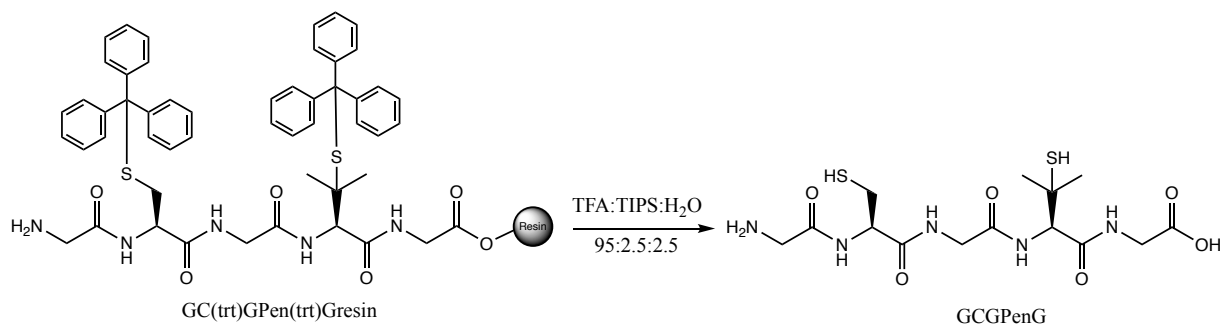


Figure 5: Side chain deprotection and resin cleavage of GC(trt)GPen(trt)G-TGT resin.

Side chain deprotection and resin cleavage: React GC(trt)GPen(trt)G-TGT resin with TFA, TIPS, and H₂O with the ratio of 95:2.5:2.5 for 20 minutes to cleave the TGT resin from the N-terminus and the trityl side chain from penicillamine and cysteine, in figure 5. Then, the product is blew dry with Ar gas and lyophilization with water and ACN with the ratio of 1:1.

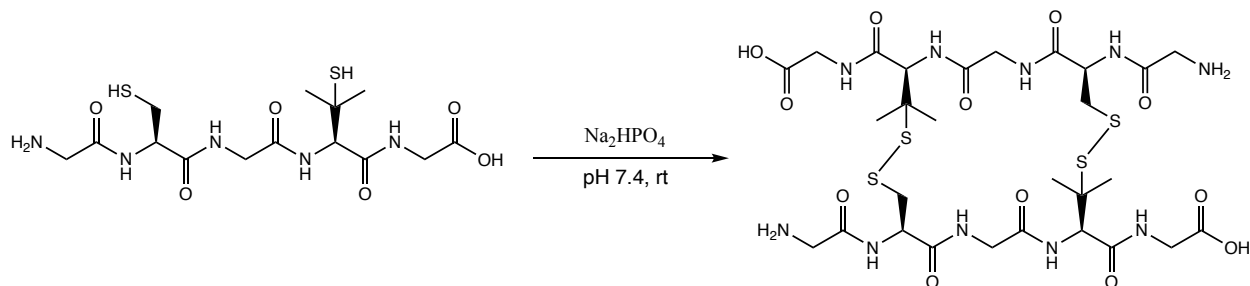


Figure 6: Formation of intermolecular disulfide bond of GCGPenG.

Formation of Disulfide Bond: Dissolve 0.5 mM of purified peptide into 100 mM of phosphate buffer (pH7.4) to achieve a concentration of ~50 μM for 1 day at room temperature, add 5% of TFA, in figure 6. Purify the product with HPLC (a flow rate of 10 mL \cdot min⁻¹ of H₂O (+0.1 TFA) and ACN (+0.1 TFA); isocratic with 5% ACN (+0.1 TFA) for 5 minutes. followed by a linear gradient of 5.0% to 95% ACN (+0.1 TFA) over 40 minutes) and analyzed by mass spectrometry.

2.2. Synthesis of Protecting group-free Thiol-containing Oligopeptides

2.2.1 Materials

3-Acetamidotetrahydro-2-thiophenone, Tetrahydrofuran (THF), tert-butyl isocyanide, Boc-cys-OH, Isobutyl chloroformate (i-BuOCOCl), Triethylamine (Et₃N), Trityl Chloride Resin, N-Methylmorpholine (NMM), Dichloromethane (DCM), Methanol (MeOH), Amino Acids (Glycine, Alanine, Fmoc-Ala-OH, Fmoc-Gly-OH, Fmoc-Cys(Trt)-OH, Boc-Cys(Trt)-OH, Fmoc-Phe-OH, and Fmoc-Pen(Trt)-OH), Piperidine, 1,8-Diazabicyclo (DBU), Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium (HATU), Ethyldiisopropylamine N,N- Diisopropylethylamine (DIPEA), Dimethylformamide (DMF), Trifluoroacetic Acid (TFA), Triisopropyl silane (TIPS), and Water (H₂O).

2.2.2 Methods

2.2.2.1 Preparation of β -Thiolactone

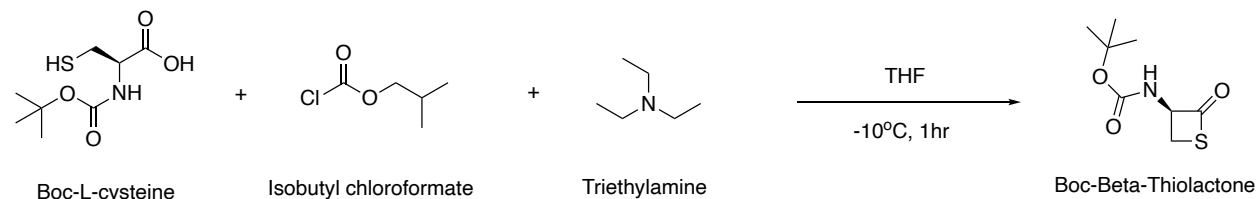


Figure 7: Preparation of Boc- β -Thiolactone

Preparation of β -Thiolactone: In Figure 7, dissolve 1 equivalent of Boc-L-cysteine, 1.5 equivalent of Isobutyl chloroformate, and 2 equivalent of Triethylamine into 0.1 mol/l of THF for 1 hour at -10°C. Use liquid-liquid extraction to obtain the organic layer of the product. Then, purify the product with column chromatography, ethanol: ethyl acetate, 3:1, to obtain pure Boc- β -thiolactone.

2.2.2.2 Cyclized peptide synthesis in the basic aqueous buffer

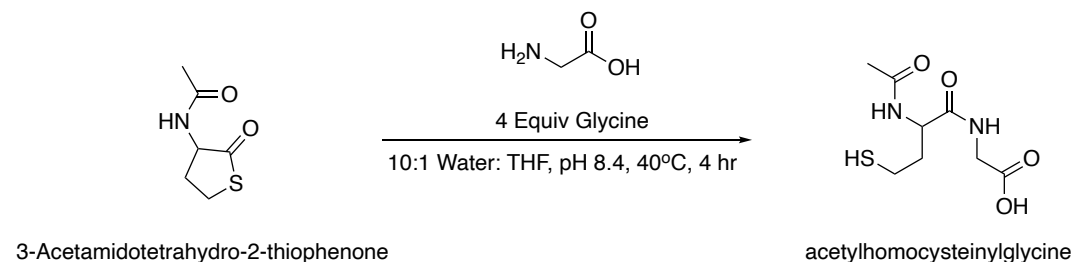


Figure 8: Peptide synthesis in the basic aqueous buffer of acetylhomocysteinyglycine

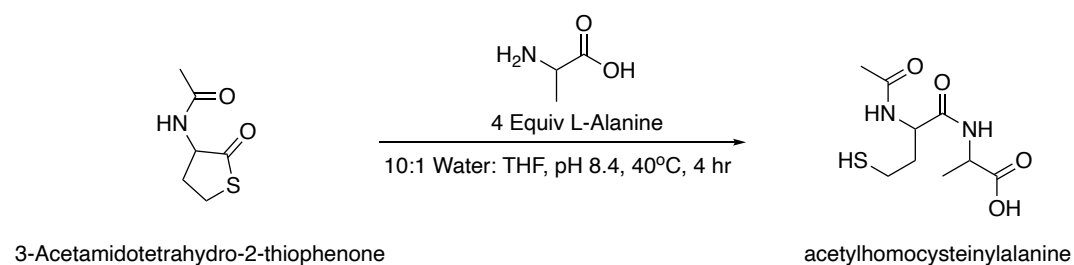


Figure 9: Peptide synthesis in the basic aqueous buffer of acetylhomocysteinyalanine

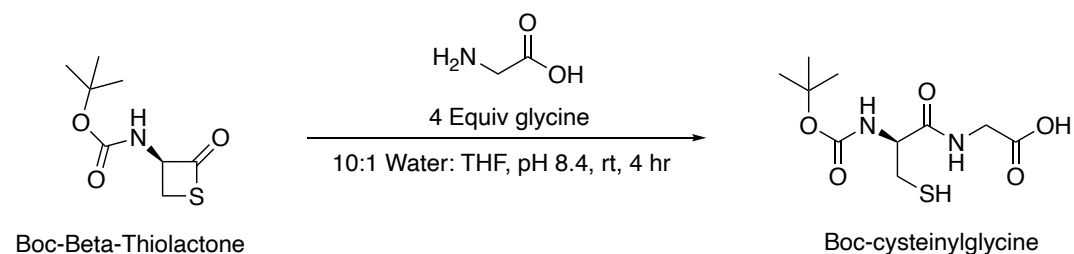


Figure 10: Peptide synthesis in the basic aqueous buffer of Boc-cysteinyglycine

Cyclized peptide synthesis in the basic aqueous buffer: In Figures 8,9 and 10, dissolve 1 equivalent of starting material (3-Acetamidotetrahydro-2-thiophenone or Boc-β-Thiolactone) and 4 equivalent of free-amino acid (Glycine or Alanine) into 10:1 of Water: THF with a pH of 8.4 for 4 hours at 40°C or room temperature. Purify the product with HPLC (a flow rate of 10 mL ·min⁻¹ of H₂O (+0.1 TFA))

and ACN (+0.1 TFA); isocratic with 5% ACN (+0.1 TFA) for 5 minutes. followed by a linear gradient of 10.0% to 40% ACN (+0.1 TFA) over 20 minutes) and analyzed by mass spectrometry.

2.2.2.3 Resin preparation

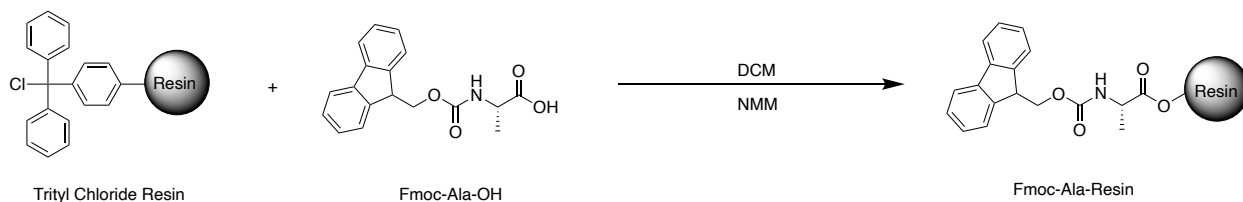


Figure 11: Preparation of Fmoc-Ala-TGT Resin

Preparation of Fmoc-Ala-TGT Resin: In figure 11, dissolve 300mg (1 equivalent) of Trityl Chloride Resin and 4 equivalent of Fmoc-Ala-OH into 8ml of DCM and 0.3ml of NMM. Put it on the shaker and shake overnight at room temperature. Follow with adding 0.4ml of 25% NMM in MeOH into the reaction and shake for another hour. Then, wash it with DMF, MeOH, and DCM twice. Last, use the vacuum to dry the resin.

2.2.2.4 Solid-phase peptide synthesis & Side chain deprotection and resin cleavage:

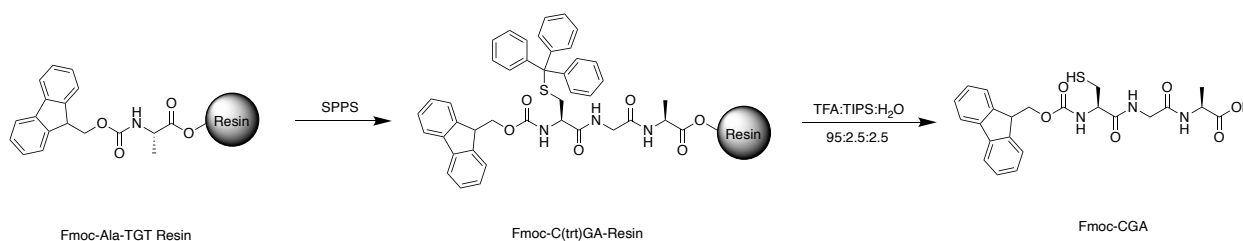


Figure 12: Solid-phase peptide synthesis & side chain deprotection and resin cleavage of Fmoc-CGA

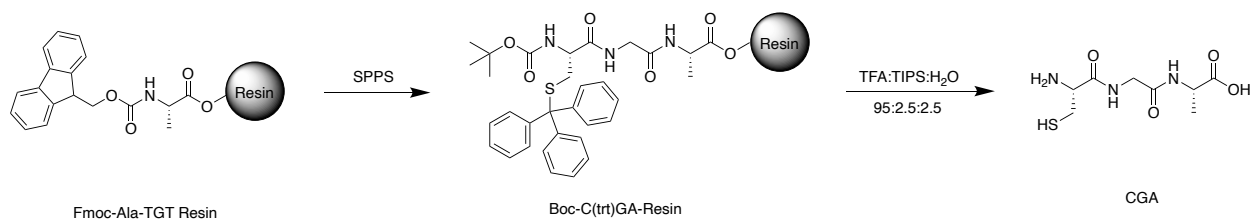


Figure 13: Solid-phase peptide synthesis & side chain deprotection and resin cleavage of CGA

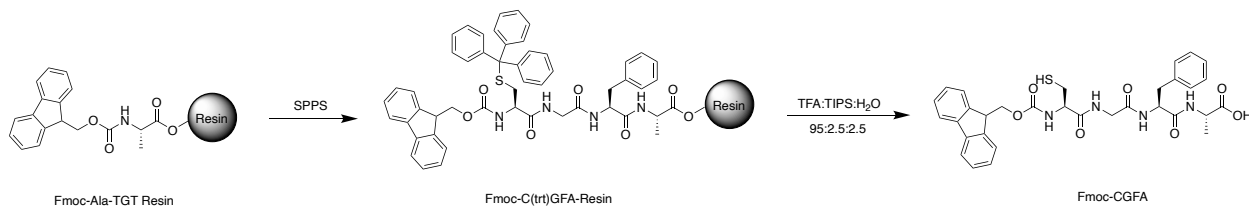


Figure 14: Solid-phase peptide synthesis & side chain deprotection and resin cleavage of Fmoc-CGFA

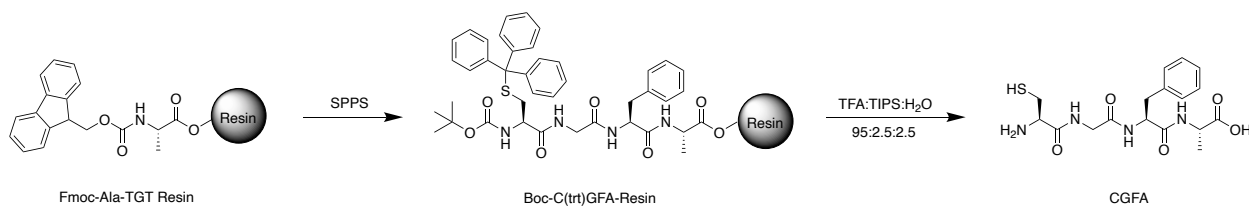


Figure 15: Solid-phase peptide synthesis & side chain deprotection and resin cleavage of CGFA

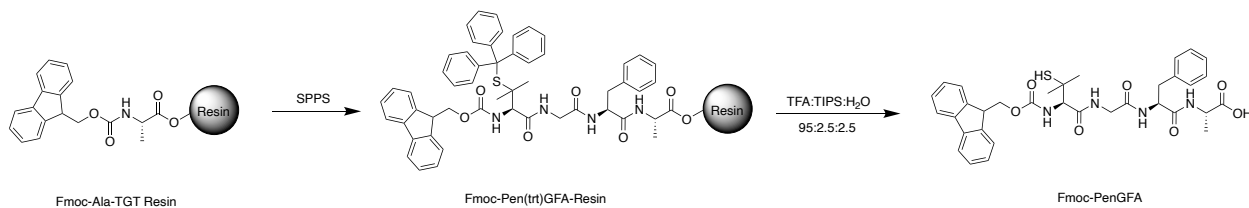


Figure 16: Solid-phase peptide synthesis & side chain deprotection and resin cleavage of Fmoc-PenGFA

Solid-phase peptide synthesis: In Figures 12, 13, 14, 15, and 16, the solid-phase peptide synthesis (machine) started with deprotection of Fmoc from Fmoc-ala-TGT resin by 2% of piperidine and 2% of DBU in DMF for 7 minutes at room temperature, then a protected-amino acid is coupled with ala-TGT

resin by HATU and DIPEA with the equivalent of 1:4:8:8 resin, amino acids, HATU, and DIPEA for 9 minutes at 40°C. The product is then washed with DMF x3 for 30 minutes. This process is repeated with other protected-amino acids (Fmoc-Gly-OH, Fmoc-Cys(Trt)-OH, Boc-Cys(Trt)-OH, Fmoc-Phe-OH, and Fmoc-Pen(Trt)-OH) in the order it's set. Then, the product is ready to obtain.

Side chain deprotection and resin cleavage: React the product obtained by the Solid-phase peptide synthesis (Fmoc-C(trt)GA-TGT Resin, Boc-C(trt)GA-TGT Resin, Fmoc-C(trt)GFA-TGT Resin, Boc-C(trt)GFA-TGT Resin, and Fmoc-Pen(trt)GFA-TGT Resin) with TFA, TIPS, and H₂O with the ratio of 95:2.5:2.5 for 20 minutes to cleave the TGT resin from the N-terminus, the Boc from the C-terminus, and the trityl side chain from penicillamine and cysteine, in figure 7, 8, 9, 10, and 11. Then, the product is blown dry with Ar gas and lyophilization with water and ACN with a ratio of 1:1. Last, analyzed by mass spectrometry.

2.2.2.5 Peptide Cyclization by Isocyanide Chemistry:

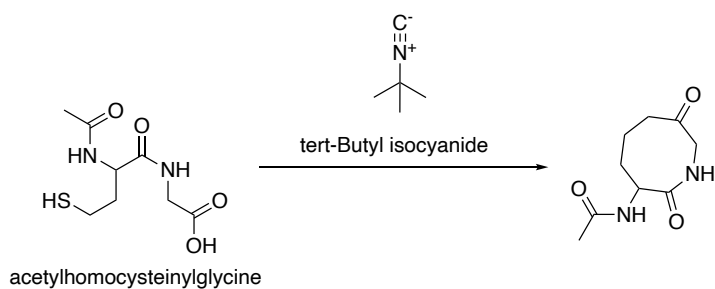


Figure 17: Peptide cyclization of acetylhomocysteinyglycine by tert-Butyl isocyanide

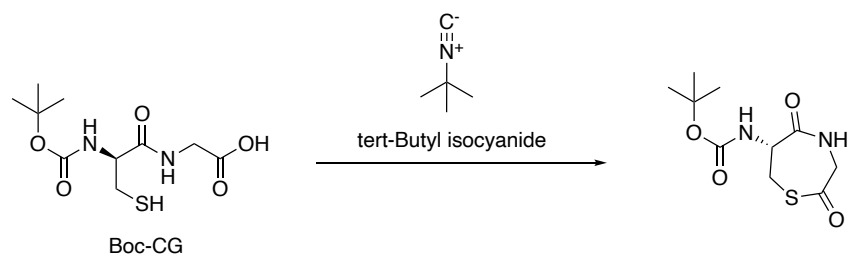


Figure 18: Peptide cyclization of Boc-CG by tert-Butyl isocyanide

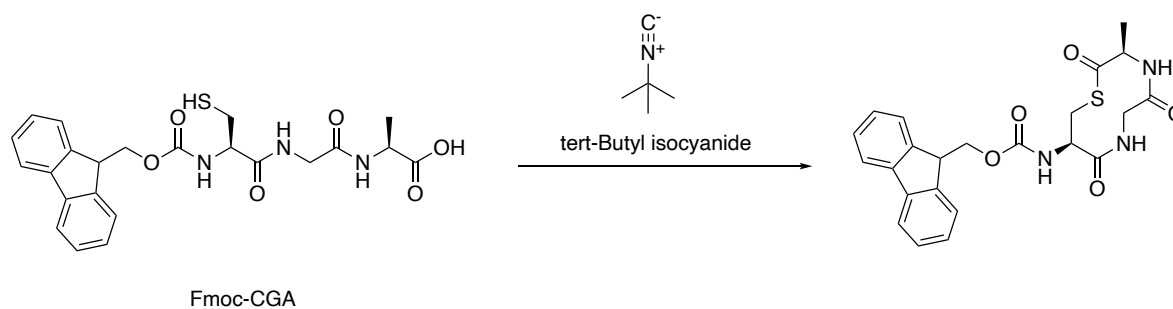


Figure 19: Peptide cyclization of Fmoc-CGA by tert-Butyl isocyanide

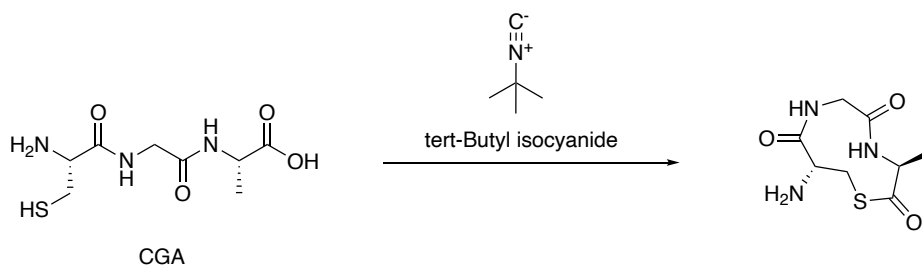


Figure 20: Peptide cyclization of CGA by tert-Butyl isocyanide

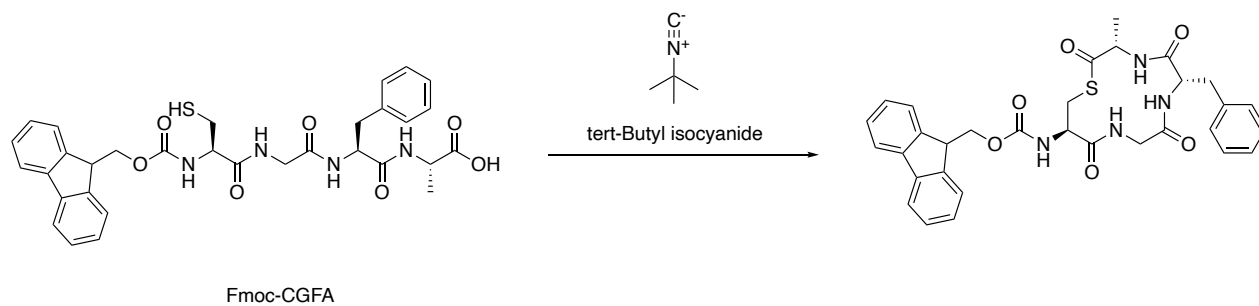


Figure 21: Peptide cyclization of Fmoc-CGFA by tert-Butyl isocyanide

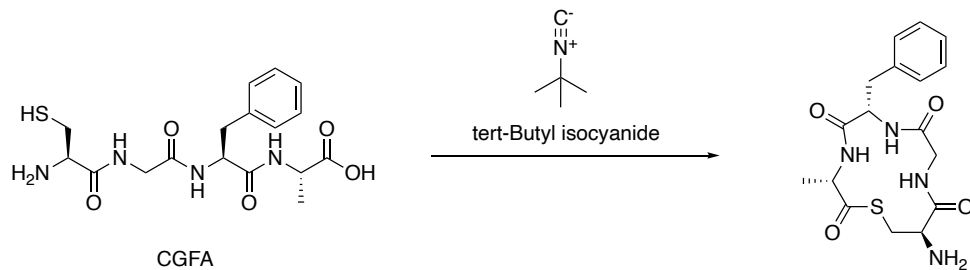


Figure 22: Peptide cyclization of CGFA by tert-Butyl isocyanide

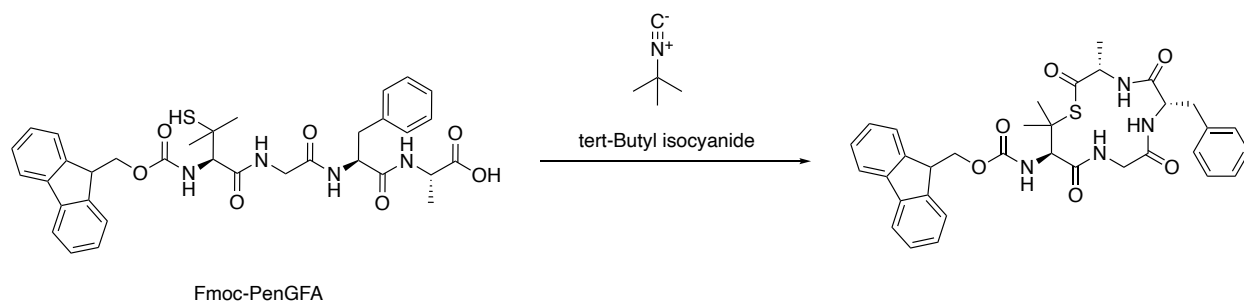


Figure 23: Peptide cyclization of Fmoc-PenGFA by tert-Butyl isocyanide

Peptide Cyclization by tert-Butyl isocyanide: In Figures 17, 18, 19, 20, 21, 22, and 23, dissolve the 1 equivalent of starting material (acetylhomocysteinylglycine, Boc-CG, Fmoc-CGA, CGA, Fmoc-CGFA, CGFA, and Fmoc-PenGFA) with 5 equivalent of tert-butyl isocyanide into 5 mmol/l of water at room temperature for 4 hours. Purify the product with HPLC (a flow rate of 10 mL · min⁻¹ of H₂O (+0.1 TFA) and ACN (+0.1 TFA); isocratic with 5% ACN (+0.1 TFA) for 5 minutes. followed by a linear gradient of 10.0% to 40% ACN (+0.1 TFA) over 20 minutes (without Fmoc) and a linear gradient of 20.0% to 80% ACN (+0.1 TFA) over 20 minutes (with Fmoc)) and analyzed by mass spectrometry

3. Results and Discussions

3.1. Synthesis of Disulfide bond

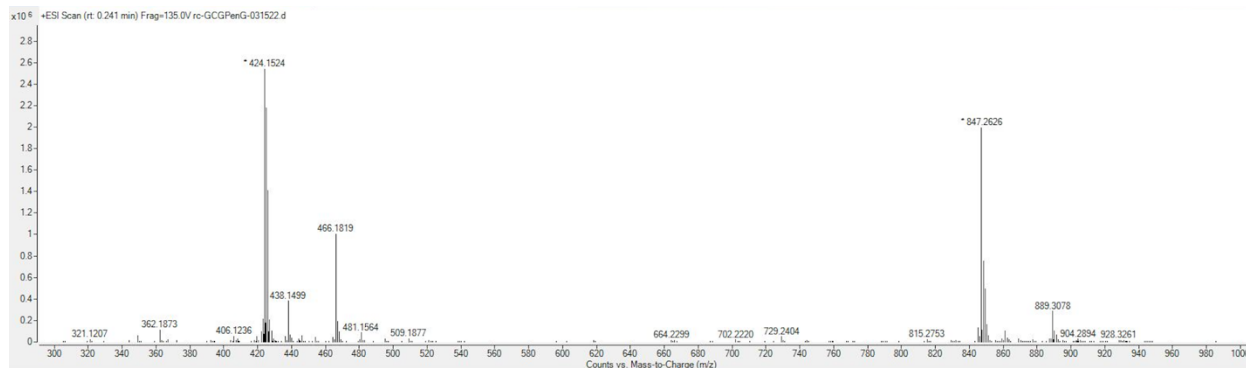


Figure 24: Mass spectrometry of peptide (GCGPenG, $[M + H]^+$: 424.13)

The calculated m/z value for GCGPenG is $[M + H]^+$: 424.13. In the mass spectrometry (figure 24), we observed a m/z value with a high peak of 424.15 at 0.241 minute, therefore we can conclude that the peptide (GCGPenG) is successfully made from solid-phase peptide synthesis.

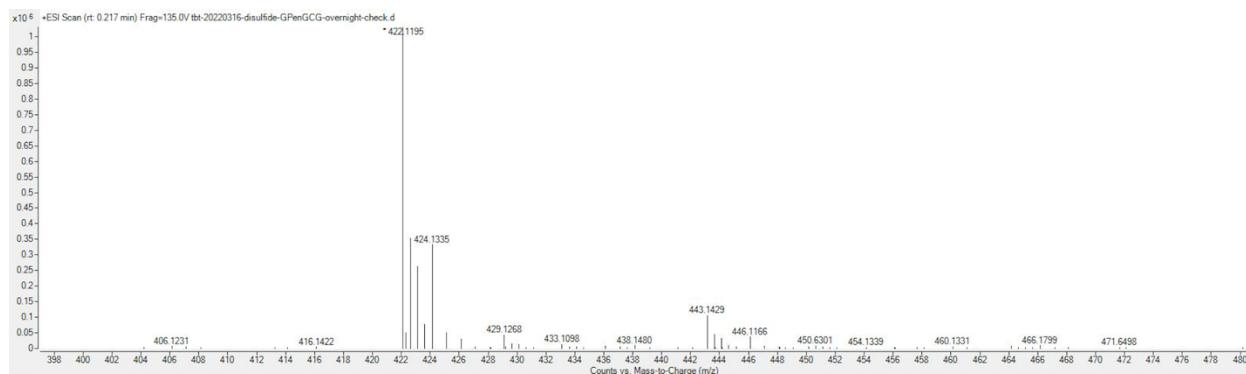


Figure 25: Mass spectrometry of intermolecular disulfide bond in two GCGPenG, $[M + H]^+$: 843.22 and $[M/2 + H]^+$: 422.11

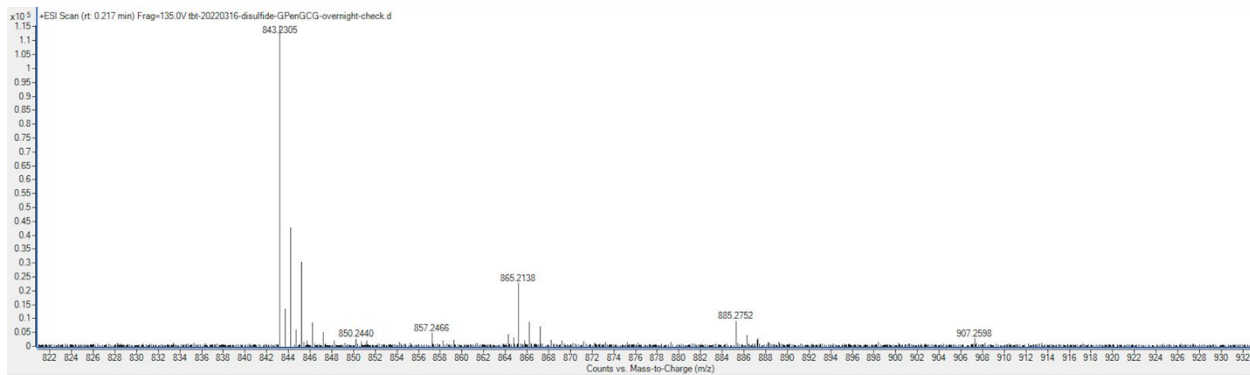


Figure 26: Mass spectrometry of intermolecular disulfide bond in two GCGPenG, $[M + H]^+$: 843.22 and $[M/2 + H]^+$: 422.11

The calculated m/z value for intermolecular disulfide bond in two GCGPenG between cysteine and penicillamine are $[M + H]^+$: 843.22 and $[M/2 + H]^+$: 422.11. After running the peptide reacted in the phosphate buffer through HPLC, we run the product with the mass spectrometry and we observed a m/z value a high peak of 422.12 (figure 25) and a m/z value with a lower peak of 843.23 (figure 26) at 0.217 minute, we conclude that the intermolecular disulfide bond in two GCGPenG between cysteine and penicillamine.

3.2. Synthesis of Protecting group-free Thiol-containing Oligopeptides

3.2.1 Cyclized peptide synthesis in the basic aqueous buffer:

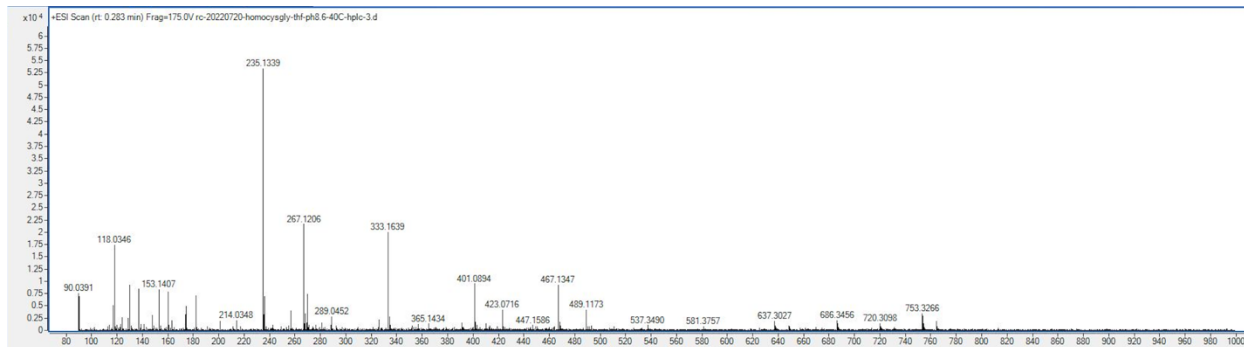


Figure 27: Mass spectrometry of peptide (acetylhomo-cysteinylglycine, $[M + H]^+$: 235.07)

The calculated m/z value for acetylhomo-cysteinylglycine is $[M + H]^+$: 235.07. In the mass spectrometry (figure 27), we observed a m/z value with a high peak of 235.13 at 0.283 minute, therefore we can conclude that the peptide (acetylhomo-cysteinylglycine) is successfully made from cyclized peptide synthesis in basic aqueous buffer.

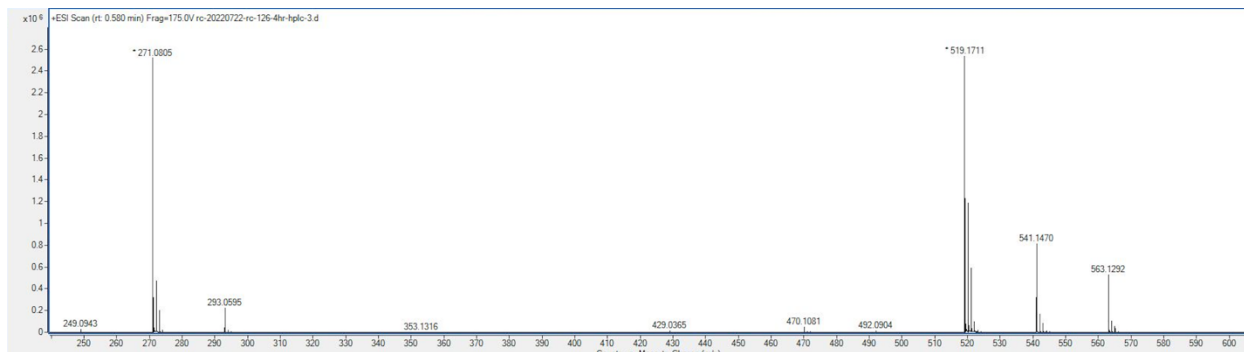


Figure 28: Mass spectrometry of peptide (acetylhomo-cysteinylalanine, $[M + Na]^+$: 271.09)

The calculated m/z value for acetylhomo-cysteinylalanine is $[M + Na]^+$: 271.09. In the mass spectrometry (figure 28), we observed a m/z value with a high peak of 271.08 at 0.580 minute and a high

peak of 519.17 is the value of $[2M + Na]^+$, therefore we can conclude that the peptide (acetylhomocysteinylalanine) is successfully made from cyclized peptide synthesis in basic aqueous buffer.

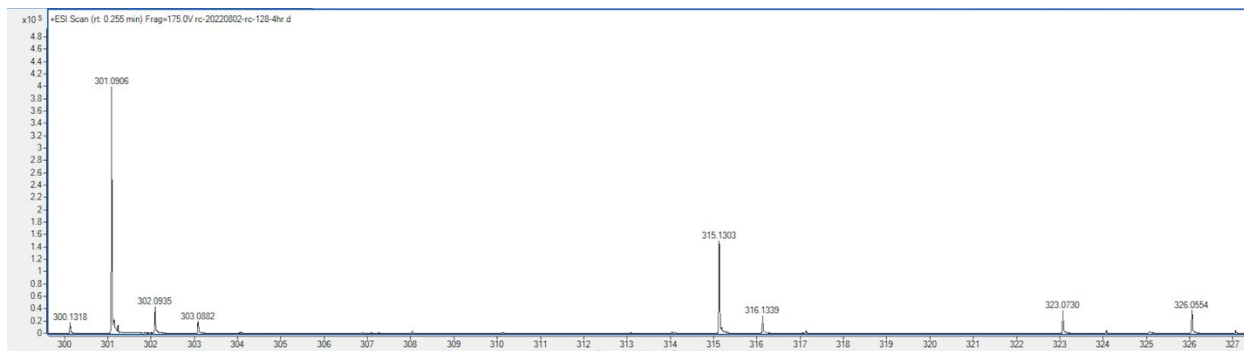


Figure 29: Mass spectrometry of peptide (Boc-cysteinylglycine, $[M + Na]^+$: 301.10)

The calculated m/z value for Boc-cysteinylglycine is $[M + Na]^+$: 301.10. In the mass spectrometry (figure 29), we observed a m/z value with a high peak of 301.09 at 0.255 minute, therefore we can conclude that the peptide (Boc-cysteinylglycine) is successfully made from cyclized peptide synthesis in basic aqueous buffer.

3.2.2 Solid-phase peptide synthesis & Side chain deprotection and resin cleavage:

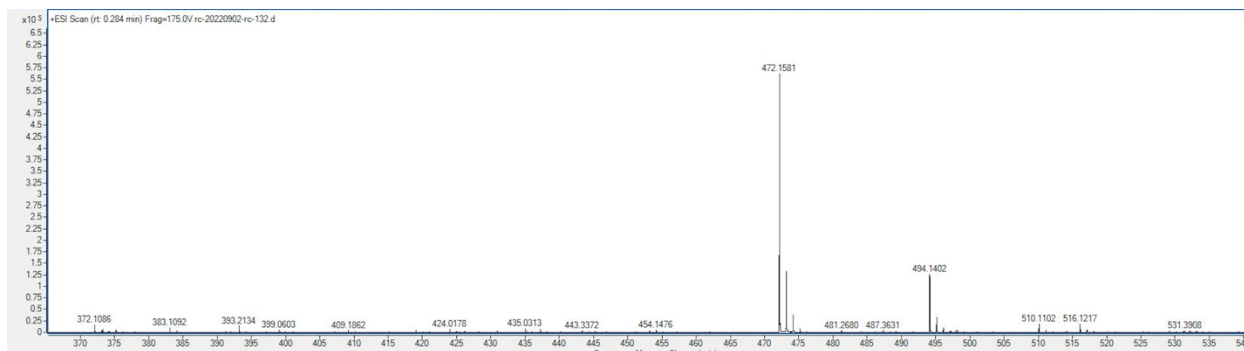


Figure 30: Mass spectrometry of peptide (Fmoc-CGA, $[M + H]^+$: 472.15)

The calculated m/z value for Fmoc-CGA is $[M + H]^+$: 472.15. In the mass spectrometry (figure 30), we observed a m/z value with a high peak of 472.15 at 0.284 minute, therefore we can conclude that the peptide (Fmoc-CGA) is successfully made from solid-phase peptide synthesis.

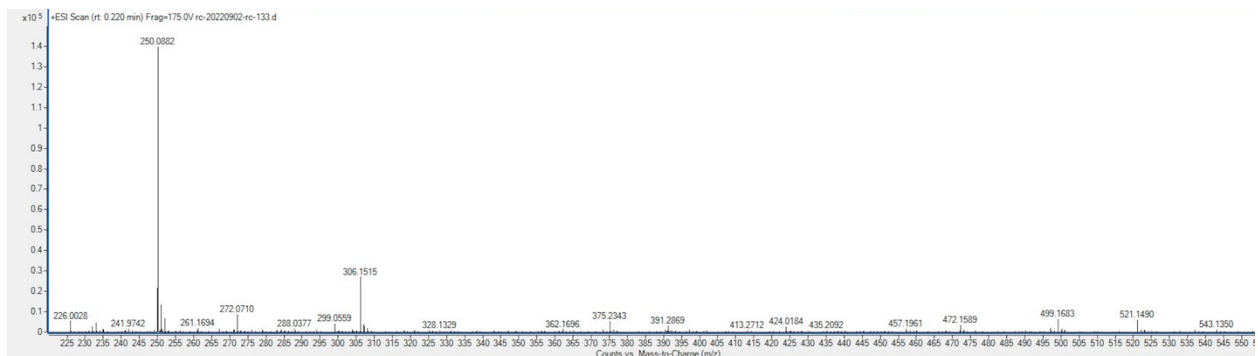


Figure 31: Mass spectrometry of peptide (CGA, $[M + H]^+$: 250.08)

The calculated m/z value for CGA is $[M + H]^+$: 250.08. In the mass spectrometry (figure 31), we observed a m/z value with a high peak of 250.08 at 0.220 minute, therefore we can conclude that the peptide (CGA) is successfully made from solid-phase peptide synthesis.

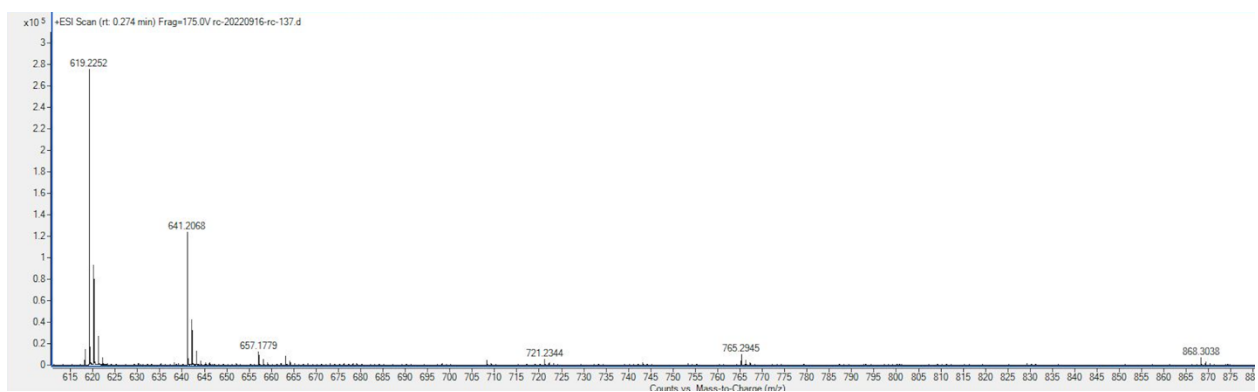


Figure 32: Mass spectrometry of peptide (Fmoc-CGFA, $[M + H]^+$: 619.22)

The calculated m/z value for Fmoc-CGFA is $[M + H]^+$: 619.22. In the mass spectrometry (figure 32), we observed a m/z value with a high peak of 619.22 at 0.274 minute, therefore we can conclude that the peptide (Fmoc-CGFA) is successfully made from solid-phase peptide synthesis.

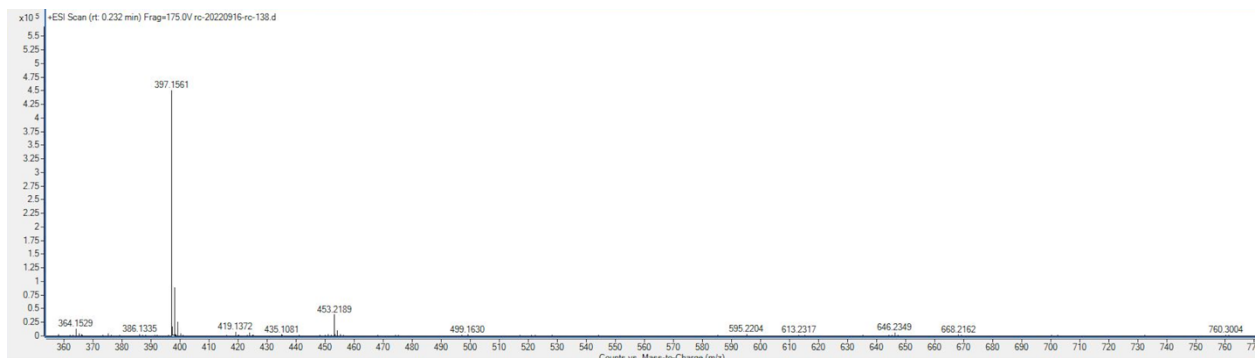


Figure 33: Mass spectrometry of peptide (CGFA, $[M + H]^+$: 397.15)

The calculated m/z value for CGFA is $[M + H]^+$: 397.15. In the mass spectrometry (figure 33), we observed a m/z value with a high peak of 397.15 at 0.232 minute, therefore we can conclude that the peptide (CGFA) is successfully made from solid-phase peptide synthesis.

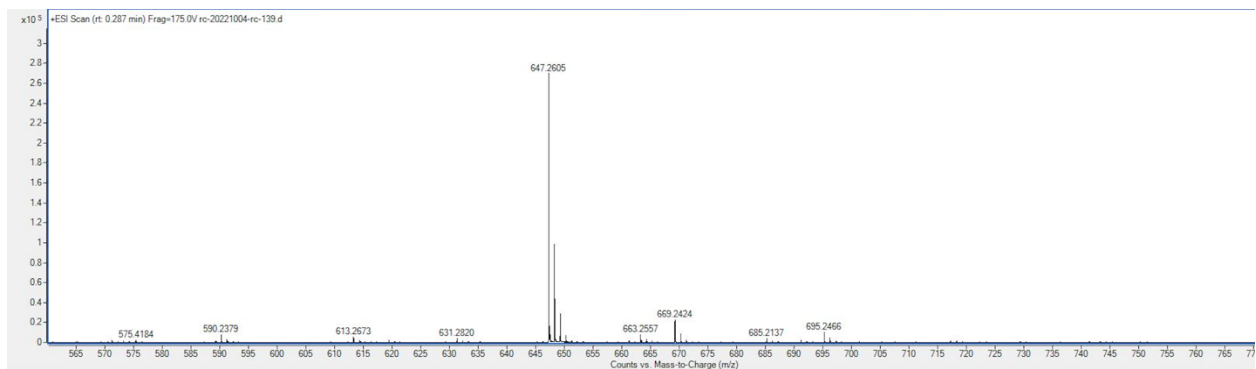


Figure 34: Mass spectrometry of peptide Fmoc-PenGFA, $[M + H]^+$: 647.25

The calculated m/z value for Fmoc-PenGFA is $[M + H]^+$: 647.25. In the mass spectrometry (figure 34), we observed a m/z value with a high peak of 647.26 at 0.287 minute, therefore we can conclude that the peptide (Fmoc-PenGFA) is successfully made from solid-phase peptide synthesis.

3.2.3 Peptide Cyclization by Isocyanide Chemistry:

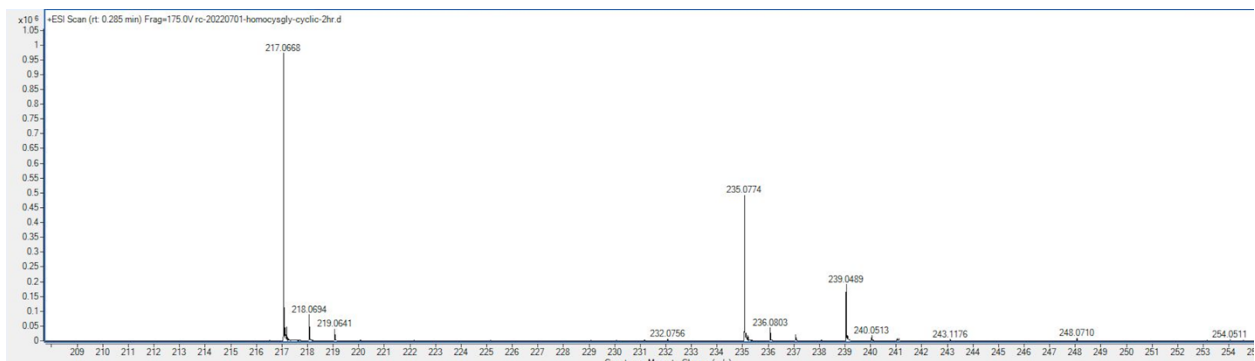


Figure 35: Mass spectrometry of cyclized acetylhomocysteinylglycine, $[M + H]^+$: 217.06

The calculated m/z value for cyclized acetylhomocysteinylglycine is $[M + H]^+$: 217.06. In the mass spectrometry (figure 35), we observed a m/z value with a high peak of 217.06 at 0.285 minute, therefore we can conclude that the cyclized acetylhomocysteinylglycine is successfully made from peptide cyclization by tert-butyl isocyanide.

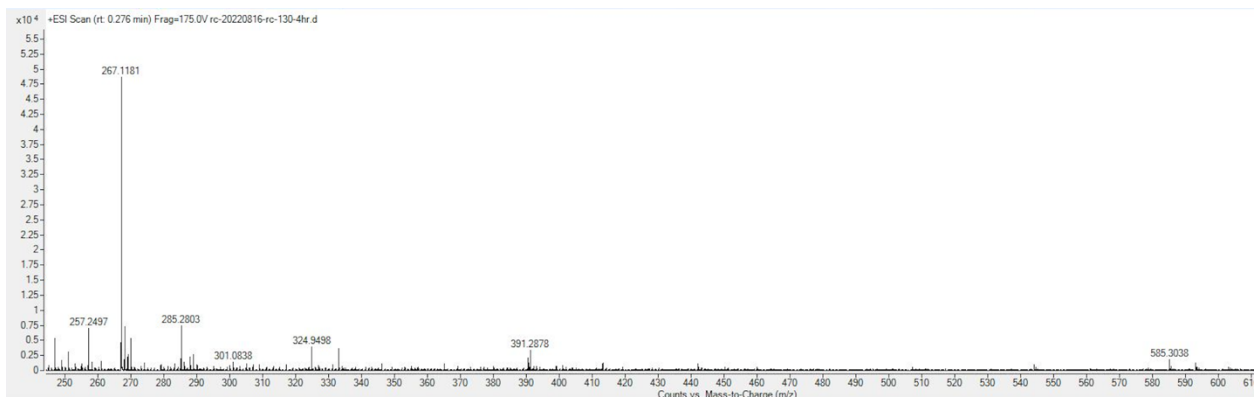


Figure 36: Mass spectrometry of cyclized Boc-CG, $[M + H]^+$: 261.09

The calculated m/z value for cyclized Boc-CG is $[M + H]^+$: 261.09. In the mass spectrometry (figure 36), we did not observe a m/z value with a high peak of 261.09, therefore we conclude that the cyclized Boc-CG is not successfully made from peptide cyclization by tert-butyl isocyanide. Due to Boc is not fully dissolved.

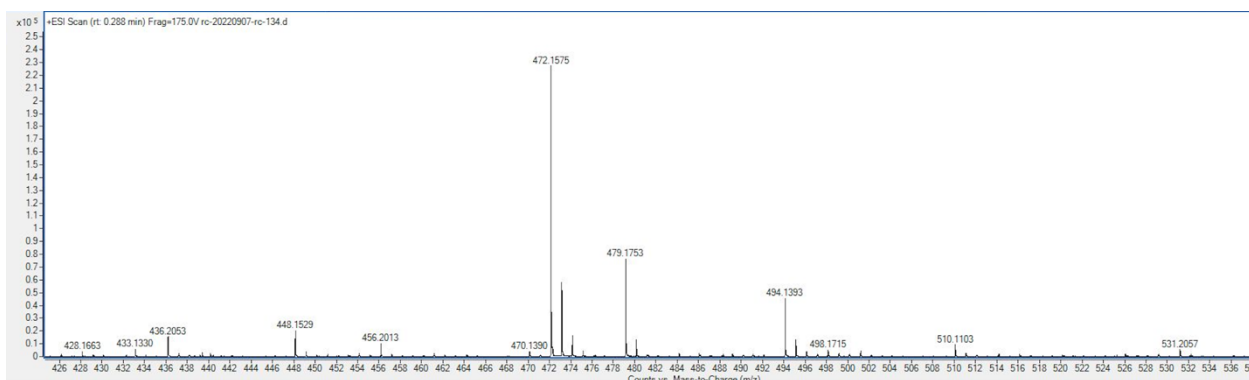


Figure 37: Mass spectrometry of cyclized Fmoc-CGA, $[M + H]^+$: 454.14

The calculated m/z value for cyclized Fmoc-CGA is $[M + H]^+$: 454.14. In the mass spectrometry (figure 37), we did not observe a m/z value with a high peak of 454.14, therefore we conclude that the cyclized Fmoc-CGA is not successfully made from peptide cyclization by tert-butyl isocyanide. Due to Fmoc is not fully dissolved.

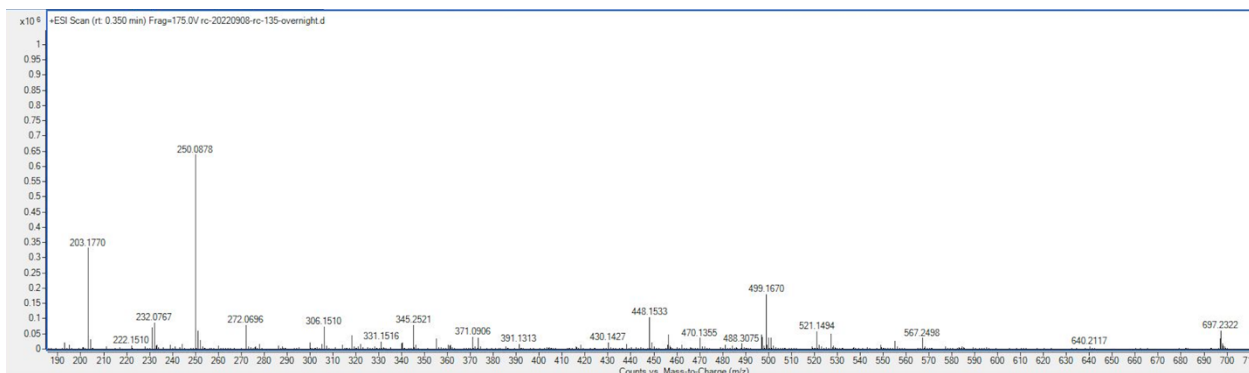


Figure 38: Mass spectrometry of cyclized CGA, $[M + H]^+$: 232.07

The calculated m/z value for cyclized CGA is $[M + H]^+$: 232.07. In the mass spectrometry (figure 38), we observed a m/z value with a peak of 232.07 at 0.35 minute, therefore we can conclude that the cyclized CGA is made from peptide cyclization by tert-butyl isocyanide, but since it is not the highest peak, purification and nmr is needed.

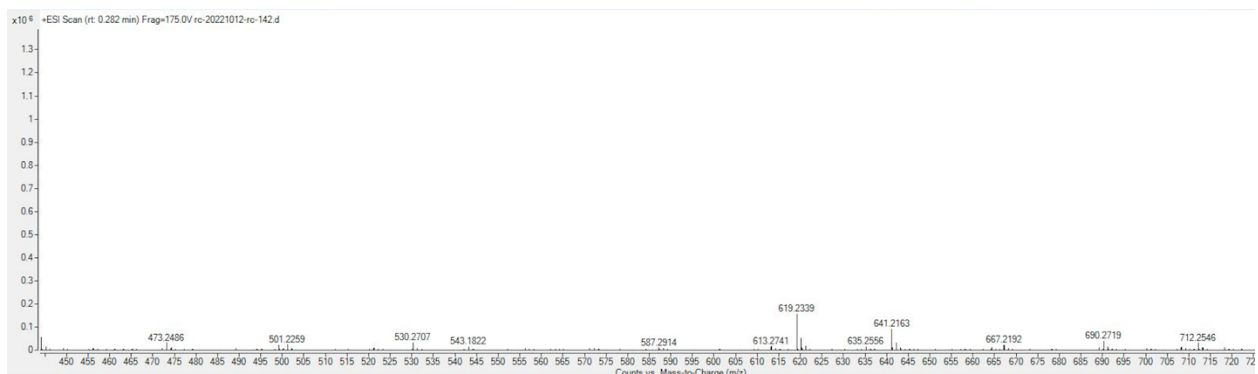


Figure 39: Mass spectrometry of cyclized Fmoc-CGFA, $[M + H]^+$: 601.21

The calculated m/z value for cyclized Fmoc-CGFA is $[M + H]^+$: 601.21. In the mass spectrometry (figure 39), we did not observe a m/z value with a high peak of 601.21, therefore we conclude that the cyclized Fmoc-CGFA is not successfully made from peptide cyclization by tert-butyl isocyanide. Due to Fmoc is not fully dissolved.

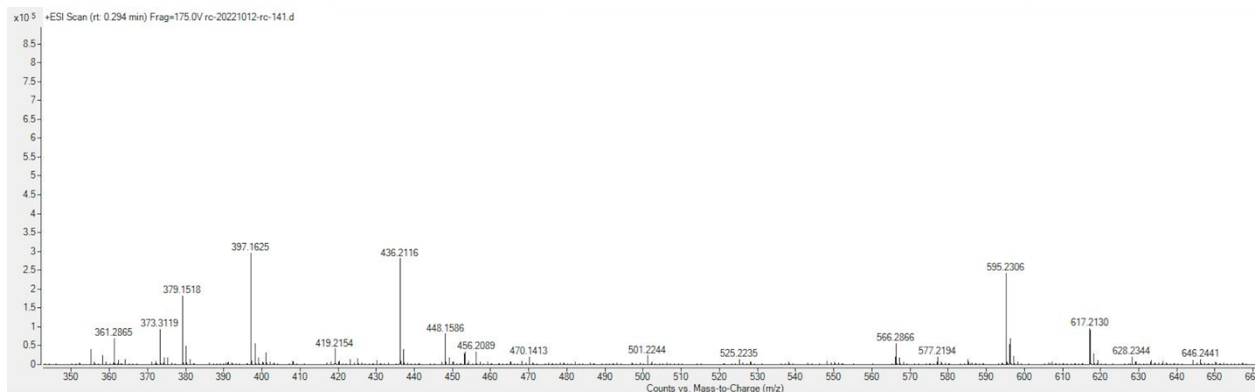


Figure 40: Mass spectrometry of cyclized CGFA, $[M + H]^+$: 379.14

The calculated m/z value for cyclized CGFA is $[M + H]^+$: 379.14. In the mass spectrometry (figure 40), we observed a m/z value with a peak of 379.15 at 0.294 minute, therefore we can conclude that the cyclized CGFA is made from peptide cyclization by tert-butyl isocyanide, but since it is not the highest peak, purification and nmr is needed.

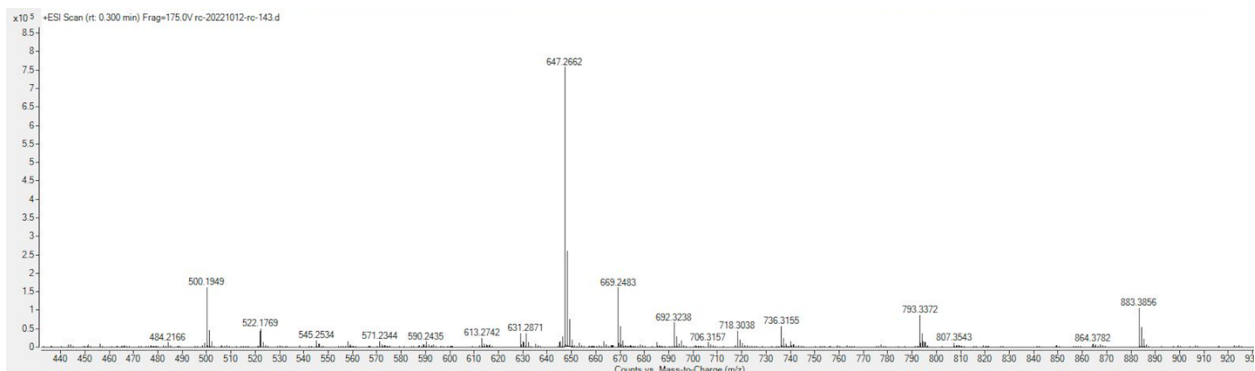


Figure 41: Mass spectrometry of cyclized Fmoc-PenGFA, $[M + H]^+$: 629.24

The calculated m/z value for cyclized Fmoc-PenGFA is $[M + H]^+$: 629.24. In the mass spectrometry (figure 41), we do not observe a m/z value with a high peak of 629.24, therefore we conclude that the cyclized Fmoc-PenGFA is not successfully made from peptide cyclization by tert-butyl isocyanide. Due to Fmoc is not fully dissolved.

4. Conclusion

In conclusion, disulfide-rich peptides have been widely studied and used in various fields, including biochemistry and medicine, due to their unique structural features and biological activities. The formation of intermolecular disulfide bonds is crucial for the precise folding of disulfide-rich peptides, which can be achieved by using molecular folding chaperones and enzymes, as well as specific chemical compounds and peptide synthesis techniques. Disulfide-rich peptides have shown potential as therapeutics for the treatment of diseases and in the development of novel biomaterials with unique properties. We speculated that the GCGPenG motif, when exposed to phosphate buffer, would not form an intramolecular disulfide bond, due to chemical properties of the motif, which consists of five amino acids, including one cysteine residues and a penicillamine residue. We noted that the steric hindrance of penicillamine could prevent the motif from forming an intramolecular disulfide bond. Penicillamine has a bulkier side chain compared to other amino acids, which can create a crowded environment around the cysteine residues, hindering their ability to form a disulfide bond within the peptide chain.[8] In addition, the short length of the peptide (penta-mer) also contributes to the inability to form a cyclic peptide, which could further constrain the formation of an intramolecular disulfide bond. Cyclic peptides, which have a closed ring structure, often have a higher propensity to form disulfide bonds due to their conformational rigidity. However, with only five amino acids, the GCGPenG motif is not long enough to form a cyclic peptide, thus limiting its ability to form an intramolecular disulfide bond. In addition, the mass spectra data exclude the intramolecular disulfide formation. The calculate m/z value for intramolecular disulfide bond is $[M + H]^+$: 423.11, and in the mass spectrometry (figure 25 and 26) we did not observed a m/z value a high peak of 423.11, therefore, we conclude that the motif will only form two intermolecular disulfide bond between cysteine's sulfur bond and penicillamine's thiol, forming an antiparallel dimerization. The disulfide bond bearing peptide is now ready for further experiments such as Raman spectroscopy. Overall, this study provides important insights into the chemical behavior of the GCGPenG motif and lays the foundation for future studies aimed at understanding its biological activity.

Moreover, the development of new, efficient methods for synthesizing peptides continues to be an important task for synthetic and medicinal chemists. Peptides have emerged as attractive therapeutic agents due to their high target selectivity, low toxicity, and potential for targeting diseases associated with protein-protein interactions. Solid-phase peptide synthesis is a well-established method for peptide synthesis, but it can be atom uneconomical due to the use of protected amino acids and coupling agents. Additionally, the protection and deprotection steps for thiol-containing peptides can be harsh and environmentally unfriendly. With the novel protocol proposed, we were able to form acetylhomocysteinylglycine, a dipeptide, from 3-Acetamidotetrahydro-2-thiophenone, an acetyl- γ -thiolactone, with a 69% yield, which is step 2 of our hypothesis (figure 3). Next, we obtained a cyclized acetylhomocysteinylglycine, a thiolactone, by tert-butyl isocyanide, but the yield is only 15%. With only a 15% yield, we did not proceed to the next step. We continue to optimize step 2 (figure 3) with different starting materials and amino acids. the acetylhomocysteinylalanine was prepared from 3-Acetamidotetrahydro-2-thiophenone, an acetyl- γ -thiolactone with only 20% yield. On the other hand, Boc-CG was produced from Boc- β -thiolactone with 14% yield. Both reactions did not work well in step 3.

Therefore, we decide to first use Solid-phase peptide synthesis to make the peptides with 3-mer, 4-mer, and 5-mer. To test whether longer peptides will be cyclized by reacting with tert-Butyl isocyanide. The results show that 3-mer, 4-mer, and 5-mer with no protecting group will be cyclized by reacting with tert-Butyl isocyanide. However, we cannot conclude that it reacts at the right position due to lack of NMR spectra evidence.. Also, the 3-mer, 4-mer, and 5-mer with Fmoc group will not cyclize by reacting with tert-Butyl isocyanide. We hypothesize that fmoc is too large and will not fully dissolve in water, therefore the peptides did not react.

This work proposes a novel strategy for the synthesis of thiol-containing peptides that could provide a more efficient and sustainable method for peptide synthesis. The development of such a protocol could have significant implications for the synthesis of peptides as therapeutic agents, enabling the production of novel peptides with unique pharmacological properties. Furthermore, this work could also have broader

implications for the field of peptide chemistry, inspiring the development of more sustainable and efficient methods for peptide synthesis. But this project is incomplete. Lastly, this project is an on-going investigation, as an alternative strategy, we propose to use 2-(tritylthio)acetic acid at the N-terminus of the peptide and observe if the peptide could be cyclized by reacting with tert-Butyl isocyanide. Steps 4, 5, and 6 (figure 3) will be evaluated in the future experiments.

5. References

- [1] Damodaran, S. (2008). Amino acids, peptides and proteins. *Fennema's food chemistry*, 4, 425-439.
- [2] Corey, R. B., & Pauling, L. (1953). Molecular models of amino acids, peptides, and proteins. *Review of Scientific Instruments*, 24(8), 621-627.
- [3] Walsh, G., & Jefferis, R. (2006). Post-translational modifications in the context of therapeutic proteins. *Nature biotechnology*, 24(10), 1241-1252.
- [4] Kent, S. B. (2017). Total chemical synthesis of proteins. *Chemical Society Reviews*, 46(17), 5347-5360.
- [5] Merrifield, R.B. (1963). Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *Journal of the American Chemical Society* 85, 2149-2154.
- [6] Fields, G.B., Noble, R.L., and Perlmutter, P. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *International Journal of Peptide and Protein Research* 35, 161-214.
- [7] Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. (1994). Synthesis of proteins by native chemical ligation. *Science* 266, 776-779.
- [8] Fass, D., & Thorpe, C. (2018). Chemistry and enzymology of disulfide cross-linking in proteins. *Chemical reviews*, 118(3), 1169-1198.
- [9] Huang, Z., Wu, Y., Dong, H., Zhao, Y., & Wu, C. (2020). Design and synthesis of disulfide-rich peptides with orthogonal disulfide pairing motifs. *The Journal of Organic Chemistry*, 85(17), 11475-11481.
- [10] Yiwu, Z., Linxiang, Z., Yibing, Z., & Chuanliu, W. (2015). Orthogonal Cysteine–Penicillamine Disulfide Pairing for Directing the Oxidative Folding of Peptides.
- [11] Jiang, S. Y., Hsieh, W. T., Chen, W. S., Liao, J. S., Chiang, P. Y., & Lin, Y. A. (2020). Synthesis of Thiol-Containing Oligopeptides via Tandem Activation of γ -Thiolactones by Silver-DABCO Pair. *Asian Journal of Organic Chemistry*, 9(10), 1638-1649.
- [12] Fields, G. B. (2001). Introduction to peptide synthesis. *Current protocols in protein science*, 26(1), 18-1.