Spectroscopic Analysis of Resin-Bound Peptides: Glutathione and FK-13

by

Michael Chan

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Chemistry

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

High-resolution magic angle spinning (HRMAS) NMR spectroscopy is used to study solid samples that are normally difficult to analyze due to broadening of peaks. Solid-phase peptide synthesis can bind peptides to an insoluble resin that can be analyzed with HRMAS NMR spectroscopy. A combination of HRMAS NMR and IRMPD spectroscopy, along with computational chemistry, was applied to analyze and evaluate the structure of resin-bound glutathione. Two-dimensional ^{1}H-^{1}H NMR experiments such as COSY, TOCSY, and ROESY were employed to assign and predict the structure of the resin-bound peptide. IRMPD results were used along with calculated protonated structures and spectra to evaluate the conformation of the peptide. The experimental spectrum was compared to the spectra and structures of the protonated species to hypothesize the most favoured structure. Molecular mechanics, molecular dynamics and DFT calculations were implemented to collect structures that best resembled the free and resin-bound glutathione peptide. The results from these methods were compared to determine the structure that is most probable for the glutathione peptide. A semi-folded conformation is the structure the resin-bound GSH most preferred as concluded from the NMR and DFT results. The IRMPD results were analyzed as separate from the resin-bound experiments and suggested protonated GSH had a folded conformation.

FK-13 was another peptide synthesized using the solid-phase peptide synthesis technique. The peptide was synthesized using a modified technique different from conventional methodology used in the past. The peptide was also analyzed using COSY, TOCSY, and ROESY to confirm that the synthesis was done correctly and hypothesize a structure. The low substitution of the peptide on the resin gave rise to minimal NOE interactions, but there was some evidence suggesting that the synthesis was successful and the peptide adopted a cyclic

conformation. These initial results are useful for future analyses and conformational studies of this resin-bound peptide.

Further work needs to be done for both peptides to explore the structures in more detail. The explicit model of solvation should be used to explore the effect of solvent molecules on the conformation of the glutathione peptide as opposed to the implicit model that PCM provides. FK-13 could be synthesized better so that a higher substitution is achieved and better NMR results are obtained. The IRMPD results obtained by the McMahon group can then be compared to the NMR results and computational calculations can be performed to obtain realistic structures of the peptide.

Acknowledgements

First I would like to thank Dr. William (Bill) Power for being a mentor to me for the past 5 years, through my undergraduate and graduate careers. I thank him for giving me the opportunity many years ago to work alongside him and his group which has allowed me to be in the position I am in today. Bill has taught me everything I know about NMR spectroscopy and has inspired me to work my hardest and give a full effort on everything I do. It is because of Bill, I knew my passion in life was with chemistry and NMR spectroscopy. He has been able to help me with all of the problems I have encountered in NMR, solid-phase and any courses I had issues with. I am extremely proud to call Bill a friend and thankful everything he has done for me in my life.

Dr. Terry McMahon and Dr. Richard (Rick) Marta for all their help with all of the IRMPD experimental procedures and interpretation of the data obtained. Namely Rick, for putting time aside from his busy days to assist me with interpretation of the IRMPD spectra and what to do with the information I obtain. With both of their help, I was able to accomplish a large goal I had set for myself in this project and am very grateful for their help.

Dr. P.N. Roy and Kevin Bishop of his group for also assisting in accomplishing another goal set out in this project. P.N.'s help and advice with the molecular dynamics aspects of this project was very helpful with starting a portion of this project. Also Kevin Bishop for providing a lot of assistance in writing the scripts for the experiments and helping me understand how to extract and interpret the data to be used in this thesis. Without their help, these foreign concepts would have been overwhelming.

Moaraj Hasan, of Dr. Scott Hopkins' group, who is not only a great scientist, but also a great friend that helped so much with this study. Mo had no obligation to this project, yet he was still willing to help with a lot of the molecular mechanics and DFT calculations. Whenever he was available, Mo would help with the many calculations I needed and allowed me to create many structures that could be used in understanding the spectroscopy data better. Mo is a great person always willing to go out of his way to help those in need even when he has a lot on his plate as well.

Jan Venne has also been a major help throughout my years while working in the Power NMR group. She has taught me so much on how to conduct an NMR experiment, set-up the probe and magnet to carry out the experiment, and how to process the data. Her help and support from the beginning has allowed me to carry out experiments with great confidence on my own. She always assists me when I am confused with problems with all aspects of NMR spectroscopy. I thank her so much for teaching me how to properly use the NMR magnet and how to solve the problems I constantly faced.

Ryan Donnelly of EMD Millipore® also played a large role in the synthesis of the peptides for these studies. His advice through emails and visit to the University of Waterloo shed light on the problems were faced during the synthesis. He then proposed a synthesis method that would possibly improve the success of the FK-13 peptide with his vast knowledge of peptide synthesis. Without his help, this project would not have been successful and his advice will be used in all future syntheses conducted by this group.

I'd like to thank my family as well, my mom, dad and brother. My parents have always supported me throughout my Waterloo career. They encouraged me to try my hardest, never give up and always seek help. Even to this day they do not understand what I do, but that does not matter because their support and encouragement pushed me to accomplish all the goals I set for myself. Also my brother, Kenny Chan, whose support I definitely needed to complete this task. He was always there to talk to and relax with whenever I got the chance to visit my home. I thank my family for everything they have done for me throughout my university career and especially during my Masters.

Finally, I also thank all of my friends throughout the eight years I have been at the University of Waterloo. All of my fellow classmates, when we entered in 2006, pushed me so hard to complete my undergraduate degree in biochemistry. Also to all of my other friends who I met outside of my program through the various clubs and events on campus, I thank them for being there for me, to help me laugh, relax and have a good time, especially towards the end of my career.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

α	Alpha
AMBER	Assisted Model Building with Energy Refinement
Arg, R	Arginine
Asp, D	Asparagine
β	Beta
Boc	*tert*-butoxycarbonyl
COSY	Correlation Spectroscopy
Cys, C	Cysteine
DCBC	2,6-dichlorobenzoyl chloride
DCM	Dichloromethane
δ	Delta
DFT	Density Functional Theory
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-(*N*,*N*-dimethylamino)pyridine
DMF	*N*,*N*-dimethylformamide
ε	Epsilon
FID	Free Induction Decay
Fmoc	9-fluorenylmethoxycarbonyl
γ	Gamma
Gln, Q	Glutamine

Glu, D	Glutamic Acid
Gly, G	Glycine
GSH	Glutathione
HBTU	*N*-[(1*H*-benzotriazol-1yl)(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HMBC	Heteronuclear Multiple-Bond Correlation Spectroscopy
HMQC	Heteronuclear Multiple Quantum Coherence
HOBt	1-hydroxybenzotriazole
HPLC	High-Performance Liquid Chromatography
HRMAS	High-Resolution Magic Angle Spinning
Ile, I	Isoleucine
IR	Infrared
IRMPD	Infrared Multiple Photon Dissociation
kJ	Kilojoules
Leu, L	Leucine
Lys, K	Lysine
µL	Microliters
mL	Millilitres
MM	Molecular Mechanics

Mmol	Millimoles
Mmt	4-methoxytrityl
MMTK	Molecular Modelling Tool Kit
MS	Mass Spectrometry
MSNT	1-(mesitylene-2-sulphonyl)-3-nitro-1*H*-1,2,4-triazole
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
O-2-PhiPr	2-phenylisopropyl ester
ODmab	4-{N-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino} benzyl ester
OH	Hydroxyl
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulphonyl
PCM	Polarizable Continuum Model
Phe, F	Phenylalanine
Ppm	Parts Per Million
r.f.	Radiofrequency
ROESY	Rotating Frame Nuclear Overhauser Effect Spectroscopy
SPPS	Solid-Phase Peptide Synthesis
TFA	Trifluoroacetic Acid
TIS	Triisopropyl Silane

TOCSY	Total Correlation Spectroscopy
Trt	Trityl
UV	Ultraviolet
Val, V	Valine

1.0 Introduction

1.1 Peptide Structure and Analysis

1.2 Solid-Phase Peptide Synthesis

1.3 Solid-Phase Peptide Synthesis Procedure

1.4 Nuclear Magnetic Resonance Spectroscopy

1.1 Peptide Structure and Analysis

Proteins are important cellular macromolecules consisting of one or more amino acid chains. Chains consisting of 40 or fewer amino acids are designated as peptides. Peptides and proteins are key molecules due to their ability to perform a variety of cellular functions, including transport, enzymatic, and structural roles. Determining the structure of peptides and proteins can provide a better understanding of the overall function of these macromolecules.

The fundamental building blocks of proteins are amino acids. There are 20 different standard amino acids each with a general chemical formula H_3N^+-CH(R)-COO^-, where 'R' is the side chain which varies for each amino acid – except for proline which has a cyclic R group connecting the amino group and α carbon. Amino acids can act as an acid or base at physiological pH due to their nature that the α amino and carboxyl group may be ionized. The state in which an amino acid α amino group bears a positive charge and the carboxyl group bears a negative charge is its zwitterionic form. The 'R' group of these amino acids are different for each of the standard 20 amino acids fall under three groups: nonpolar, uncharged polar and charged polar side chains. Individual amino acids can form a peptide bond, through a condensation reaction, between the amino group of one residue and the carboxyl group of another. With 20 different amino acids, an exponential variety of peptides can be formed depending on the number of residues in the chain.[1]

Peptide formation begins with the amino acids making up the primary structure, which is simply the sequence of the amino acids, by convention described numerically, from the N-terminus to the C-terminus. The change in geometry of the backbone of the primary peptide sequence gives rise to the secondary structure of the peptide. The peptide bond that joins two amino acids tends to exhibit a planar structure that put the 'R' groups in a *trans* conformation

with respect to one another. Due to a variety of steric interactions between the 'R' groups, the peptide as a whole will not stay planar and the torsion angles between residues will be altered. Certain 'R' groups cannot have a staggered conformation with respect to one another; therefore they adopt an eclipsed conformation as well. In addition to the neighbouring steric effects between amino acids, a variety of intramolecular forces cause the linear peptide to adopt a secondary structure. Combinations of van der Waals interactions that arise from different electrostatic situations play a critical role in peptide formation and stability. The various dipole interactions (Figure 1.1.1 a i) assist in the folding of peptides, such as between two permanent dipoles, a permanent and induced dipole, or between two induced dipoles. Permanent dipoles, such as those induced by carbonyl groups, can interact with other permanent dipoles or induced dipoles (i.e. from methyl groups in the side chains). Peptides that have a low dielectric constant in the core tend to experience the effects of dipole-dipole interactions between functional groups more strongly, thereby greatly influencing the folding of the peptide. Interactions between two permanent dipoles can vary in their effects (Figure 1.1.1 a i); they may repel or attract one another depending on the functional group and its orientation. Interactions between a permanent dipole and an induced dipole will always result in an attraction between the groups (Figure 1.1.1 a ii). The circulation and flow of electrons dictates whether there is an attraction or repulsion between induced dipoles. When the electrons of one functional group distort the electrons of another, an attractive force is felt, but if the group cannot induce a change in the electron flow, repulsion is experienced. The interactions between two induced dipoles (Figure 1.1.1 a iii) are very weak on their own, giving rise to London dispersion forces (such as two hydrogen atoms of the same methyl group interacting), which do not influence peptide folding very much in small numbers, but if the peptide has many of these London forces close together, they can

influence the peptide folding in some manner. London forces occur between two induced dipoles and create an attractive interaction between the two groups due to the instantaneous shift of electrons between two induced dipoles.[2]

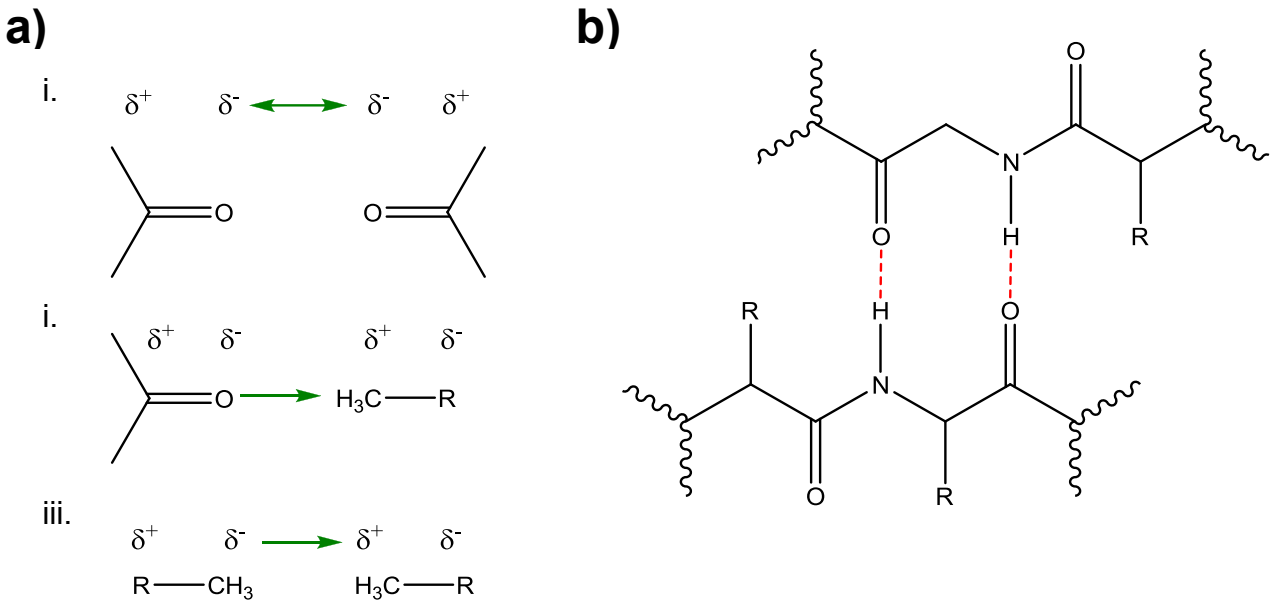


Figure 1.1.1: Amino acids that make up primary sequences can fold into secondary structures through a variety of stabilizing forces such as a) van der Waals forces and b) hydrogen bonds. A variety of dipole interactions can influence how peptide can fold a) i. Two permanent dipoles can influence one another to either cause an attraction or repulsion (in this case a repulsion is shown), ii. Interactions between a functional group with a permanent dipole can induce a dipole on a different group creating an attraction, iii. When two groups that do not have permanent dipoles are in close contact, the flow of electrons from those groups can create a London dispersion force, creating an attraction between the groups. B) Hydrogen bonding (dashed red lines) can occur between a hydrogen bond donor (in this case the amide proton of the backbone) with a hydrogen bond acceptor (a carbonyl group of a different segment of the peptide or different chain), and create a force that weakly stabilizes secondary structures, but in large numbers, hydrogen bonds play a key role in stabilizing secondary structure.

Another key interaction that dictates peptide folding is the interaction between a hydrogen donor and a hydrogen acceptor, commonly called hydrogen bonds (Figure 1.1.1 b). A hydrogen bond donor (such as from an –OH, $-NH_2$, or even $-CH_3$) creates an electrostatic interaction with a hydrogen bond acceptor (i.e. atoms that have a free lone pair of electrons contributed by carbonyls, alcohols, amines, etc…), forming strong interactions that are weaker than covalent bonds. In folded peptides and proteins, hydrogen bond donors are not exclusively restricted to interacting with one other hydrogen bond acceptor; in many cases the hydrogen will interact with multiple donors, such as the backbone carbonyl of an adjacent residue and another one that is several residues away. On the other hand, peptides in their unfolded state do not form as many hydrogen bonds with other residues and the hydrogen atoms will interact with the surrounding aqueous environment. As a result, hydrogen bonds in an unfolded peptide are very weak or do not exist, though in folded conformations, hydrogen bonds greatly influence the structure of the peptide. In the peptide's folded state, hydrogen bonds that associate tend to occur in clusters along the sequence, meaning there will be sections of the folded peptide where many interactions give rise to certain structures, but other portions with no hydrogen bonding. For example, some larger peptides and especially proteins will contain multiple secondary structures (helices, sheets and turns), and in these sections there are many hydrogen bond interactions that lay the foundation for such structural features.[2]

Hydrophobic effects also dictate peptide folding based on the interactions of non-polar residues with aqueous environments. The driving force behind non-polar groups excluding polar groups (such as water) is largely enthalpic (i.e. it is more favourable for the non-polar group to dissociate in a non-polar solvent than an aqueous one) and also is entropically driven, whereas a non-polar medium is relocated to a non-polar environment. In the case of a single non-polar

residue, when it tries to disrupt the well-ordered matrix of an aqueous environment, the water molecules will reorient themselves around the non-polar chain as it cannot form any sort of hydrogen bond with the side-chain. This same example can also be applied to larger systems, such as peptides and proteins, where a large group of non-polar side chains will be surrounded by the water molecules creating a hydrophobic cavity. A large number of these hydrophobic interactions caused by the amino acid residues play a large role in dictating the folding of peptides and which side chains are localized within the protein to avoid the aqueous environments. One other specific interaction that contributes to peptide folding and stability is the disulphide bridge. With regards to peptides and proteins, these interactions only occur once the peptide has folded into its native state and only readily occur between the sulphur atoms of cysteine residues.[2]

Many types of secondary structures exist, but some of the most predominant structures are helices, β-sheets, coils and loops (Figure 1.2.1). As the peptide backbone begins to twist the α carbons in a uniform matter, it gives rise to a helical conformation. Depending on the residues, the conformation of the helix can vary in the number of amino acids per turn, the distance between each turn along the axis, and can turn to the left or right. The helix can only form if certain angles in each turn are sterically favourable, stable and can be held together, i.e. stabilized with hydrogen bonds. One secondary structure that is very stable in a helical form is the α-helix, a right-handed helix that is stabilized by hydrogen bonds with 3.6 amino acids per turn, and torsion angles of $\Phi = -57\ ^\circ$ and $\Psi = -47^\circ$. The stabilizing force from the hydrogen bonding comes from the amide proton of residue *n* interacting with a carbonyl group of residue *n*-4. Along with hydrogen bonding, the R groups of each amino acid point out from the centre of the helix, in order to avoid steric complications, and the core is tightly packed so the atoms

are within van der Waals contact with one another. There are a variety of helices that make up this group of secondary structures, but the α-helix is the most stable of all the helices.[2]

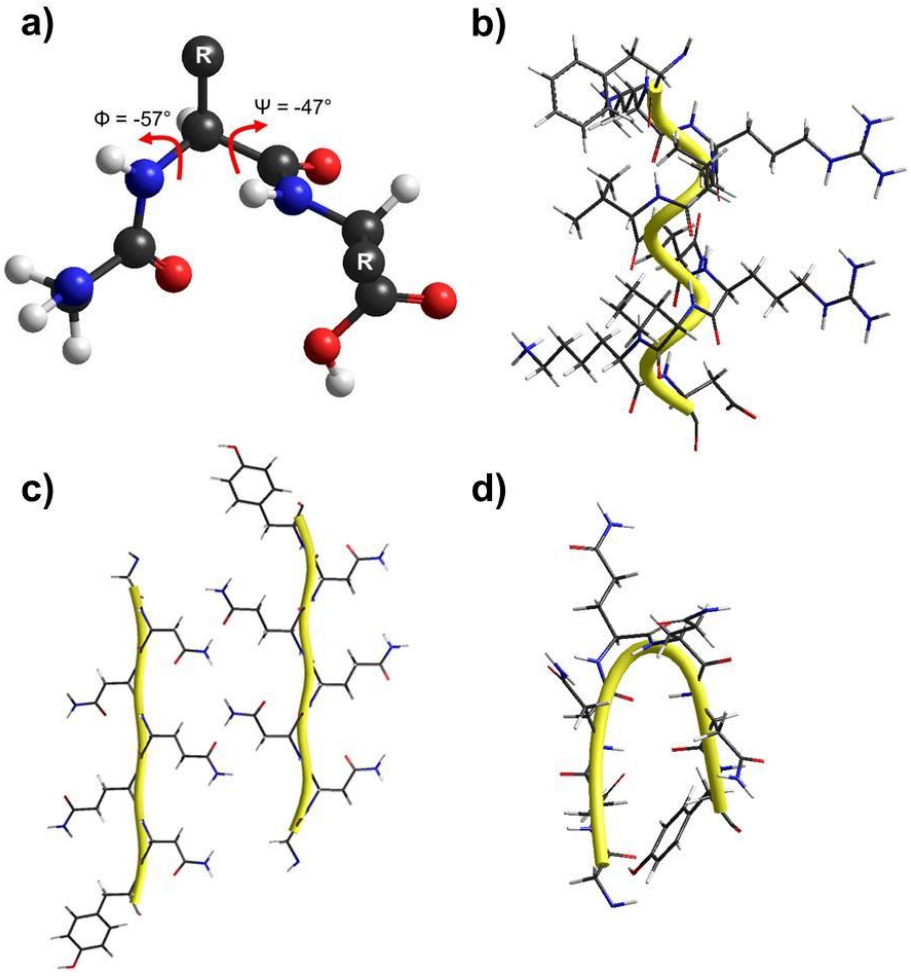


Figure 1.1.2: Various types of secondary structures that primary sequences tend to adopt –a) Φ and Ψ torsion angles in an α-helix b) α-helices, c) β-sheets, and d) β-turns are the most common types of secondary structures. These structures are stabilized by a variety of electrostatic forces, hydrogen bonds and hydrophobic effects. Smaller peptides will fold into a secondary structure and in larger proteins multiple secondary structures can comprise its entirety.

A second group of secondary structures is the β pleated sheet, which has Φ and Ψ angles that give the peptide a sheet-like structure. These structures are also stabilized by hydrogen bonds, but they usually occur between two monomers of the sequence as opposed to internally (as with the α-helix). These sheets can adopt two forms: parallel and anti-parallel. In a parallel formation, the two sheets run in the same direction (i.e. the C-termini for both chains are on the same side and run towards the N-termini) and the sheets are stabilized by hydrogen bonds across the chain. In the anti-parallel conformer, the direction of the C- to N-terminus is opposite for the two chains, but they are still stabilized by hydrogen bonds between the amide protons and carbonyl groups of the backbone. Many of these monomeric sheets can continuously form hydrogen bonds with one another to create an aggregate of β-sheets, and can occur between parallel or anti-parallel sheets. It has been seen that aggregates formed from parallel sheets are less stable when compared to sheets formed from anti-parallel peptides.[2] In larger proteins, these β-sheets on the same chain can be connected through other peptide chains, such as helices or turns. Furthermore, sheets tend to have a natural right hand curve (similar to a helix turn) and this can cause the sheet to loop around and be stabilized through hydrogen bonding.[3]

Helices and sheets make up approximately half of the secondary structures in fibrous proteins, and the remaining structures are non-repetitive structures such as coils, loops, and turns. These structures are just as organized as helices and sheets, but their conformations are more random. β-turns have four amino acids that comprise the turn and can be stabilized by hydrogen bonds between the backbone residues. Another common non-repetitive structure is the loop, which is usually 6 to 16 residues and almost always exists in larger proteins. Even though this group of structures is less common than the helices and turns, they usually are on the surface of proteins, meaning they play a large role in the overall biological function.[3]

Sequential and structural determination of peptides is a key step in the understanding of the function of the peptide and the structure of the protein it is a part of. Traditionally, peptide chain analysis was done by hydrolyzing the peptide bond between amino acids using an acid, base, or peptidase. The use of peptidases over acids and bases was preferred since different enzymes were more specific and would cleave a particular peptide bond between certain amino acids, whereas the acids or bases will attempt to break all the peptide bonds in the chain without any specificity. This method of analysis has been replaced with automated amino acid analyzers. The peptide can be broken down before analysis and observed with a combination of HPLC and UV spectrophotometry. When the amino acids are placed into the HPLC column they will elute at different rates; as they are broken down they can form a fluorescent adduct with 2-mercaptoethanol and *o*-Phthaladehyde, and will give a unique absorbance for each amino acid when measured by UV spectrophotometry. This breakdown does not give an exact indication of the structure of the protein as a whole, but by understanding the sequence of the peptide being analyzed, conclusions about the nature of the peptide can be drawn. For example, if many non-polar amino acids are observed in the peptide analysis, it suggests that this portion of the peptide may exist as part of the hydrophobic core of a protein. The methods outlined here can be used to analyze fragments and sequence of the peptides with the aid of HPLC and UV spectrophotometry.[4]

One of the first techniques used to establish a protein structure was X-ray crystallography. This method requires that the sample be crystallized, then a beam of X-rays bombards the crystal to produce a diffraction pattern. The resulting film (or digital image, in modern diffractometers) of the X-ray pattern shows the interactions between the X-ray and electrons producing dark spots on the film, which are indicative of areas where electron density

is concentrated, usually atomic centres. Traditionally complex mathematical functions were used to convert the diffraction pattern into a crystal structure of the sample, but this has been simplified with the aid of computers to generate the three-dimensional crystal structure. Proteins have been analyzed on numerous occasions using X-ray crystallography, but the technique is limited by the size and resolution of the crystal that is obtained for the sample. Protein crystals tend to be composed mostly of water which makes the overall crystal less ordered and the resulting diffraction pattern uninformative. Furthermore, the side chains of the individual residues are difficult to map since some of them cannot be differentiated as they have approximately the same size and shape, but fortunately the backbone of the protein can be visualized, allowing for the orientations of the side chains to be determined. Some of these issues can be overcome by knowing the sequence of the protein beforehand and, with the aid of mathematical functions, the crystal structure of the peptide can be accurately established. The accuracy of X-ray structures to the crystalline native structure of the protein is often a very good representation of that protein in solution. Most proteins normally exist in an aqueous environment inside cells and the protein crystal that is formed is mostly made up of water. Upon crystallization, these proteins tend to exist in their most native state just as in many enzymatic systems. The proteins will adopt the conformation that still makes them functional in the crystal, thereby making the X-ray pattern a good representation of what it looks like in a solution state.[3]

Mass spectrometry (MS) has also evolved into an important tool in the analysis of peptides and proteins. An ionization source fragments the peptide into ions that are passed through a vacuum into a mass analyzer and then measured by the mass spectrometer. Traditional methods of ionization in MS include electron impact and chemical ionization, but these methods were poor as they were prone to completely destroying the peptide. Fast atom

bombardment (FAB) was introduced as a method, where the sample typically is bombarded with xenon or argon atoms. This method usually gives back a molecular weight of the sample, but works best below a 7 kDa limit. Electrospray ionization is another very common method used in the analysis of peptides by MS. The sample is introduced as a solution and after interacting with a nebulizing gas, it exits as an aerosol of charged droplets. Eventually, the droplets will evaporate, explode and release the ions into the vapour phase to be detected by the mass analyzer. Before passing into the detector, the ions generated must be separated so they can be read individually. Ion traps are just one of the many types of mass analyzers available for this technique; these in particular are very useful for peptide analysis and are used in the IRMPD technique (to be discussed later). Fragmented ions that make it into the ion trap are held until being ejected into the detector. These traps can be selective for ions of specific mass-to-charge (m/z) by controlling the voltage so that one ion, or even many ions from the same parent molecule, can be detected. Tandem mass spectrometry (two mass spectrometers that work in series together) is another useful method of trapping ions for analyzing peptides. After fragmentation, a single ion of the parent molecule is isolated in the first mass spectrometer. This lone peptide fragment is isolated from other fragments of differing m/z ratios and contaminants. The fragment interacts with a collision cell where the fragment can break about any of its peptide bonds, and the individual fragments will be measured by the second mass spectrometer. As each ion is measured in the mass spectrometer, one can compare the sizes of each fragment and eventually deduce the sequence of the peptide.[5]

Infrared multiple photo dissociation (IRMPD) spectroscopy is a recently developed technique that has evolved into powerful tool for peptide structure determination. This technique works by irradiating a molecule with a laser at a specific frequency inducing

fragmentation of the peptide. If the molecule vibrates at the same frequency as the irradiating laser, the molecule will absorb as many photons as possible before it can no longer hold any more internal energy and dissociates. The fragment of the peptide is tracked as a change in the m/z ratio measured using a mass spectrometer. This technique also produces an IR spectrum showing the frequencies of the fragments produced after dissociation. IRMPD spectroscopy is a combination of mass spectrometry and IR spectroscopy; different sources of ionization can be employed to dissociate the peptides and a variety of ion traps can be used for differing times for laser and sample interactions. The different types of lasers that are available are, but not limited to, carbon monoxide, free electron (FEL) and optical parametric oscillator (OPO) lasers. Depending on the frequency range that is to be observed and the radiation power available, the choice of laser can vary from sample to sample. This technique has a mass spectrometer component to it, so different types of ion traps are available. The most common ion traps for analysis are Penning and Paul-type traps, each with its own pros and cons. The Penning trap allows for longer irradiation periods, producing more ions. The disadvantage to this type of trap is that the set-up of this apparatus focuses the laser on at certain point as opposed to irradiating the entire ion cloud. The Paul-type trap improves where the Penning trap falters in the sense that the entire ion cloud is irradiated by the laser. In a general sense, the Paul-type traps seem to be more efficient than the Penning traps. Unfortunately, the Paul-type trap is not as versatile as the Penning trap for many experiments. IRMPD spectroscopy is a newly established technique, but has great implication for the study of peptide structure.[6]

One of the most powerful tools in protein and peptide structure determination is nuclear magnetic resonance (NMR) spectroscopy. The theory of NMR spectroscopy is to be covered in detail in section 1.6, but here the application to structure determination of peptides and proteins

will be discussed briefly. By placing a sample into a magnetic field and irradiating it with a radio-frequency pulse, the magnetic nuclei will be excited and will generate a signal that is its own electromagnetic radiation. When an electronic cloud of a nucleus experiences a change from the applied magnetic field, it gives rise to differing internal magnetic fields and an overall change in the magnetic field. This difference from the applied magnetic field and the effects felt by the nuclei are tracked as a change in the chemical shift. The most common nuclear isotopes possessing magnetic moments that are analyzed include hydrogen-1 (proton), carbon-13, nitrogen-15, phosphorus-31, and fluorine-19 isotopes (just to name a few). With regards to peptides, ^{1}H, ^{13}C and $^{15}N$are the most commonly observed nuclei as these elements occur frequently in peptides and proteins. By determining the frequency at which the protons resonate for a particular peptide sequence, the connectivity of the atoms can be determined within certain residues. Two-dimensional NMR experiments like COSY and TOCSY can establish the connectivity of protons within a peptide; the difference between the two is that COSY detects protons that are 3 bonds or less from one another, while TOCSY can detect interactions between protons that are within the same spin system. The COSY and TOCSY experiments usually identify the sequence of the peptide or protein being analyzed and can assign a frequency to all the protons in the system. Experiments that utilize NOE enhancement, such as NOESY or ROESY, can be used to determine the structure of the sample. These experiments establish through-space interactions between protons (usually restricted to 5 Å or less). By compiling a list of through-space interactions, the three-dimensional structure of the peptide or protein can be pieced together. In addition to proton-proton two-dimensional NMR experiments, heteronuclear experiments such as HMQC and HMBC are useful for assigning frequencies to heteroatoms and establishing structure. These experiments can be used to determine the

frequency at which a nucleus (not proton) resonates, and correlates it to a proton that is bound or nearby. For example, in an HMQC experiment, the magnetization of a heteronucleus, carbon or nitrogen being the most common, is exploited by using a proton that is directly bound to it and produces a signal for this interaction. This technique is especially useful for distinguishing between protons that may have the same chemical shift by correlating them to the different carbon or nitrogen atoms they are bound to. It may also resolve two protons that are chemical shift inequivalent that are bound to the same carbon. HMBC experiments use the through bond *J*-coupling between the heteronucleus and protons, i.e., protons that are usually three bonds or less away from a carbon will produce a signal. This technique is quite useful for establishing the structure of the peptide or protein since it may support some of the NOE interactions and eliminate any ambiguity in assignment. There are a variety of 2D NMR experiments that can be used to determine the three-dimensional structure of a peptide or protein, but analysis is not limited to 2D experiments as 3D and even 4D experiments are becoming increasingly more popular to aid in this determination. These 3D and 4D experiments just add another nucleus to the experiment and the analysis involves looking at another axis on the spectrum. By combining the different homo- and heteronuclear NMR techniques, the NMR structure of a peptide can be hypothesized based on frequency of the nuclei, through-bond connectivity and through-space interactions.[7]

With all of these tools available to hypothesize a structure based on the evidence collected from these experiments, the final step would be to generate a visual representation of the peptide computationally. When creating a variety of structures that can represent the protein from the data that was acquired, structural constraints need to be applied so that the electronic structure does not violate any physical laws. Torsion angles need to be restricted so that when

sampling different structures, atoms do not rotate on top of other atoms or bring them too close in contact, which would be sterically or electrostatically unfavourable. These constraints also need to be applied when evaluating the hydrogen bond distances in these proteins. The next step to creating the input structure is to include all of the connectivity information and NOE interactions that were gathered from the various NMR experiments. These interactions will help provide an initial structure to work from as opposed to randomly sampling different structures from a straight chain. Once all the physical and NOE constraints are in place and an initial structure is generated, it can be optimized with the aid of a variety of molecular modelling programs. Each program that can be used for calculating protein structures is unique and uses different algorithms to approach one or multiple energetically and geometrically optimized structures. The results of calculations can be evaluated based on a multitude of criteria, namely root mean square deviation (RMSD) energy, and how well it complements the X-ray, MS, IRMPD and NMR data. If multiple ideal structures are generated, which is not uncommon for peptides as they can experience a variety of stable conformers, the most favourable one can be chosen by analyzing the aforementioned criteria. Calculated structures that tend to have the lowest RMSD value and energy of all the projected conformers will be the favourable structures, and those with the lowest values will most likely be the dominating native structure of the peptide or protein. These results can then be compared to the various spectroscopic results obtained to verify the most favourable structure. Molecular modelling of peptides and proteins is the final step in a long process of determining peptide structure from the initial step of the simple peptide sequence. All of these tools combine to better understand the structure of peptides and proteins and ultimately can be used to aid in the understanding of their functions in a larger biological system.

1.2 Solid-Phase Peptide Synthesis (SPPS)

Over 50 years ago, synthesis of short biologically active peptides was not a difficult task, but the approach used to make these short peptides did not lend itself well to the synthesis of longer peptides. Using theory based on pre-existing peptide synthesis techniques and chromatography theory, Robert Bruce Merrifield developed a technique in 1963 to synthesize and purify longer peptides in high yields. His method began by covalently linking an amino acid to an insoluble polymer resin matrix that contained a functionality to which the first amino acid could be linked. Merrifield's choice of polymer was a chloromethylated copolymer mix of styrene and divinylbenzene, which had a mesh-like microporous structure that allowed reagents and solvents to pass through it. The attachment of the first of four amino acids in the tetrapeptide was done by adding a carbobenzoxy protected amino acid to the chloromethylated functional group in a dry solvent.[8] His procedure then required that any of the unreacted sites were esterified using triethylammonium acetate. The cleavage of the carbobenzoxy protecting group on the first amino acid was carried out using 10% HBr in acetic acid to give a free amino terminus. The addition of the next amino acid for Merrifield's synthesis utilized the *N,N'*-dicyclohexylcarbodiimide method described by Sheehan[9], which required an excess amount of the carbobenzoxy protected amino acid, DMF and diimide. Again, the amino acid was deprotected and all excess reagents were removed with acetic anhydride and triethylamine. This process of deprotecting and coupling a protected amino acid was one cycle of the process Merrifield described and was repeated three more times to complete the synthesis. The product was then cleaved from the resin, purified and isolated from incomplete chains using chromatographic methods. The next year, Merrifield demonstrated the ability of his technique in the synthesis of a nonapeptide, bradykinin, but used *t*-butyloxycarbonyl (Boc) protected amino

acids instead of the carbobenzoxy amino acids. This allowed for milder conditions for deprotection and purification throughout the synthesis. Using the method that he had developed previously, the nonapeptide was synthesized on resin and when cleaved displayed normal biological activity of the natural bradykinin peptide.[10] Merrifield's Boc method was a staple of peptide synthesis for about 8 years until Carpino introduced the possibility of using the 9-fluorenylmethyloxycarbonyl (Fmoc) group to protect amino acids which required milder conditions to remove. Their study utilized methods that were used in Merrifield's study, but instead employed the Fmoc protecting group on the amino acids to synthesize a dipeptide.[11] The Fmoc protecting group only required that a secondary amine, such as piperidine, be used to cleave the group from the amino acid to create a dibenzoylfulvene adduct. As the field of peptide synthesis progressed, so did the understanding of the benefits of Fmoc peptide synthesis over Boc methods. A major leap forward in the field of peptide synthesis and Fmoc synthesis was the jump to automated peptide synthesis using Merrifield's techniques and Carpino's implementation of Fmoc chemistry. Traditionally, the synthesis was carried out manually using a Merrifield reaction vessel, shakers and valves to add or remove solutions. Manual synthesis was carried out using continuous flow set ups, where the reaction vessel would have the reagents of choice and solvents continuously pumped through the vessel and reacting with the resin. As the reagents react with the resin, it is continuously monitored by a UV spectrophotometer to determine the molecular substitution on the resin. The reagents have to be manually switched for specific steps in the synthesis, but the cycle of deprotecting, coupling and checking is done using these continuous flow setups.[12] The concepts of continuous flow synthesis were then applied to semi-automatic and fully automated solid-phase peptide synthesis. Semi-automatic synthesis is very much like the continuous flow mechanics, but most

parts of the synthesis are done automatically or are computer controlled. The apparatus would pump in a selected reagent and solvent into the reaction vessel, will control the amounts that get circulated through the machine, and monitor the reaction progress and constantly recirculate reagents until a specific substitution is met. All of these processes are controlled by using a computer with the proper programming and commands to allow for the synthesis to occur. The only aspect of this method that is manual is that the user must switch the amino acid that is to be used for each step of the synthesis, as automations adds a single amino acid at a time. These semi-automatic methods tend to only be viable for single amino acid additions, so the jump to fully automated peptide synthesis was made possible by the ability to add all amino acids sequentially without changing any parts of the set-up in a fully automatic synthesizer. These machines follow the same cycle of adding DMF then the amino acid, mixing, adding piperidine, washing and checking reaction progress until the peptide is fully synthesized. All of the amino acids that are to be used in the synthesis are housed within the machine and all of the processes that the control the flow of the reagents is done through a computer.[13]

Another major leap forward for the field of peptide synthesis was the introduction of on-resin cyclic peptide synthesis. The synthetic procedure follows the same concepts outlined by Merrifield in order to cyclize a peptide on resin using a "head-to-tail" method. Early reports of cyclic synthesis were not able to keep the peptide bound to the resin for cyclization, but upon cleavage from the resin, lactamization of the carboxyl terminus and amino terminus created the cyclic peptide. Fridkin demonstrated this by the synthesis of polyglycine and polyalanine cyclic peptides with the aid of polyalcohol carriers to ensure that the cyclic product is favoured and formed, preventing other intermolecular interactions or secondary structure formations. This method worked well upon liberation of the peptide sequence from the carrier; the yield of the

cyclic peptide was moderately high with some contamination from oligopeptides.[14] Other methods for the formation of cyclic peptides involved the bond formation between the side chains of specific amino acid residues. One instance of this type of cyclization was shown by Lebl through the formation of cysteine disulphide bridges across the peptide chain. While this method is very different from the "head-to-tail" methodology done by Fridkin, the cyclized peptide was formed on resin before being liberated from the resin.[15] Side-chain to side-chain cyclization has also been done on some occasions with Schiller performing the cyclization between the ε amino group and the β carboxyl group.[16] This synthesis more closely resembles the "head-to-tail" synthesis in which the terminal amino and carboxyl ends are joined, but this was done through the side chains as opposed to the backbone of the peptide. "Head-to-tail" synthesis of a peptide through the backbone functionalities while being bound to a resin was described by Sklyova through their synthesis of the glycine analog gramicidin S. By affixing a pre-synthesized tetrapeptide to a polymer, they were able to avoid inter- and intra-chain acylation that would give unwanted side products. The tetrapeptide (which contained a single unusual amino acid) of L-ornithine-L-leucine-D-phenylalanine-L-glycine was bound to the resin through the 5 amino group of L-ornithine. With the tetrapeptide bound to the resin, unwanted side reactions would be limited and the extension of the peptide could be carried out to complete the decapeptide sequence. The peptide was then cyclized on the resin to complete the synthesis of gramicidin S before being cleaved.[17] Another study in which the N-terminus and C-terminus of a peptide were cyclized on resin before cleavage was also conducted by Isied, where the hexapeptide $(Gly\text{-}His)_3$ was synthesized on resin. A 2,4-dinitrophenyl (Dnp) protected His residue was used to anchor the hexapeptide to a resin while Boc and OBzl were protecting the N- and C-terminus, respectively. Upon cleavage of the Boc and OBzl groups, the

peptide sequence was cyclized while remaining bound to the resin, but still had all of the Dnp protecting groups on the His residues.[18] These two examples of cyclic peptide synthesis are the "head-to-tail" type of synthesis being performed on resin that are very different from conventional linear solid-phase peptide syntheses.

The synthesis of a peptide using Merrifield's method begins with the choice of a resin. The resin bead is not an entirely solid bead in which a reaction occurs, but rather it is a porous bead that allows for reagents to pass through the insides of the mesh. The common term "solid-support" simply refers to the insolubility of the resin bead in solution. The resin that is used must be: a stable support that can be easily filtered off from any solvent that is used during the synthesis, must not react with any of the solvents or reagents that are used in the synthesis, should be able to swell in the solvent of choice so that upon expansion the reagents can access all of the available reaction sites on the resin, and should have an anchor to allow for the attachment of the first reagent or amino acid. A variety of resin types exist such as gel type resins, which generally are a polymer of different compounds, surface supports, composite gels and brush polymers. The connecting branch between the first amino acid and the resin is the anchoring support or a linker. A linker provides space between the resin and the peptide that is being synthesized; it also contains a functional group to allow the first amino acid to be coupled. The functional group can vary depending on the goals of the synthesis and the first amino acid – it can be an amino group, hydroxyl or even a halogen for some reactions. There are even some linkers that are specialized for syntheses of cyclic peptides that may involve linkage through a side chain.[12]

The coupling of the first amino acid to the linker has changed vastly over the years as the understanding of SPPS has evolved. In Merrifield's first described synthesis, an amino acid,

triethylamine and ethyl acetate were used to activate the reaction mixture so that linkage could occur.[8] In older methods, pre-activated acyl derivatives were prepared to form an ester bond between the linker and the first amino acid, such as acyl halides[19], esters[20], and anhydrides[21]. Methods for *in situ* activation involved *N,N'*-dicyclohexylcarbodiimide (DCC) which had been a cornerstone of solid-phase synthesis for quite some time. Other reagents for *in situ* activation that are more commonly used in the modern era are 2,6-dichlorobenzoyl chloride[22] and 1-(2-Mesitylenesulfonyl)-3-nitro-1H-1,2,4-triazole with 1-methyl-imidazole[23]. In some scenarios, during pre-activation, side reactions can occur between the free carboxyl group of the amino acid and the activation reagent to form an *O*-acylurea side product. These side reactions can be avoided through the use of benzotriazole derivatives and phosphonium or aminium salts, such as HOBt and HBTU/HATU.[24,25] The mixtures of these reagents along with the activating reagent are the requirements that are needed in order to couple the first amino acid to the linker successfully. After the first amino acid is bound to the resin, more amino acids can be coupled to elongate the chain until completion, but first the protecting group on the α amino group must be removed to form a new peptide bond to a second amino acid. Most techniques use Fmoc as a protecting group for the amino end of the residue, which is a very common protecting group because it only requires mild basic conditions to be cleaved. This allows for protecting groups on amino acid side chains to be acid-labile or more strongly basic labile. Cleavage of the Fmoc protecting group is carried out with treatment of the residue with 20% piperidine in DMF (Figure 1.2.1) creating a dibenzylfulvene-piperidine adduct that can be observed using UV spectrophotometry to analyze the resin substitution level.

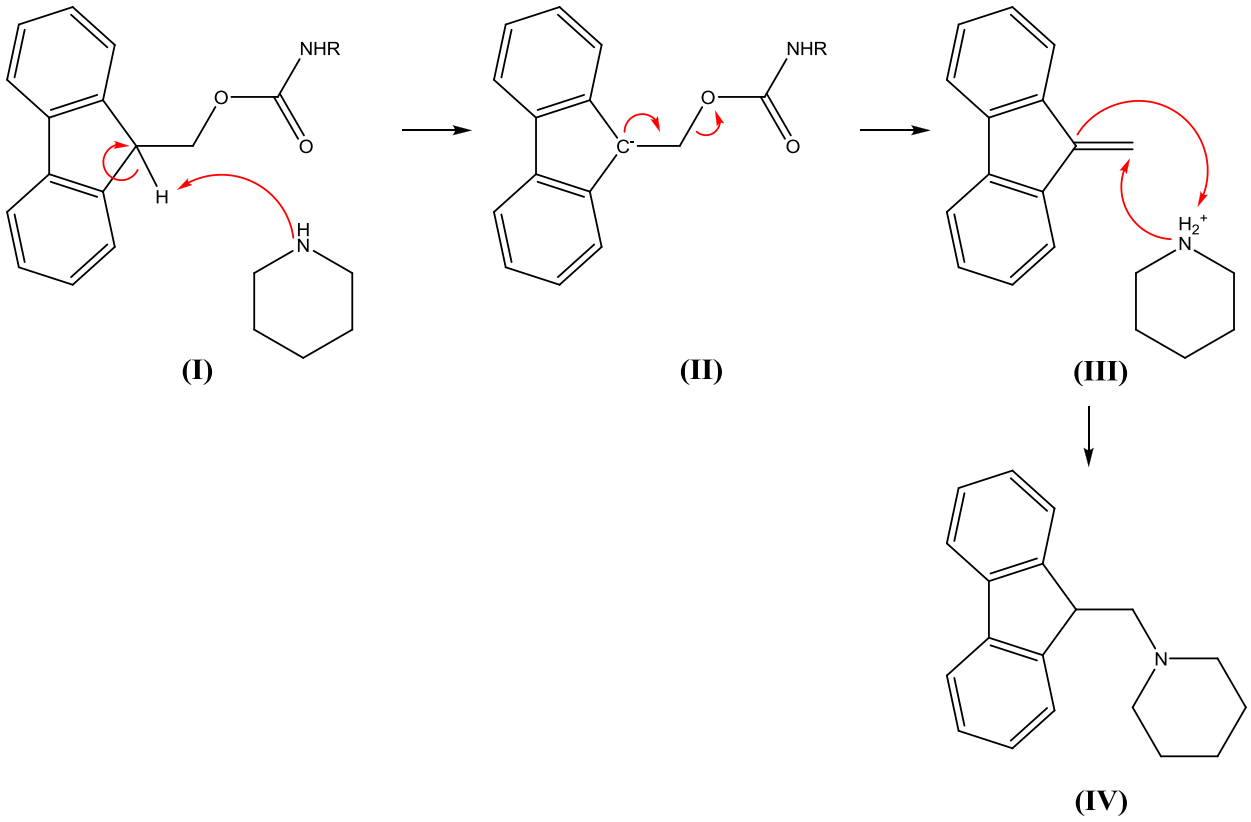


Figure 1.2.1: Reaction mechanism for the cleavage of a 9-fluorenylmethoxycarbonyl (Fmoc) protecting group from an amino acid. The 'R' group at the end of the chain represents the amino acid residue in which the Fmoc protecting group is bound to. Cleavage is initiated by treatment of piperidine (I) to create the carbanion (II). Upon elimination of the amino acid, dibenzylfulvene (III) is formed which then creates a dibenzylfulvene-piperidine adduct (IV) that can be measured using UV spectrophotometry to determine the substitution level of the amino acid coupling.

Analysis of the peptide and the coupling efficiency can be accomplished by measuring the absorbance of the dibenzylfulvene-piperidine adducts since it has an absorption maximum of 301 nm (with a ε of 7800 M^{-1} cm^{-1}). With a UV spectrophotometer, the absorbance of this cleaved product can be measured using the following equation[12]:

$$\text{Degree of Substitution } \left(\frac{\text{mmol}}{\text{g}}\right) = \frac{(0.01\text{L})(\text{Absorbance at 301 nm})(1000\frac{\text{mmol}}{\text{mol}})}{(\text{mass of sample})(7800\ \text{L M}^{-1}\text{cm}^{-1})(1\ \text{cm})} \quad (1.2.1)$$

Measuring the absorbance of the resin bound peptide is a very good quantitative check for the degree of substitution, along with other methods such as an elemental analysis and reverse-phase HPLC. Qualitative methods such as the ninhydrin test[26] and bromophenol blue test[27] give an idea of the coupling success based on a visual test.

After the check for the substitution level for the first amino acid and the cleavage of the Fmoc protecting group, the second amino acid can be attached to continue chain elongation. Traditionally, a mix of *N,N'*-diisopropylcarboiimide (DCC), HOBt, and the Fmoc protected amino acid were used as coupling reagents in DCM or DMF yielded good coupling results. Unfortunately anhydrous crystals of HOBt have been classified as an explosive and cannot be transported by sea or air.[28] Newer methods of peptide elongation use aminium or phosphonium salts, such as HBTU, HATU, HCTU, PyBOP, PyAOP and TBTU, to activate the reaction along with DIPEA to create the peptide linkage between the first and second amino acid. Reaction of the Fmoc protected amino acid with these aminium or phosphonium salts creates an ester product that can be easily coupled to the free amino end of the first residue. When the second amino acid is coupled, the quantitative and qualitative checks can be performed on the resin to determine the degree of substitution then washed through with DMF. The sample is then subjected to treatment with 20% piperidine in DMF to remove the Fmoc protecting group on the second amino acid and the third amino acid can be added. The cycle of deprotecting, adding the amino acid, checking and washing is done continuously until the chain has fully be elongated to

complete the peptide chain (Figure 1.2.2). The same methods can be utilized when synthesizing a cyclic peptide, but the final step is cyclization of the peptide by joining the N-terminus with the C-terminus by using aminium or phosphonium salts to activate the reaction.

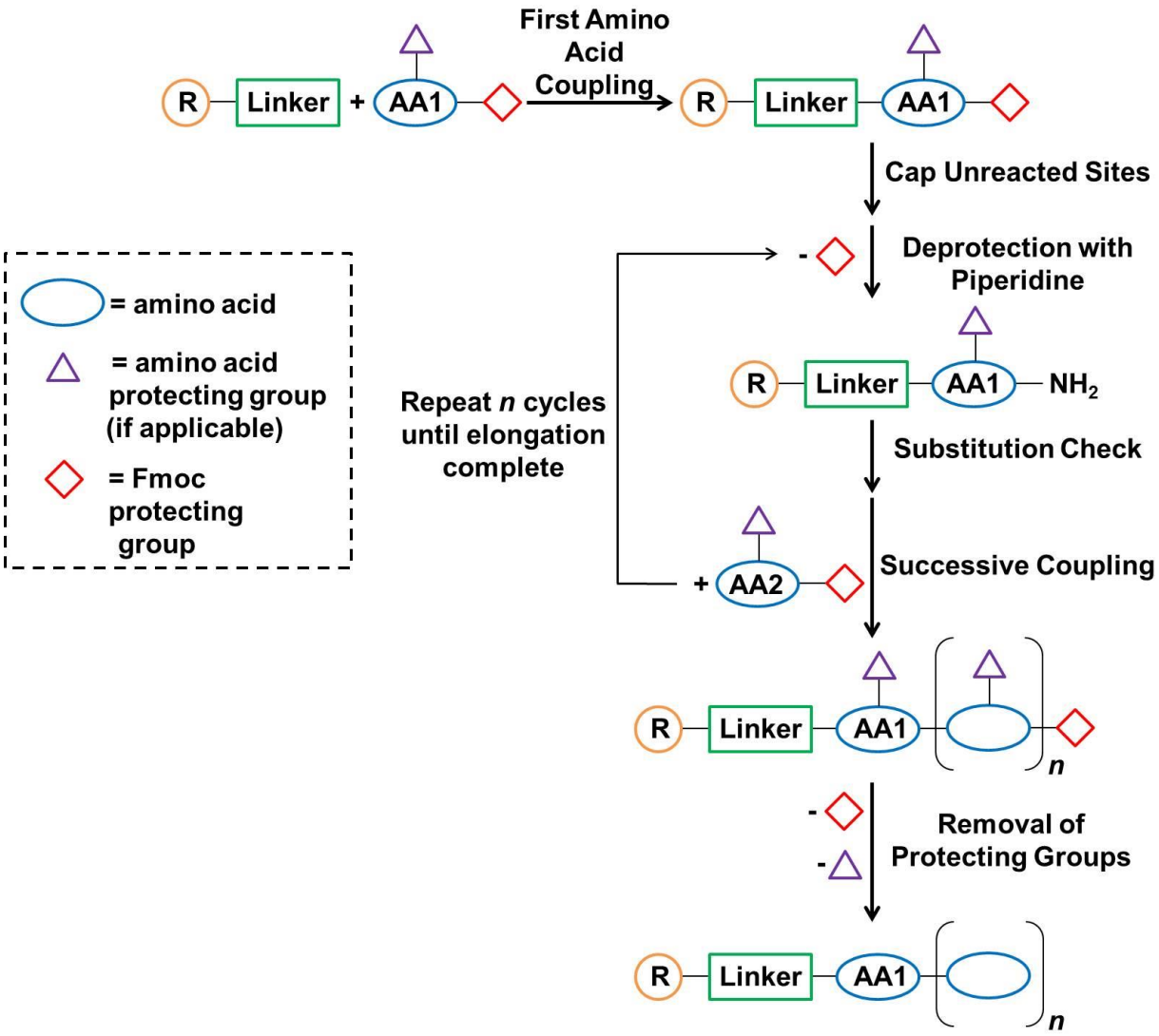


Figure 1.2.2: General schematic of the solid-phase peptide synthesis as proposed by Merrifield adapted from Kates and Albericio.[12] The resin, orange circle denoted "R", has an anchoring linker attached to it, green rectangle labelled "Linker", which contains a functional group in

which the first amino acid can be coupled to. The dashed box is the legend and shows what the blue circle, purple triangle and red diamond represent in the flow chart.

Merrifield's solid-phase peptide synthesis has evolved into a powerful tool in the field of peptide chemistry since it was first introduced back in 1963. His work to synthesize peptides on resin have given chemists a powerful tool to assemble segments of larger proteins that may not have been viable in solution state or that had extensive sequences. Because of his effort, Merrifield was awarded a Nobel Prize in Chemistry in 1984 for his discovery of solid-phase peptide synthesis. Even to this day, the foundation of his work is still being exploited in the synthesis of peptides in manual and automated manners.

1.3 Solid-Phase Peptide Synthesis Procedure

All of the amino acids needed for the synthesis of all peptides in this project were purchased from EMD Millipore Chemicals® and all solvents used were supplied by Sigma-Aldrich® and EMD Millipore Chemicals®. All reactions required for synthesis were carried out in a custom designed glass vessel known as a Merrifield vessel (Figure 1.3.1), which is 10 cm long, has a fritted disk at the bottom and can have a 2 mm glass stopcock inserted at the top.

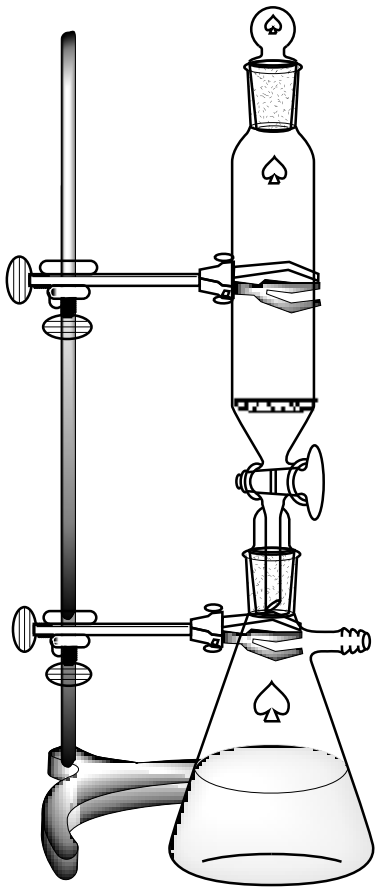


Figure 1.3.1: Manual solid-phase peptide synthesis is carried out in a Merrifield vessel inserted into a filter flask so that the removal of excess solvent and reagents can be collected easily.

The synthesis began by reacting 5 molar equivalents of a Fmoc-protected amino acid with 3.3 molar equivalents of DCBC and pyridine to couple the amino acid to the resin. The

vessel was allowed to shake on a Burrell Wrist-Action Shaker for a minimum of 18 hours for the reaction to take place. After the coupling attempt was complete, the resin was washed 4-5 times with DMF to remove any excess DCBC, pyridine and unreacted amino acid. At this point a very small portion of the resin-bound amino acid was removed and allowed to dry in a fume hood for approximately 45 minutes. Approximately 2-3 mg of the dry resin was weighed into an Eppendorf tube and 1 mL of 20% piperidine in DMF was added to cleave the Fmoc from the amino acid. The tube was then shaken for approximately 30 minutes to prepare the sample for a substitution check. After 30 minutes passed, the contents of the tube were passed through a glass pipette with glass wool to isolate fulvene adduct from the resin. Knowing the amount of fulvene in solution indirectly determined the molar substitution of the amino acid to the resin. A Beckman Coulter DU 640B Spectrophotometer set to read at 301 nm was used to determine the amount of fulvene. The solution was diluted 1:10 with DMF to make a 2% piperidine solution then placed in the spectrophotometer to determine the degree of substitution using Equation 1.2.1**.**

Measuring the degree of substitution needed to be performed after every coupling step with the goal of achieving a substitution of 80% or greater of the previously available sites. If the coupling was less than 80%, the reaction needed to be redone until at least the minimum requirement was met or if the coupling does not change after multiple attempts. If the substitution was equal to or greater than 80% the next amino acid coupling could be done.

After the first amino acid had been coupled, all of the unreacted sites needed to be capped to ensure that chain elongation properly occurred and no truncated chains were synthesized. The capping was done by treating the resin with 5 molar equivalents of acetic anhydride and DIPEA and allowed to shake for 20 minutes. The excess reagents were washed

out with DMF 4-5 times, a small portion of the resin was removed, dried and the substitution was measured to ensure nothing had changed since the capping reaction began. The next step was to cleave Fmoc from the amino acid to leave a free amino terminus so that the next amino acid could be coupled. Fmoc cleavage was carried out by treating the resin with 20% piperidine in DMF and allowing it to shake for a minimum of 30 minutes. The resin was then washed with DMF 4-5 times to remove all the fulvene that had been cleaved, then a small amount of the resin was removed, dried and a substitution check was performed to ensure all of the Fmoc had been removed from the amino acid. An indication for when all of the Fmoc had been cleaved from the amino acid was when the absorbance for the substitution check is negative or zero.

Chain elongation continued by reacting 4 molar equivalents of an Fmoc protected amino acid, HBTU and HOBt, 1.5 molar equivalents of DIPEA and the resin bound amino and allowing the mixture to shake for a minimum of 18 hours. Immediately following the reaction time, the resin was washed 4-5 times with DMF to remove any excess reagents and amino acid that may not have reacted. At this point a small amount of the resin-bound peptide was removed and allowed to dry in a fume hood for approximately 45 minutes. A substitution check was performed next to determine the success of coupling of the amino acid to the previous one. The procedure to do this utilized the same method for checking the first amino acid coupling and Equation 1.2.1 as well. The coupling reactions were done numerous times until the coupling of the previously available sites was 80% or greater, or until the percentage did not improve with multiple attempts. When the coupling of the amino acid to the previous one was deemed sufficient, the Fmoc protecting group could be cleaved and the next amino acid was added. This continuous cycle of deprotecting an amino acid, coupling the next one, checking the substitution, and deprotecting again was done until the peptide had been fully synthesized.

After chain elongation was complete, any protecting groups that remained on the peptide were removed. There were a wide variety of protecting groups for each amino acid and each required different conditions to be cleaved. The Fmoc protecting group was simply removed with 20% piperidine in DMF, but other protecting groups required mild to harsh acidic or basic conditions depending on the reactivity. Once all the protecting groups were removed and washed out, the peptide was fully synthesized free from the effects of protecting groups.

An alternate method was also used to synthesize the FK-13 peptide. An HMBA-AM resin was used instead of the Wang resin, which allowed for the amino acids to possess strongly acidic labile protecting groups. The HMBA-AM resin is a base-labile resin, which requires a strong base to cleave the peptide from the resin. Thereofre at the end of the synthesis, the protecting groups can be treated with a strong acid while the peptide remains bound to the resin.[21] Instead of using a combination of HBTU and HOBt, just HCTU was substituted in place of both of these reagents. In addition to this, the amount of molar equivalents needed to do each coupling attempt was increased to 5 molar equivalents for the Fmoc-protected amino acid and HCTU and 10 molar equivalents for DIPEA. Also, the time needed to shake the vessel was greatly reduced from 18 hours to just 2. Coupling of the first amino acid and substitution checks all remained the same for this slightly modified procedure.

1.4 Nuclear Magnetic Resonance Spectroscopy

1.4.1 Basic Principles

One of the most powerful tools for chemical analysis of organic and inorganic compounds is nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy analyzes the response of magnetic dipoles of atomic nuclei to an external applied magnetic field. For nuclei to possess a nuclear magnetic moment, **μ**, the nuclear spin angular momentum, ***I***, must be greater than zero. Common nuclear isotopes that have an ***I*** = ½ are ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P, while nuclei that have an ***I*** = 0 are said to be NMR inactive. In the absence of a magnetic field, all orientations of the magnetic moments are energetically equivalent, but in the presence of an applied external magnetic field, **B**, the energy can be defined by:

$$E_{mag} = -\mu \cdot B \tag{1.4.1}$$

If this external field is applied, the z component of the ***I*** vector has $2I+1$ possible orientations of the magnetic moment. Also with spin and magnetism proportional to one another, the magnetic moment can be defined by:

$$\mu = -\gamma \cdot I \tag{1.4.2}$$

In equation 1.4.2, **γ** is the magnetogyric ratio, which is unique for every nuclear isotope. In an applied magnetic field, there are two states which a magnetic moment can adopt: α and β state (Figure 1.4.1). These different energy states are known as Zeeman levels and can be described as:

$$\Delta E = (h\gamma \div 2\pi) B_0 \tag{1.4.3}$$

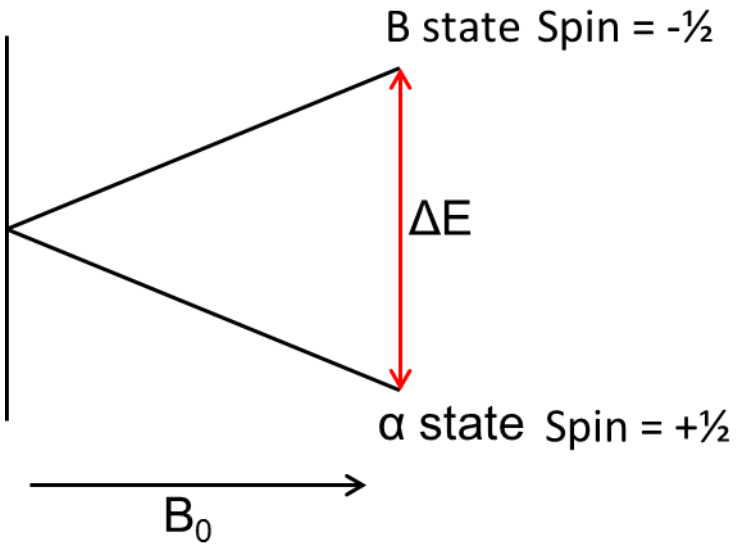


Figure 1.4.1: For a spin ½ nucleus, two different states can be occupied by the spins (α and β). The energy it takes for a spin to move from the α state to the β state is defined by the ΔE and this energy different is unique for every nucleus.

While the magnetic field remains off, the spin polarization axes point in a variety of directions (not just up or down), but once the field is turned on the spin polarization begins to spin around the field at a constant angle between the field and the spin magnetic moment. This movement of spinning is called precession and is analogous to a top spinning. If the top is spinning upright it will be stable, but as soon as the top becomes tilted, it begins to topple to the ground. If the top on an angle is spinning fast enough though, it will circle around trying to keep all the momentum it has and not topple over. The frequency of precession is proportional to the magnetic field and is commonly known as the Larmor frequency (given in rad s^{-1}).[29]

$$\omega^0 = -\gamma \cdot B^0$$

(1.4.4)

Many nuclei have a positive γ which gives the Larmor frequency a negative value, meaning the nuclei will precess in a counter-clockwise manner. On the other hand, when the γ has a negative value, the Larmor frequency will be positive and precession will be in the clockwise direction. While the magnetic field is on, the spins will fluctuate at different magnitudes and directions in tiny amounts. Even though these fluctuations are minuscule, they are needed in the long term to observe the difference in magnetism. Eventually when the spins stop precessing, they return to a thermal equilibrium where the distribution of spin polarizations is stable. There will still be slight precession occurring in the nuclei and the net distribution of spins will likely be in the same orientation as the magnetic field.[29]

As soon as the magnetic field is turned on, the macroscopic nuclear magnetization gradually builds up due to the random wandering of the spin polarizations. This exponential build up occurs exponentially (a first-order process) over a time described by T_1, is known as the spin-lattice relaxation time constant or the longitudinal relaxation time constant. When the polarizations are rotated by a $\pi/2$ pulse along the x-axis, the net spin polarizations that occur along the z-axis would rotate into the –y-axis. Over time though, these spins will want to return to the thermal equilibrium populations and the precessing nuclear magnetizations will begin to decay. This decay of the macroscopic nuclear magnetizations is characterized by another first-order time constant called T_2, also known as spin-spin relaxation time constant or the transverse relaxation time constant.[29]

The electronic cloud of a nucleus responding to change due to an applied magnetic field is known as chemical shift. The applied magnetic field induces the circulation of electrons within the nucleus which gives rise to magnetic fields surrounding the nucleus as well. Since these electrons produce their own magnetic field, the Larmor frequency also changes depending

on the environment of the nucleus and the presence of other magnetic nuclei that may be interacting with the electrons. Chemical shift, **δ**, can be calculated by measuring the Larmor frequency of a specific nucleus and comparing it to a reference compound (such as trimethylsilane), and is normally measured in Hz or ppm:

$$\delta = \frac{\omega^0 - \omega^o_{ref}}{\omega^0_{ref}} \tag{1.4.5}$$

Chemical shielding, **σ**, is the influence of the electric current on a nucleus, and is driven by the magnetically-induced circulation of electrons. It is measured with respect to the "bare" nucleus of an atom and is composed of diamagnetic (σ^d) and paramagnetic (σ^p) terms:

$$\sigma = \sigma^d + \sigma^p \tag{1.4.6}$$

Orientation of the molecule in the field plays a role in how the chemical shift and chemical shielding are affected. This orientation dependence can be best described using the chemical shielding tensor, which is a 3x3 matrix that relates the orientation of the molecule to how the electric currents are being induced:

$$\sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \tag{1.4.7}$$

The principal components of the principal axis system are the eigenvalues σ_{11}, σ_{22}, and σ_{33}. Larmor frequency is dependent on the chemical shielding tensor and can be related back to Equation 1.4.4:

$$\omega^0 = \gamma(1-\sigma_{iso})\cdot B^0 \tag{1.4.8}$$

σ_{iso} is the average of the principal components: σ_{11}, σ_{22}, and σ_{33}. The chemical shielding tensors can be used to hypothesize the structure and orientation of the molecule within the magnetic field.[29]

In many cases, there will be multiple nuclei in a molecule to be analyzed. Indirect spin-spin coupling, more commonly known as *J*-coupling, is a mechanism in which nuclear spins are able to communicate with one another through chemical bonds. The term "coupling" is used in multiple scenarios throughout this thesis, in solid-phase synthesis, and for *J*-coupling. Therefore, in this thesis, the term coupling will refer to synthetic reactions during peptide synthesis and *J*-couplings will be referred to as interactions. These interactions depend on the orientation and structure of the molecule being analyzed, but are independent of the external magnetic field. Since *J*-coupling is electron mediated, it can be used to establish the connectivity between different nuclei. The *J*-coupling is reported in Hz and usually is designated as $^{X}J_{AB}$ where X is the number of bonds between nuclei A and B. The interaction between nuclei gives rise to the splitting in NMR spectra that describes the association of one nucleus with other nuclei that it may be interacting with. Theoretically, *J*-coupling can be described as the sum of four individual terms:

$$J_{\text{Total}} = J_{\text{orbital}} + J_{\text{spin-dipolar}} + J_{\text{Fermi-contact}} + J_{\text{spin-dipolar-FC}} \quad (1.4.9)$$

J_{orbital} is the product of nuclear spin angular momentum coupling and electronic orbital momentum, and has diamagnetic and paramagnetic terms. $J_{\text{spin-dipolar}}$ are interactions between nuclear and electron spins that are not near the nucleus. $J_{\text{Fermi-contact}}$ is dictated by the density of electrons at the nucleus, and usually only describes s-character bonding orbitals. $J_{\text{spin-dipolar-FC}}$ is a cross term between spin-dipolar and Fermi-contact interactions.[29]

Dipolar coupling is dependent on the through-space interactions between two magnetic nuclear spins. Energetically, this is equivalent to:

$$E = \left(\frac{\mu_X \mu_Y}{r^3} - 3\frac{(\mu_X \cdot r)(\mu_Y \cdot r)}{r^5}\right)\left(\frac{\mu_0}{4\pi}\right) \quad (1.4.10)$$

Equation 1.4.10 is the classic description of nucleus A interacting with nucleus B, where **r** is the distance between the two nuclei, but the quantum mechanical equivalent is:

$$\hat{H}_{DD} = \gamma_X \gamma_Y \hbar^2 \left(\frac{\hat{I}_X \hat{I}_Y}{r^3} - 3\frac{(\hat{I}_X \cdot r)(\hat{I}_Y \cdot r)}{r^5}\right)\left(\frac{\mu_0}{4\pi}\right) \quad (1.4.11)$$

With the use of spherical polar coordinates, equation 1.4.11 can be converted to:

$$\hat{H}_{DD} = \frac{\gamma_X \gamma_Y \hbar^2}{r^3}[A+B+C+D+E+F]\left(\frac{\mu_0}{4\pi}\right) \quad (1.4.11)$$

The term in the front, before the six different terms (Equation 1.4.12), is usually represented as R (in Hz) and is the dipolar coupling constant.

$$A = \hat{I}_{Xz} \cdot \hat{I}_{Yz}(3\cos\Theta - 1) \quad (1.4.12)$$

$$B = \frac{1}{4}(\hat{I}_{X+} \cdot \hat{I}_{Y-} + \hat{I}_{X-} \cdot \hat{I}_{Y+})(3\cos\Theta - 1)$$

$$C = -\frac{3}{2}(\hat{I}_{Xz} \cdot \hat{I}_{Y+} + \hat{I}_{X+} \cdot \hat{I}_{Yz})\sin\Theta\cos\Theta e^{-i\Phi}$$

$$D = -\frac{3}{2}(\hat{I}_{Xz} \cdot \hat{I}_{Y-} + \hat{I}_{X-} \cdot \hat{I}_{Yz})\sin\Theta\cos\Theta e^{+i\Phi}$$

$$E = -\frac{3}{4}\hat{I}_{X+} \cdot \hat{I}_{Y+}\sin^2\Theta e^{-2i\Phi}$$

$$F = -\frac{3}{4}\hat{I}_{X-} \cdot \hat{I}_{Y-}\sin^2\Theta e^{+2i\Phi}$$

In a rapidly tumbling solution, the six terms (A to F) are orientation-dependent and average out to zero. On the other hand, solids that are locked in space do not average out to zero in the terms A to F.[29]

1.4.2 High-Resolution Magic Angle Spinning (HRMAS)

The NMR of solids is very different from solutions. Due to the lack of free tumbling motion, dipolar coupling is the dominant mechanism (as opposed to *J*-coupling) and the resonance frequency depends on the orientation of the nuclei in the magnetic field. In addition to these differences, solids are difficult to work with as they have longer T_1 times, are insensitive, exhibit line broadening due to dipolar coupling, and possess anisotropic line shapes. Luckily, there are many techniques to improve solid-state NMR spectroscopy, one being high-resolution magic angle spinning (HRMAS). HRMAS (and other techniques used to improve solid-state NMR) is beneficial in the sense that solubility (and usually volatility) is not a problem, the data collected can be compared to X-ray results, Fourier-transform NMR is not limited by freezing or boiling point of solvents, and just a powder is needed (not entire crystals). Since solid-state NMR is orientation-dependent, the chemical shift anisotropy term is affected by these solids and causes broadening in spectra. Factors such as the magnetic field strength, flux and solvent-polymer interactions also add a magnetic susceptibility broadening factor. Each of these issues that cause spectral broadening can be addressed by spinning the sample at the magic angle of 54.74° with respect to the magnetic field (Figure 1.4.2). This angle is a product of the orientation-dependent term $P_2(\cos\Theta)$ equalling zero[30]:

$$P_2(\cos\Theta) = (3\cos^2\Theta - 1)/2 = 0 \qquad (1.4.13)$$

Where $\cos^2\Theta = 1/3$

$\Theta = \cos^{-1}(1\sqrt{3}) = 54.74°$

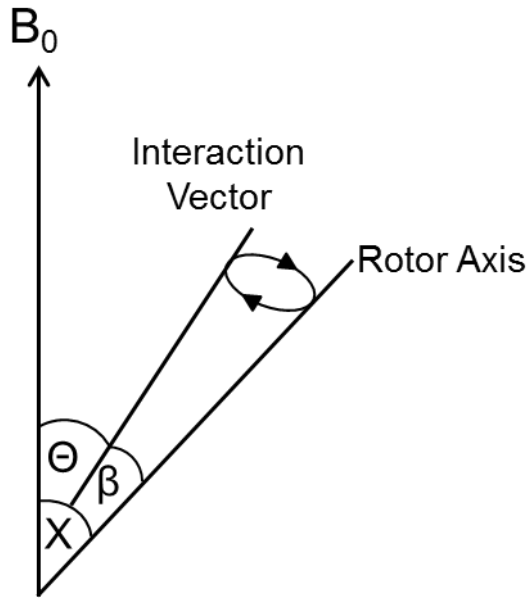


Figure 1.4.2: Magic angle spinning is an orientation dependent technique that spins the sample at $\Theta = 54.74°$ with respect to the magnetic field. Line broadening is common amongst solid-samples, but by spinning the sample at the magic angle narrow lines can be obtained.

The result is a reduction of line broadening, yielding narrow lines at the isotropic frequency if the spinning rate is sufficiently high compared to the breadth of the interactions being averaged. If the spinning rate is less than the total breadth of the interaction, then spinning side bands (which look like narrow peaks) will also occur at multiples of the spinning rate from the isotropic peak.

1.4.3 Pulse Sequences

The nuclear spin polarizations must be excited in order for a signal to be produced. This is executed by initially exciting the nuclear spins with a radiofrequency (r.f.) pulse. In the simplest case, the pulse excites the nuclear spins, and as they return to thermal equilibrium, the weak r.f. signals radiated by the nuclei are amplified and detected. An r.f. pulse is used to excite the sample, which then reacts by released transverse nuclear magnetization. Once the r.f. pulse is turned off, the nuclear spins give off their own r.f. NMR signal (free induction decay, FID), but they are very weak and need to be amplified before they can be processed. The signal is then digitized and can be converted from a function of time to a function of frequency with a mathematical operation known as Fourier transformation. The result is an NMR spectrum that is displayed on the computer.[29]

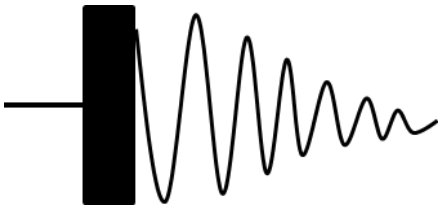


Figure 1.4.3: Block diagram of a single pulse NMR experiment, where the black rectangle with a phase that is a multiple of $\pi/2$ and the signal acquisition follows (wave line).

The pulses used are in phases that are multiples of $\pi/2$ (or sometimes 90°). For a single pulse experiment, such as the one in Figure 1.4.3, the nuclear spins are excited once and then detected, but by adding a second r.f. pulse, more spectral information can be obtained. After the first r.f. pulse is used, a delay time, t_1, is initiated before the second pulse is released and then detection occurs. Figure 1.4.4 shows a two pulse sequence, but multiple pulses per sequence are not uncommon.

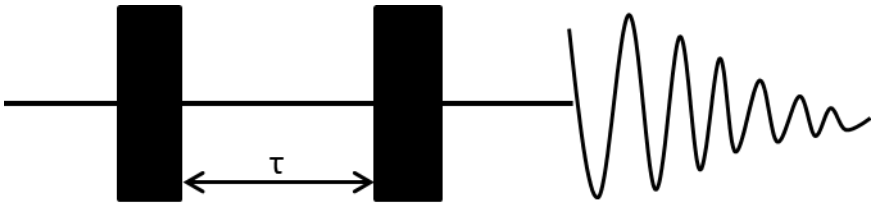


Figure 1.4.4: In a multiple pulse sequence, two different pulses (black rectangles) are used with a time delay in between pulses (τ). Acquisition occurs after the second pulse. Two pulses can be used to irradiate nuclei on two separate occasions to achieve different goals (depending on the pulse phases used).

Until now, the examples given were for a single channel experiment where only one nucleus was being irradiated and observed. In many cases, this type of experiment is not sufficient and a second channel is added. This second channel is commonly a different nucleus than in the first channel, i.e. an atom such as ^{13}C or ^{15}N. These types of experiments involve multiple pulses on both channels, multiple time delays, often decoupling and the acquisition can occur on either channel. Figure 1.4.5 shows an example of one of these two channel experiments (INEPT). After acquiring the signal from these two channel experiments, two-dimensional spectra are normally produced, but by adding more channels, extra dimensions can be added into the spectrum. It is not uncommon to add a third dimension to obtain 3D NMR and there was an instance where a 7D NMR spectrum was obtained.[31] Adding more channels and dimensions to a spectrum allows for more information about the sample being analyzed to be revealed, depending on the pulse sequences used.

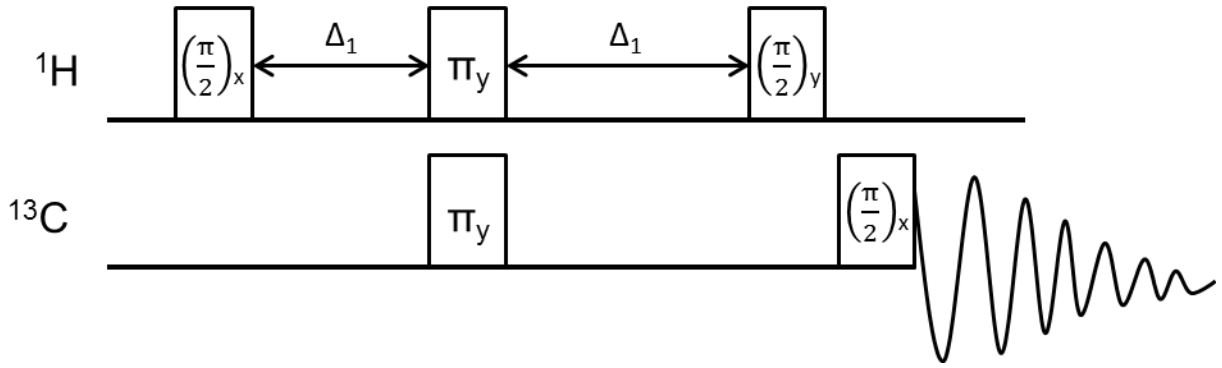


Figure 1.4.5: Block diagram of the INEPT pulse sequence is an example where second channel is added to observe another nucleus. In this case the second channel is ^{13}C and data acquisition occurs on this channel as opposed to on the ^{1}H channel.

1.4.4 COSY, TOCSY and ROESY

There are many different pulse experiments available, but the three focussed on in this thesis are COSY, TOCSY, and ROESY. These are all single channel experiments that observe ^{1}H signals. COSY, or correlated spectroscopy, establishes connections between nuclei that are *J*-coupled. The block diagram from this pulse sequence is shown in Figure 1.4.6; the pulse begins with a $\pi/2$ pulse, followed by a variable t_1 delay time. Next, a second $\pi/2$ pulse is used and the signal is collected during a time t_2. The FID is collected for both time delays and Fourier transformed in both dimensions. The result is frequencies that will align with one another to create "cross-peaks" in the spectrum. The signals in a COSY spectrum only represent nuclei that are connected by *J*-coupling. In the simplest case, COSY will detect ^{1}H atoms that are three bonds or less away from one another. This experiment is especially useful for determining connectivity between nuclei in molecules.[32]

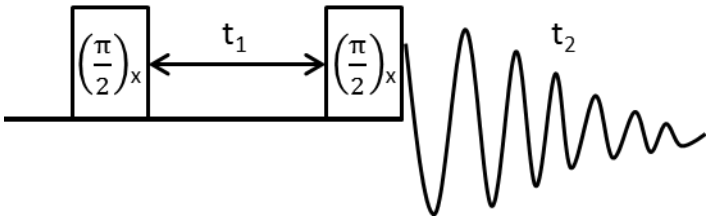


Figure 1.4.6: Block diagram of the COSY pulse sequence. A π/2 pulse along the x-axis is applied, then a time delay occurs followed by another π/2 pulse along the x-axis. The COSY technique allows for nuclei that are *J*-coupled to be observed.

Total correlated spectroscopy (TOCSY) is a technique that mixes all of the magnetization for a set of coupled nuclei. Nuclei that are within the same spin system will be observed interacting with one another. The method to this approach is that a π/2 pulse is applied and then allowed to evolve during t_1. Another π/2 pulse is then applied and right after, a mixing time begins. This mixing time is a period when the magnetization is transferred throughout the spin system of the molecule. Instead of being restricted by *J*-couplings (as COSY is), all of the spins that are within the system that are interacting with one another will be observed as cross-peaks in the spectrum. However, magnetization does not travel across carbonyl, ethers and other heteroatoms, but this can be beneficial as it can be used to identify individual residues within larger peptides. After the mixing time, another π/2 pulse is applied and the signal is acquired. In the case of larger proteins and peptides, this technique can prove to be more beneficial than COSY since it can identify all the spins within individual residues.[32]

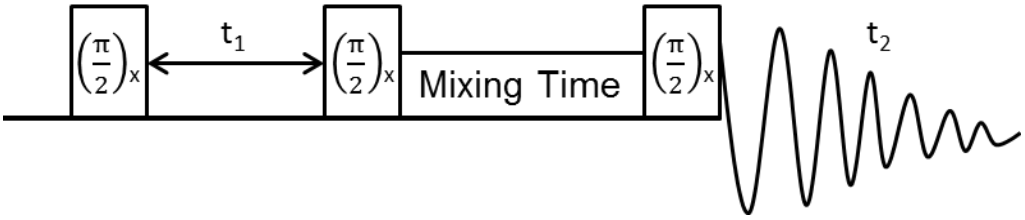


Figure 1.4.7: Block diagram of the TOCSY pulse sequence. A π/2 pulse along the x-axis is applied, followed by t_1 delay and then another π/2 pulse. The mixing of the magnetization throughout the spin systems then occurs during the "mixing time". Nuclear magnetization is transferred from nuclei to nuclei allowing for all of the spins in an entire spin system to be observed.

Rotating-frame nuclear Overhauser effect correlation spectroscopy (ROESY) is a technique that provides information concerning distances between nuclei. This technique is based on dipolar interactions between nuclei (NOEs) in the rotating-frame. The rotating frame is a coordinate system that rotates with the field to make the r.f. field look static. The rate of change for the net magnetization, with an applied magnetic field (B_1), in the rotating frame can be represented as:

$$\frac{d\mathrm{M}(t)}{dt} = \mathrm{M}(t)\, x\, \left(\mathrm{B}_0 + \frac{\omega}{\gamma} + \mathrm{B}_1\right) \qquad (1.4.14)$$

This results in M(t), being independent of time. While in the rotating frame, the magnetization will not experience any precession if it is in an applied field. Relaxation in the rotating frame will be achieved by B_1 decaying with a time, $T_{1\rho}$, or the spin-lattice relaxation time in the rotating frame.[29] Initially a π/2 pulse along the x-axis is applied then a time t_1 occurs. Following the t_1 time, there is a spin-lock along the y-axis. This spin-lock, keeps the nuclear polarization spinning along the axis the entire duration of this event. Once the spin-lock is complete, signal

acquisition occurs. If the pulse for the spin-lock is too strong then magnetization is able to transfer to *J*-coupled spins, which would be give the same information as COSY and subsequently reduce the dipolar interactions in the molecule. The NOEs that are produced from ROESY all have a positive sign, but those interactions that come from shared magnetization through *J*-coupling also have the same sign which makes interpreting the data more difficult. ROESY is beneficial when analyzing intermediate sized molecules (1000 to 3000 MW) but NOE enhancement is not as effective for large or very small molecules. In small peptides, ROESY is useful and can be used to establish the three-dimensional structure by analyzing and evaluating the different NOE interactions that are seen in the spectra.[32]

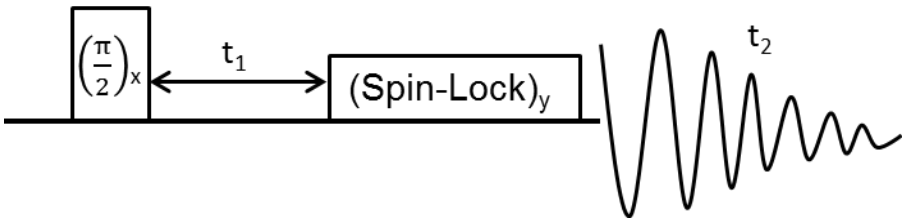


Figure 1.4.8: The ROESY pulse sequence begins with a π/2 pulse along the x-axis and then a time delay, t_1. While the bulk magnetization is along the y-axis, a "spin-lock" is applied on that same axis to keep the magnetization along that axis. This allows for through-space NOEs (dipolar interactions) to evolve and be observed once the data is acquired.

In the following sections of this thesis, the results of the work done on glutathione and FK-13 will be presented. Chapter 2 contains an introduction to the GSH peptide followed by the results. The results of the solid-phase synthesis, NMR, IRMPD and structure calculations on the GSH peptide are presented first, followed by a discussion summarizing all of the results and drawing conclusions about the work. Chapter 3 begins with an introduction to the LL-37 peptide and FK-13 which is then followed by the results. The results of the FK-13 peptide consist of the solid-phase synthesis and NMR data. Following these results, a discussion about the results and concluding statements are given. Finally, a conclusions chapter, Chapter 4, will provide statements about the work done in this thesis. These conclusions include a summary of the work done on both peptides, the value of the work accomplished, and the future directions of the work.

2.0 Glutathione

2.1 Introduction

2.2 Experimental Procedures

2.3 Results

2.4 Discussion

2.1 Introduction

Glutathione (GSH) is an essential peptide in the protection against reactive oxygen species and free radicals through its role in certain enzymes. The structure of GSH has been studied with a variety of methods including NMR, molecular dynamic calculations, mass spectrometry and IRMPD spectroscopy. However, this peptide is still being studied in a variety of manners and new aspects of the tripeptide are being discovered. It is a simple peptide that can be analyzed using a variety of techniques and still has great implication in the field of peptide chemistry.

Glutathione (GSH) is a tripeptide that with the amino acid sequence γ-glutamic acid-cysteine-glycine (ECG), but there is an unusual peptide linkage between the γ-carboxyl group of glutamic acid and the α-amino group of cysteine (Figure 2.1.1). This peptide primarily acts as an antioxidant and antitoxic agent.

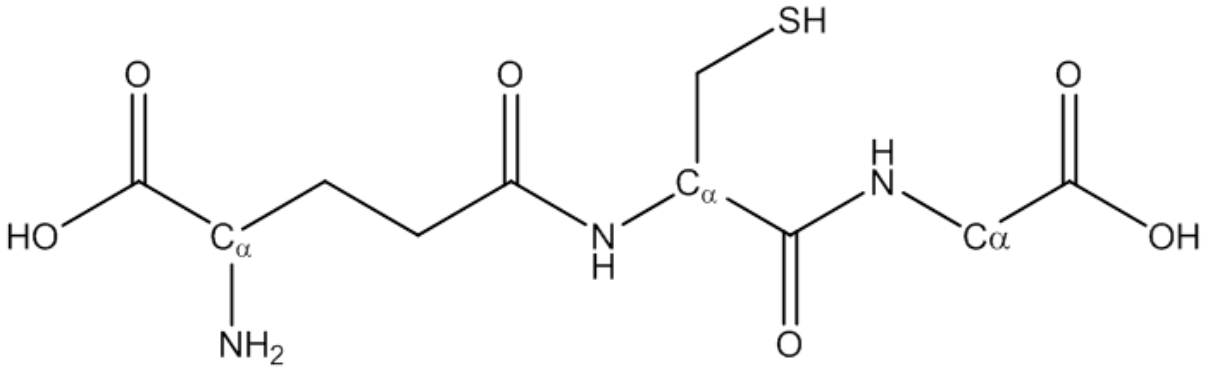


Figure 2.1.1: Skeletal structure of neutral glutathione with the α carbons labelled. A conventional peptide bond between amino acids would be between the α-carboxyl group of one amino acid and the α-amino group of the next amino acid. In GSH, the γ-carboxyl group of glutamic acid (Glu) forms a bond with the α-amino group of cysteine (Cys).

Glutathione is a biosynthetic peptide synthesized primarily by two enzymes: γ-glutamylcysteine synthase (γ-GCS) and glutathione synthase (GS). In the synthesis of GSH, γ-GCS synthesizes γ-glutamylcysteine via an ATP-mediated condensation reaction between glutamate and cysteine. This formation of γ-glutamylcysteine is the rate-determining step in the biosynthesis of GSH, and is limited by the amount of cysteine available in the cell. The next step is to create GSH by adding glycine to the C-terminal end of the dipeptide. GS is the enzyme responsible for this step in the synthesis and is also an ATP driven reaction.[33]

One role of GSH is to act as an antioxidant by scavenging free radicals or reducing harmful peroxides, but this produces a potentially harmful dimer, GSSG (oxidized GSH). The antioxidant ability of GSH is due in large part to the peptide's thiol group that can donate a H^+ and electron for these reactions. The dimer can then be converted back to GSH with the help of glutathione reductase, which is an NADPH dependent enzyme. GSH is also a co-factor for GSH-peroxidases (GPx), which are another important class of enzymes that protect the cell against oxidants. These enzymes function by reducing hydrogen peroxide and other peroxides to water, an alcohol and GSSG. Glutathione S-transferases (GST) are another major class of enzymes that relieve the oxidative stress on cells from xenobiotics. They act against α,β-unsaturated carbonyls, epoxides and peroxides that are introduced into the cell. By reducing these harmful agents, GSTs prevent them from reacting with other important macromolecules such as proteins and nucleic acids.[34] The antitoxic function of GSH is very similar to its antioxidant role. GSH has nucleophilic and reducing properties so it can interact with electrophiles and oxidizing agents to prevent harm to the cell. Most of the detoxifying roles that GSH possesses work in tandem with GS-S-transferases to reduce the effect that electrophilic xenobiotics have in the cell.[35] These GSTs have also been shown to provide defense against

carcinogens and affect antineoplastic drugs.[36] GSH is an important cofactor for the function of a variety of enzymes that help in the defense against oxidizing agents and potentially harmful toxic compounds.

The structure of glutathione is one that has been studied in the past by NMR and other computational methods. The first experiment to determine the structure of GSH was performed by Fujiwara by analyzing a fully ionized sample with the aid of ^{1}H NMR spectroscopy. Analysis of 1D ^{1}H NMR spectra and the *J*-coupling of the peaks suggested that the most stable conformer of the tripeptide was when the CO_2^- and NH_3^+ of glutamate are fully extended away from the rest of the chain.[37] A later study done by York was performed to analyze GSH in solution again, but ^{13}C NMR was used along with ^{1}H NMR to determine the structure of the peptide. The *J*-coupling in the ^{1}H NMR spectra was used along with the spin-lattice relaxation times in the ^{13}C NMR to hypothesize a structure of the tripeptide. Their studies revealed that at physiological pH, the glutamate and cysteine ends are hindered in their motions, but the glycine end is very mobile and can rotate freely. They also observed an interaction between the ^{1}H of the glutamate amino group and the carbonyl of the peptide backbone. This observation contradicts the results that Fujiwara had obtained because York's results suggest that the CO_2^- and NH_3^+ are not extended away and folded in towards the rest of the backbone.[38] Additional studies have shown that GSH has the flexibility to coordinate itself with metal ions such as Cd, Zn, Pb and Hg with high specificity as proven with ^{13}C NMR.[39] GSH acid-base chemistry has also been monitored using NMR spectroscopy while the peptide was coordinated to methylmercury.[40] High-resolution magic angle spinning (HRMAS) NMR spectroscopy combined with computational calculations were used to predict the chemical shifts and hypothesize a structure. ^{1}H, ^{13}C and ^{15}N NMR experiments were performed on crystals of GSH

to determine the experimental shifts for the individual nuclei. The experimental results were then compared to calculated ^{1}H and ^{13}C shifts, and suggested there were similarities between the two methods to validate the structure of GSH.[41]

Further analysis of the GSH structure was done with the aid of molecular dynamics experiments. GSH and 16 different charge states were analyzed by Lampela using molecular dynamics simulations and NMR spectroscopy to determine if the peptide favoured a conformation at varying pH. The results of this study revealed that the NMR results did not fully agree with the MD simulations, as the side chains of the Cys and Glu residues preferred a staggered conformation in the NMR data, and the simulations suggested that it preferred a gauche orientation. Additionally at physiological pH, the simulations revealed that the molecule did not prefer any state due to the flexibility of the chain. The study also looked at the proximity in which the thiol group of Cys comes with the Glu C_{α}-H bond. The most favoured structure for this thiol and hydrogen contact is when the contact is minimal and the Glu has NH_3^+ and COO^- groups, the thiol is uncharged, and the Gly carboxyl is COO^-.[42] Zhang also used molecular dynamic and NMR experiments to study the conformation and hydrogen bonding properties of GSH in water. NMR spectroscopy was used to determine the shift of the ^{1}H's that were responsible for taking part in the hydrogen bonding with water. The varying differences in frequency determine whether a bond was strong or weak, and could be compared to the molecular dynamic calculations. The simulations suggested that GSH has three common conformations: extended, semi-folded and folded (Figure 2.1.2). The extended form is the most favourable of all the states and can interchange with the folded state, but that rarely occurs. The hydrogen bond that was strongest in the simulations was between the amide proton of glycine and the water residues, but other water residues will coordinate to the cysteine amide proton to

form stable clusters.[43] Other molecular dynamics simulations with GSH at varying pH showed that at pH 7.0 (close to physiological conditions), the tripeptide had no favourable conformation and can exhibit a wide variety of structures similar to Lampela's work.[44]

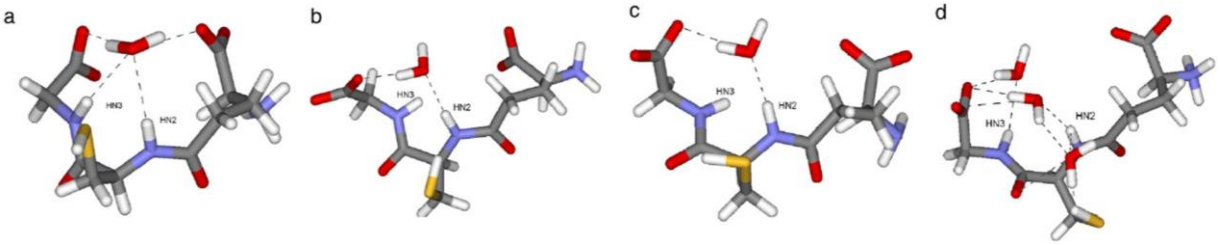


Figure 2.1.2: GSH predominantly exists in one of three states: a) folded, b) and d) extended and c) semi-folded. Zhang *et al.* conducted NMR and molecular dynamic experiments to observe the coordination of water molecules to the various hydrogen bond acceptors in GSH to observe which conformation is most favoured.[43]

Mass spectrometry and IRMPD spectroscopy have also been used to analyze the structure of GSH on multiple occasions. Electrospray ionization (ESI) mass spectrometry was used to obtain spectra of a protonated GSH and GSH dimer. A mixture of GSH and various amino acids was then observed with ESI-MS to evaluate the interaction between the tripeptide and the amino acids. GSH was able to form complexes with eight common amino acids which had implications for GSH and amino acid interactions in human blood.[45] Collision induced dissociation (CID-MS) was also used to generate GSH cations using two different chemicals. The two different methods of ionization provided spectra that were very similar to one another. Computational calculations were also done to show the capability of the radical to move around in the tripeptide, and the energy barriers associated with the radical migration.[46] IRMPD is a quickly evolving technique that is now being applied to investigating the structure of protonated peptides. Two residues of GSH (Cys and Gly) were studied using IRMPD spectroscopy in the

CysGly and GlyCys conformation by Osburn. Migration of the radical between the two dipeptides was studied by obtaining the IR spectra of the samples and comparing them to calculated IR spectra. The migration of the hydrogen atom from the sulphur atom to the α-carbon is not favoured for the Cys-Gly conformer, but does occur for the Gly-Cys conformer as supported by the IR spectra and DFT calculations. These results could then lend themselves well to the full GSH peptide when trying to understand the radical movement.[47] Recently, the reactivity and migration of the radical in the full GSH tripeptide was studied by Osburn as well. The results of the protonated tripeptide radical migration were inconclusive as the absorption region being analyzed contained too many overlapping broad peaks. The spectrum for the Glu-Cys dipeptide did show migration of the sulphur radical to the α-carbon of the Glu residue, which was expected based on their previous study. The region of 2800-3700 cm^{-1} in the spectrum for the tripeptide suggested that the radical does migrate from the sulphur atom to the α-carbon as hypothesized from their previous study.[48]

2.2 Glutathione (GSH) Experimental Procedures

2.2.1 Synthesis of Glutathione

Synthesis of the tripeptide glutathione was carried out using methods outlined and adapted from Kates and Albericio[12] and Chan and White[13]. The amino acids that were used in this synthesis were: Fmoc-Glu(ODmab)-OH, Fmoc-Cys(Mmt)-OH and Fmoc-Gly-OH, and the coupling reagents were HBTU and HOBt. Synthesis was carried out in a custom designed glass vessel, known as a Merrifield vessel, which is 10 cm long, has a fritted disk at the bottom and a 2 mm glass stopcock in the top.

Synthesis began by weighing 0.5022g of Wang-OH resin (200-400 mesh, 0.6-1.0 mmol/g substitution) into the Merrifield vessel and was shaken for one hour in DCM. The average molar substitution of the resin was used to determine the amounts of each reagent needed for the coupling of Fmoc-Gly-OH. This first coupling step required, 0.5973 g of Fmoc-Gly-OH, 287.8 μL of DCBC, 267.7 μL of pyridine and 10 mL of DCM. These reagents were added directly into the vessel and allowed to shake for a minimum of 18 hours. After this attempt, the resin was washed and the coupling success was measured (using methods outlined in section 1.3 and for each substitution check in this synthesis), but the substitution was only 0.6126 mmol/g (76.5% of the previously available sites). This coupling was less than ideal, so a solvent change was implemented and DMF was used instead of DCM. After 4 more attempts of using DMF with the same amounts of reagents, the final coupling for Fmoc-Gly-OH (G3) was measured to be 0.6802 mmol/g (85.0% of the previously available sites). The coupling of this amino acid was successful and allowed for the addition of the next amino acid, Fmoc-Cys(Mmt)-OH (C2), to be carried out. The resin-bound amino acid was treated with

approximately 10 mL of 20% piperidine in DMF to cleave the Fmoc protecting group. After a few cleavage attempts, the substitution check for Fmoc removal showed a negative value for the absorbance, and the chain elongation could proceed with the next amino acid.

For this synthesis, 4 molar equivalents of the amino acid (C2), HBTU and HOBt were used, along with 1.5 molar equivalents of DIPEA for each coupling. The amounts of reagents needed for the second amino acid coupling were based off the molar substitution of the previous amino acid, hence 0.8413 g of Fmoc-Cys(Mmt)-OH, 0.5183 g of HBTU, 0.2092 g HOBt and 89.7 μL of DIPEA were required. These reagents were weighed into a round bottom flask and dissolved in approximately 15 mL of DMF, then agitated for 5 minutes to dissolve the contents and pre-activate the mixture. The mixture in the round bottom flask was added to the Merrifield vessel that contained the glycine bound amino acid and a very minimal amount of solvent. The mixture of reagents was then allowed to react for a minimum of 18 hours. After this 18 hour period, the resin was washed 5 times with DMF to remove excess reagents. Next, a small amount of the resin was removed and dried in a fume hood for 45 minutes. The substitution check was then done to determine the degree of coupling of C2 to G3; after 5 attempts, the coupling of Fmoc-Cys(Mmt)-OH was sufficient to continue the synthesis of glutathione. The substitution of C2 to G3 was calculated to be 0.4891 mmol/g (71.9% of the previously available sites). Even with multiple attempts to couple C2 to G3, the ideal substitution of 80% was not met, but all previous attempts up to the fifth coupling attempt did not show much change in the substitution. At this point, the Fmoc protecting group on C2 was cleaved with 20% piperidine in DMF.

The coupling of the final amino acid, Fmoc-Glu(ODmab)-OH (E1), was done using 0.6688 g of the Fmoc protected amino acid, 0.3726 g of HBTU, 0.1504 g HOBt, and 64.5 μL

DIPEA. These reagents were weighed into a round bottom flask, dissolved in approximately 15 mL of DMF, and agitated for approximately 5 minutes to pre-activate the mixture. This solution was then added to the Merrifield vessel and was allowed to shake for a minimum of 18 hours. After 18 hours, the resin was washed 5 times with DMF to remove any excess or unreacted material. A small amount of resin was removed and dried to determine the coupling success of E1 to the previous amino acid C2. This coupling took three attempts to get the final substitution of 0.3987 mmol g^{-1} and marked the end of the chain elongation of glutathione.

Even though the tripeptide was completely synthesized, the protecting groups on C2 and E1 had to be removed. The first protecting group that needed to be cleaved was the Fmoc protecting group on the α amino group of E1. This was done by treating the resin with 20% piperidine in DMF and shaking the vessel for 30 minutes. After 30 minutes, the resin was washed 5 times with DMF to remove dibenzylfulvene adduct in solution. A small portion of the resin was removed and dried in a fume hood to determine if every Fmoc protecting group was cleaved. The sample was then diluted and checked in the UV spectrophotometer, and when a negative absorbance was observed, it was assumed the Fmoc was completely removed. Cleavage of the ODmab protecting group from E1 was done by treating the resin-bound peptide with 2% hydrazine in DMF 10 times. Each time the resin was treated with the hydrazine mix, the vessel was shaken for approximately two minutes and then washed through with DMF 4-5 times. Following this cleavage, the α carboxyl group of E1 was free from the effects of the protecting group. The protecting group on C2, Mmt, had to be cleaved using a 1:5:94 mix of TFA:TIS:DCM. This solvent mixture was added to the vessel and agitated for 30 minutes. Following the shaking, the resin-bound peptide was washed 5 times with DMF to remove the Mmt protecting group. The resin was washed through with DMF 4 more times to ensure that no

free protecting groups, unreacted reagents or other impurities remained in the vessel. The final product obtained was resin-bound glutathione free from the effects of protecting groups. The product was removed from the vessel and placed in a fume hood until it was completely dry and subsequently stored in a vial that was put into a desiccator.

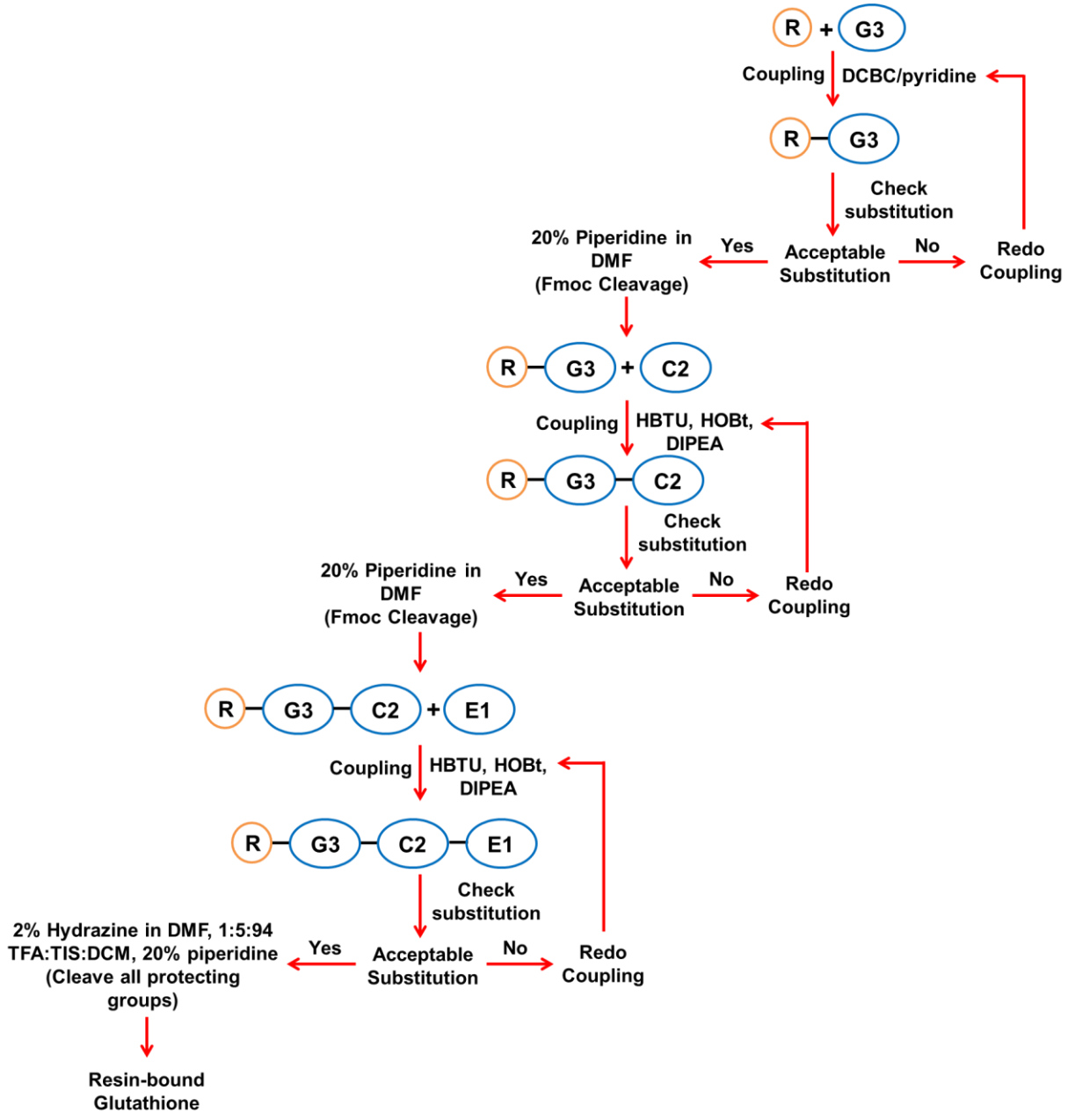


Figure 2.2.1: Flow chart for the synthesis of the resin-bound peptide glutathione. Methods used in this synthesis were adapted from Kates and Alberts[12] and Chan and White[13]. 'R' represents the Wang resin in which the peptide was synthesized on.

2.2.2 NMR Experiments Conducted on Glutathione

All NMR experiments conducted on glutathione were done at the University of Waterloo using a Bruker Avance 500 MHz NMR spectrometer equipped with a Doty Scientific triple-tuned DSI-XC4 HRMAS probe at ambient temperature (~295 K). The probe was equipped with a ^{13}C-^{2}H LF Trap Wand, ^{13}C-^{2}H MF Tune/Voltage Division Wand, ^{13}C-^{2}H LF Tune/Voltage Division Wand and a ^{13}C-^{1}H capacitor. A small amount of the dry resin-bound peptide was loaded into a XC4 40 μL Kel-f sealing cell, which filled approximately half the volume of the cell. The rest of the cell was filled with DMF-d_7 and a XC4 nominal Teflon plug was used to cap it. The cell could only be used reliably for spectroscopy if there were no air bubbles in it, so care was necessary in preparing the samples. The sealing cell was then loaded into a Doty Scientific XC4 4-mm silicon nitride thin wall rotor and capped with Torlon short turbine caps. The rotor was then placed into probe and spun at a speed of 4.5-4.6 kHz at the magic angle of 54.74°.

One-dimensional ^{1}H NMR spectra, plotted as spectral intensity versus frequency, were obtained using a simple one-pulse sequence. The spectrum was collected after 1024 transients, with an acquisition time of 0.62 seconds and a spectral width of 6613 Hz. A 90° ^{1}H pulse length of 7.75 μs at low power (0.00 dB) was used to obtain the spectrum. The resulting 8k data points were zero-filled to 16k data points for Fourier Transform and apodized using the exponential windowing function with a Lorentzian broadening constant of 0.30 Hz, sine bell shift of 0 and a Gaussian broadening factor of 0. The spectrum was referenced internally to the chemical shifts of the DMF solvent and applied to all other 1D and 2D ^{1}H experiments. All spectra were processed using the Bruker XWIN-NMR 3.5 pl6 software.

Two-dimensional ^{1}H-^{1}H COSY experiments used the cosyqf90 pulse sequence and collected a set of 256 x 1024 data points in the t_1 and t_2 dimensions, respectively. Each increment consisted of 128 transients with 4 dummy scans over a spectral width of 6009 Hz in both dimensions. A 90° pulse length of 7.55 μs was applied at a power level of 0.00 dB with an acquisition time of 0.085 seconds. Delay increments (IN0) of 0.166 ms, a dwell time of 83.2 μs and a relaxation delay of 2 seconds were also applied in the pulse sequence. After acquisition and Fourier transform was applied, the data sets underwent apodization using the SINE multiplication function with a Lorentzian broadening constant of 0.30 and 1.00 Hz in the t_1 and t_2 dimensions, respectively. In both dimensions, the sine bell shift was set to 0 and the Gaussian broadening factor was set to 0.1 in the t_1 dimension and 0 in the t_2 dimension. The COSY spectrum was calibrated using the value obtained from the 1D ^{1}H experiment for both the t_1 and t_2 dimensions. The spectrum was processed using the Bruker XWIN-NMR 3.5 pl6 software.

Two-dimensional ^{1}H-^{1}H TOCSY experiments used the DIPSI-2 pulse sequence and were collected in a set of 512 x 4096 data points in the t_1 and t_2 dimensions, respectively. A total of 56 transients and 16 dummy scans were used per increment over a spectral width of 5000 Hz in both dimensions. A 90° pulse length of 8.00 μs was used at a power level of 0.00 dB and 90° pulse at low power of 70 μsec was accompanied by a 19.12 dB power level for the TOCSY-spin lock. A TOCSY mixing time of 35 ms and an acquisition time of 0.40965 seconds were used in the pulse sequence as well. Delay increments of 0.199 ms, a dwell time of 100 μs, and a relaxation delay of 2 seconds were also used for this experiment. Upon Fourier transform, the data sets were apodized using the QSINE multiplication function with a Lorentzian broadening factor of 0.30 and 1.00 Hz in the t_1 and t_2 dimensions, respectively. In both dimensions, a sine bell shift of 2 was applied and the Gaussian broadening factor was set to 0.1

in the t_1 dimension and 0 in the t_2 dimension. The TOCSY spectrum was referenced using the value obtained from the 1D ^{1}H experiments on this sample. The spectrum was processed using Bruker XWIN-NMR 3.5 pl6 software.

Two-dimension ^{1}H-^{1}H ROESY experiments utilized the roesyph pulse sequence and collected a set of 512 x 2048 data points in the t_1 and t_2 dimensions respectively. A total of 64 transients and 16 dummy scans per increment were used over a spectral width of 5000 Hz in both dimensions. A 90° pulse length of 8.00 μs at a power level of 0.00 dB along with a 50 ms pulse for the ROESY spin-lock at a power level of 28.02 dB. Each pulse had an acquisition time of 0.2049 seconds and a 20 μs power switching delay. Upon Fourier transform, the data sets were apodized using the QSINE multiplication function with a Lorentzian broadening factor of 0.30 and 2.00 Hz in the t_1 and t_2 dimensions, respectively. A sine bell shift of 2 was applied in both dimensions and a Gaussian broadening factor of 0.1 in the t_1 and 0.05 was used in t_2 dimension. The spectrum was referenced to the same solvent reference as the 1D ^{1}H experiments for this sample. Processing of the ROESY experiments was done using the Bruker XWIN-NMR 3.5 pl6 software.

2.2.3 Infrared Multiple Photon Dissociation Acquisition of Glutathione

Reduced L-Glutathione was purchased from Sigma-Aldrich for use in the IRMPD experiments. All IRMPD experiments were performed at the Centre de Laser Infrarouge d'Orsay (CLIO) in Orsay, France by members of the McMahon group. The free electron laser (FEL) was used for the experiments and was produced using a 10-50 MeV electron beam. The FEL beam was guided into a Bruker FT MS APEX-Qe 9.4T FT-ICR mass spectrometer which was coupled to the apparatus.

The experiment was scanned continuously over the range of 900-2000 cm^{-1}. The FEL output consisted of 8 μs macropulses composed of 500 micropulses with a repetition rate of 25 Hz. An average IR power of 500 mW had macro- and micropulse powers of approximately 20 mJ and 40 μJ, respectively. The spectrum was reported as frequency versus IRMPD efficiency. The spectrum was calibrated using phosphotyrosine as a reference compound.[49] The calculated protonated structures of GSH were created by adding a proton to each of the heteroatoms and then performing a DFT calculation using the DFT/B3LYP method with a 6-31g+(d,p) basis set. This was done for each site of protonation and with different orientations of the proton around each site. The calculations resulted in a calculated structure of protonated glutathione and a calculated spectrum. The calculated spectra were adjusted using scaling factor of 0.965 for B3LYP/6-31g level of theory.[6]

2.2.4 Computational Calculations of Glutathione

Initially glutathione and DMF structures were created in GaussView 5.0 and DFT calculations were performed to optimize the structures with the aid of the Gaussian 09W program. The dielectric constant of DMF needed to be determined to be applied as a solvent box around glutathione in its calculation. Using the DFT/B3LYP method with a 6-31g+(d,p) basis set, DMF was optimized and the value for the dielectric constant that was obtained was used as a constant to surround the GSH molecules in further calculations. Another DFT/B3LYP method calculation with a 6-31g+(d,p) basis set and the PCM model for solvation was performed on the straight chain conformer of GSH. A single structure was obtained from the DFT calculation and used to explore the different conformations of GSH. A large number of unique structures needed to be generated from the initial conformer to create a large sample size of candidate structures. The approach used would generate a large number of structures that

when optimized, would provide the different possible conformations in which GSH may be found in.

To do this, the first step was to generate the AMBER force field for the GSH molecule. This required that the partial charges of each atom in the GSH molecule be determined so they could be applied to the molecular mechanic calculations in the next steps. A fitting algorithm known as ChelpG in Gaussian 09W was used to aid in this determination. Since specific bond lengths, angles and torsion angles are already known for particular atom combinations, the partial charges calculated would be the only other variables needed to create the AMBER force field.

Next, the large sample size of structures needed to be generated, and this required the DFT GSH structure to be physically altered in some systematic way. For instance, the bond lengths between atoms were altered by 1 Å or the torsion angle between four different atoms was changed by 10°. Using the molecular mechanics method and the custom AMBER force field, the altered structure was subjected to a geometry minimization along with a frequency calculation. In total, a set of 7 000 structures was generated and the accompanying MM calculation was performed to give a large collection of GSH molecular conformations. The output files of these MM calculations contained optimized structures of the GSH molecule along with the energies (in Hartrees) of these optimizations. All calculations were done using the Gaussian 09W program.

Further calculations needed to be done on GSH to simulate the effect that the linker would have on the conformation of the peptide. Molecular dynamic experiments were conducted to calculate the best angles for the peptide and the linker. The AMBER parameters

and partial charges that were generated for the molecular mechanics portions of this study had to be used for the molecular dynamic experiments as well. Some of the AMBER atom types that were in the linker were not found when the calculations were performed and atom types had to be created, based on existing atom types in the AMBER library. With the help of the Roy group, a python script was written to run these simulations using Python 2.7.5 and MMTK (molecular modelling tool kit). The script was run with a temperature of 298.15 K, 1 fs timesteps, 100000 equilibration steps, 10000000 production steps and outputting information every 100 steps. Once the simulation was complete, a text file with all the dihedral angles of the peptide was generated and converted into a histogram.

Using the free GSH molecule generated from the DFT calculation, a linker was added to acquire a structure that would be optimized to best represent the resin-bound peptide. The calculation was done with the DFT level of theory calculation using the RB3LYP method with a 6-31g+(d,p) basis set and the PCM model for solvation. This calculation was done using the Gaussian 09W program.

2.3 Glutathione Results

2.3.1 Results of Synthesis

The results of the synthesis of the tripeptide are summarized in Table 2.3.1. Coupling of the Fmoc-Gly-OH (G3) amino acid to the Wang resin was initially done using the DCBC method with pyridine and DCM as the solvent. The results from this step were not close to ideal with only a 0.6126 mmol/g coupling (76.5% of the previously available sites). A change in solvent was implemented and DMF was the solvent to be used for all subsequent couplings. After a few more attempts, the first amino acid coupling was measured to be 0.6802 mmol/g (85.0% of the previously available sites). After a third and fourth attempt, not much change was seen in the molar substitution of G3, therefore it was considered the end point of the first coupling. The resin-bound amino acid was then treated with acetic anhydride to cap any unreacted sites on the resin before the Fmoc protecting group was removed. Removal of the Fmoc protecting group was done with 20% piperidine in DMF and when the sample was measured in the UV spectrophotometer the absorbance showed a negative value.

Coupling of the second amino acid, Fmoc-Cys(Mmt)-OH (C2), was done using HBTU, HOBt and DIPEA as coupling reagents. The first three attempts for coupling this amino acid in DMF were very poor, approximately 0.4 mmol/g (58.8% of the previously available sites). A change from 15 mL of DMF to 10 mL was implemented to create a more concentrated mixture. The results of the next two attempts had greatly improved the results, increasing the substitution to 0.4891 mmol/g (71.9% coupling of the previously available sites). All of the Fmoc-Cys(Mmt)-OH that was purchased was used in the five coupling attempts, so no further attempts to try to achieve the 80% benchmark were done. Fmoc was cleaved using 20%

piperidine in DMF and when the sample was tested in the UV spectrophotometer the absorbance was negative.

Table 2.3.1: The coupling conditions and substitution needed for each attempt in the solid-phase peptide synthesis of the tripeptide, glutathione. The percentage of the total available sites was based off the initial molar substitution of 0.80 mmol/g.

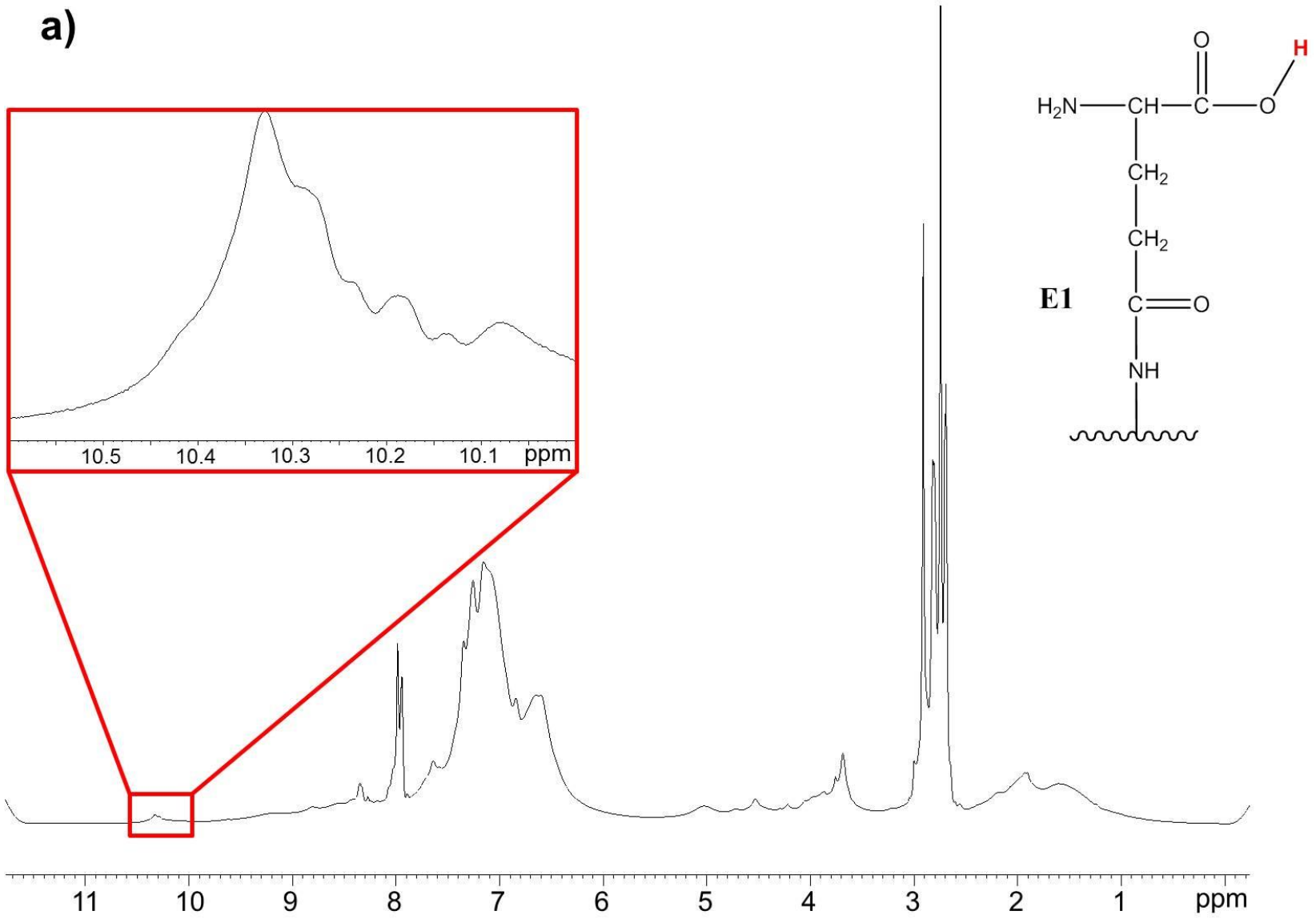
Amino Acid	Coupling Method	Solvent	# of Coupling Attempts	Resin Substitution (mmol/g)	% Subs of Previous Available Sites	% Subs of Total Available Sites
1st: Gly	DCBC	DCM	1	0.6126	76.5%	76.5%
		DMF	4	0.6802	85.0%	85.0%
2nd: Cys(Mmt)	HBTU/HOBt	DMF	5	0.4891	71.9%	61.1%
3rd: Glu(ODmab)	HBTU/HOBt	DMF	3	0.3987	81.5%	49.8%

Coupling of the final amino acid Fmoc-Glu(ODmab)-OH (E1) was done only three times with HBTU, HOBt and DIPEA to achieve a final substitution of 0.3987 mmol/g (81.5% of the previously available sites). The tripeptide glutathione had been fully synthesized, but the protecting groups on E1 and C2 had to be removed. Before removal of the protecting groups, a portion of the resin was taken out from the Merrifield vessel and allowed to dry in a fume hood for a minimum of 24 hours. Cleavage of the protecting groups began with the ODmab group on E1 using 2% hydrazine in DMF, then the Mmt group on C2 with a mix of TFA:TIS:DCM (in a 1:5:94 ratio). The peptide substitution was checked again to ensure that nothing had changed when the protecting group cleavage was done. Finally the Fmoc protecting group on E1 was cleaved with 20% piperidine in DMF and checked in the UV spectrophotometer to have a zero or negative absorbance. The final product obtained was a resin-bound glutathione free from the

effects of protecting groups. The fully synthesized deprotected resin-bound tripeptide was removed from the Merrifield vessel and allowed to dry for a minimum of 24 hours before analysis.

2.3.2 NMR Analysis of Glutathione

Proton assignment of the tripeptide was carried out with the aid of 1D ^{1}H and ^{1}H-^{1}H 2D experiments. The use of COSY allows for interactions between protons that are *J*-coupled three bonds or less from one another to produce a signal known as a "cross-peak" in the 2D spectrum. TOCSY also shows interactions between protons that are within the same spin system, but not restricted to the three bond limitation, which is one of the drawbacks of the COSY experiment. These assignments are summarized in Table 2.3.2 and are consistent with the 1D ^{1}H spectrum. Using the COSY and TOCSY spectrum (shown in Figure 2.3.2), three different spin systems were identified; all three had an amide proton interacting with an α proton, two systems had an α interacting with a β, but only one system had interactions with γ protons. One shift that was not seen in the COSY or TOCSY and was assigned using the 1D spectrum was the proton on the α carboxyl group of E1 (Figure 2.3.1). This peak was first thought to be a spinning side band, but after conducting the 1D experiment at different spinning speeds the peak was determined to be that of the α carboxyl proton of E1. The first experiment was done by spinning the sample at 4.6 kHz and acquiring a 1D spectrum of the sample and the second experiment was spun at 3.9 kHz. In both cases, a small peak suspected to be the proton on the carboxyl group was seen at 10.28 ppm. If the peak was not at 10.28 ppm for both spinning speeds, the peak would be considered a spinning side-band, but this was not the case so the peak was real.


a)
10.5
10.4
10.3
10.2
10.1
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
E1

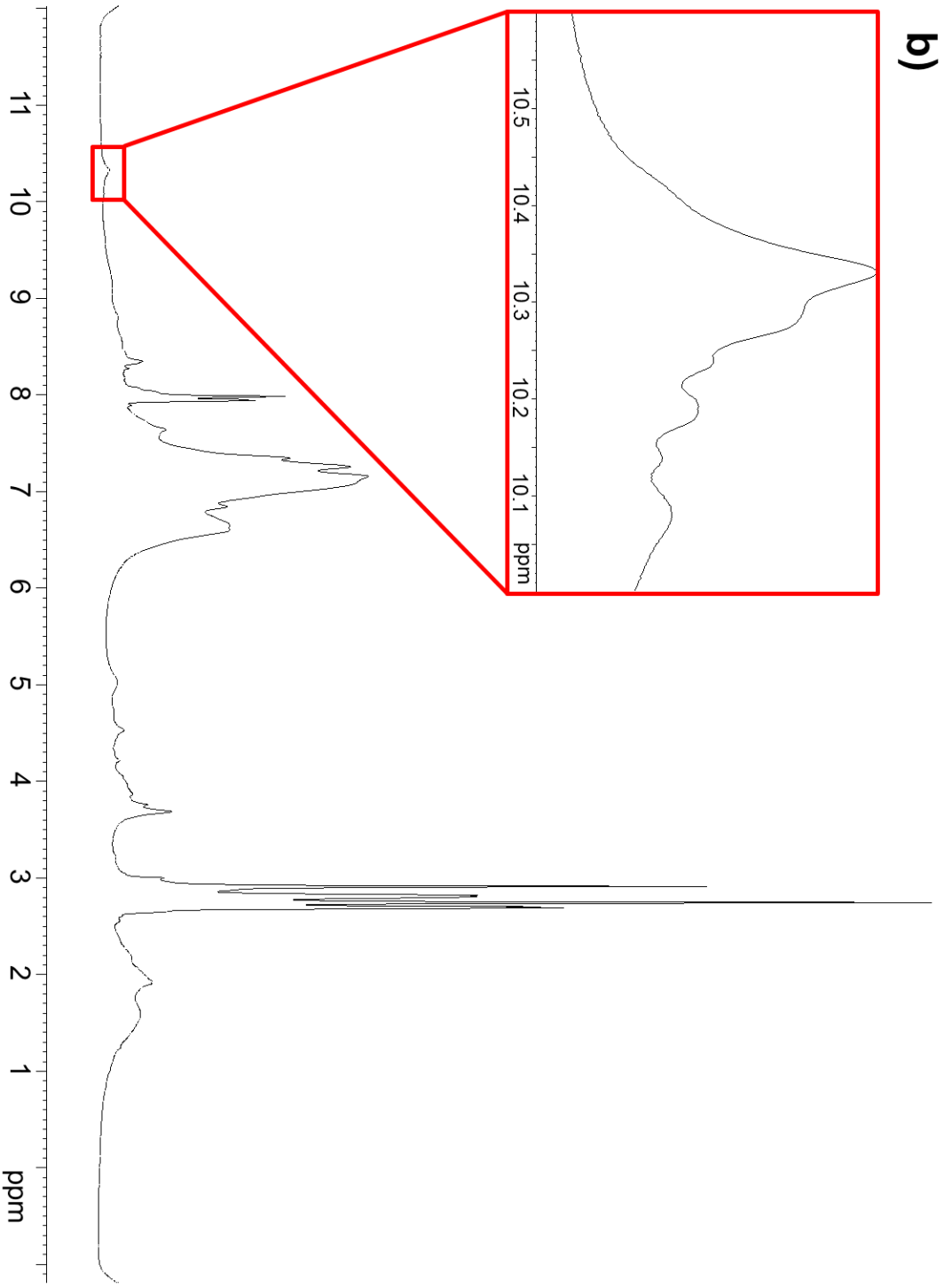


Figure 2.3.1: Comparison of 1D spectra for confirmation of the α carboxyl proton of E1 tested at two different spinning speeds. A chemical structure of the E1 residue portion is provided in the first spectrum and the proton bolded in red is the carboxyl proton of interest for both spectra. The inset in both spectra provides an expansion of the region between 10.6 ppm and 10.0 ppm to show the peak for the carboxyl proton: a) the sample was spun at 4.6 kHz, b) sample was

spun at 3.9 kHz, much lower than the first experiment to test if the peak in question was a real peak or a spinning side band.

Table 2.3.2: ^{1}H NMR chemical shift assignments of the resin-bound GSH peptide referenced internally to the solvent peaks of DMF-d_7

Amino Acid	NH	Hα	Hβ	Hγ	Other
E1	8.47	4.75	2.72 2.86	2.98 3.09	10.28 (α carboxyl)
C2	8.50	4.48	2.61 2.67	-	-
G3	8.31	3.93	-	-	-

Assignment of these spin systems to the individual amino acids was very simple, since G3 has no β protons and C2 has no γ protons. Therefore, the one system that had only an amide and α alpha interaction belonged to G3 and the system that had an amide, α and β belonged to C2. The α protons of glycine normally exhibit a chemical shift near 4.0 ppm, which is seen in the COSY and TOCSY spectra. There was a strong correlation between the amide proton of G1 (approximately 8.3 ppm) and the α proton (approximately 3.9 ppm). These were the only interactions between protons with chemical shifts at these values. It was concluded that this interaction between amide and α proton belonged to G1. A single spin system had an α proton correlating to a β proton, but no other γ protons was determined to belong to protons of C2. The amide and α protons show a strong interaction at approximately 8.5 ppm and 4.5 ppm. An interaction between the C2 α proton and two β protons was also seen in the "α region" of the COSY and TOCSY spectra. This α proton interaction with two β protons is not uncommon and is a result of the β protons of the β-methylene group being chemically inequivalent to one another. With this chemical inequivalency, these β protons have individually interacted with the α proton each producing their own cross-peak in the spectra. Another interaction seen in the

"amide region" of the TOCSY is one between the backbone amide proton of C2 and one of the β protons in the side chain. This interaction is greater than the COSY three-bond limitation between protons and was only observed in the TOCSY spectrum and provides additional proof that these set of cross-peaks belongs to the C2 system. The final amide and α proton interaction, along with all other protons in this system, belonged to the E1 residue. This spin system showed a strong correlation between an amide proton and an α proton at approximately 8.4 ppm and 4.8 ppm. The α proton then showed an interaction between 4 different signals in the "α region" representing the β and γ protons. The four signals that were seen interacting with the α proton were at 2.7, 2.9, 3.0 and 3.1 ppm. The two signals that are the most upfield were determined to belong to the β protons of C2 and the more downfield protons belonged to the γ protons. The environment of the γ protons causes them to appear more downfield; this is due to the carbonyl group that is adjacent to these protons and its proximity to the backbone C2 amide functionality. The circulation of the electrons within the carbonyl group creates a deshielding effect that is felt by the γ protons of E1. Also, these protons appear inequivalent to one another even though they are on the same carbon, giving rise to the two peaks interacting with this α proton. The β protons are not as deshielded as the γ protons since they are further from the carbonyl group, but they do experience chemical shift inequivalency due to the two different signals interacting with the α proton. The protons at approximately 3.0 and 3.1 ppm belong to the γ protons of E1 and the remaining two peaks (2.7 and 2.9 ppm) belong to the β protons of E1. Using COSY and TOCSY, the three different spin systems of GSH were identified and the individual protons for each system were able to be assigned to their respective chemical shifts.

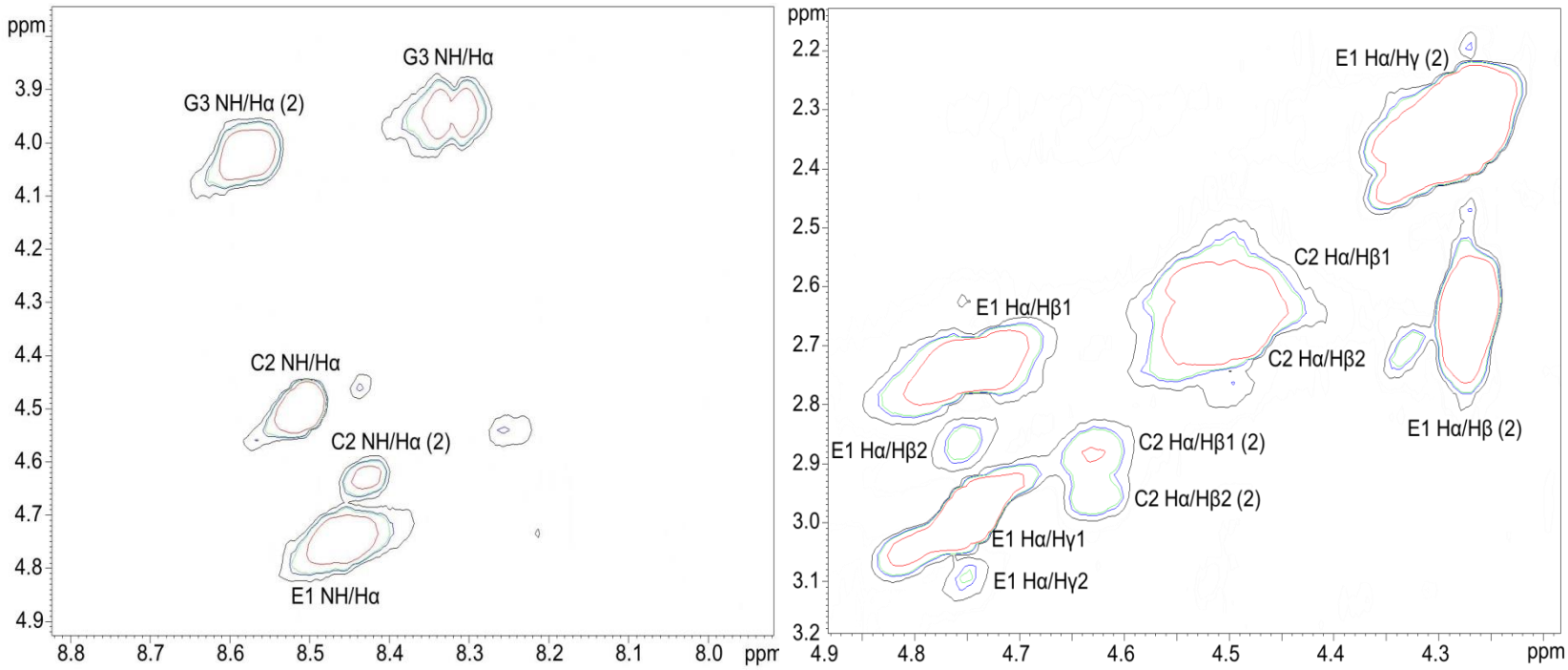


Figure 2.3.2: Expanded views of the TOCSY spectrum for the tripeptide, GSH. Glycine, cysteine and glutamate are represented as G3, C2 and E1 respectively. A) "Amide region" of the peptide shows the three different spin systems in which the α proton is interaction with an amide proton. B) Expanded view of the "Alpha region of the spectrum shows interactions between α protons and other protons in side chains, i.e. β and γ protons. One spin system lacked a β and γ interaction and another system lacked a γ interaction. Determination of correlation between the spin systems and the appropriate residues was simple with this known information. A set of secondary shifts were identified and summarized in Table 2.3.3 and are denoted with a (2). The full TOCSY spectrum can be seen in Appendix.

A set of secondary shifts was also seen for this tripeptide (Table 2.3.3), but were not as informative as the primary shifts (labelled with a (2) in Figure 2.3.2). During the synthesis, the peptide may have adopted a structure different that the most predominant conformer that is seen. A secondary shift was seen for G3, where an amide and α proton were interacting near 8.6 ppm and 4.0 ppm. This α shift is very characteristic of glycine residues and is very distinct from the first G3 interaction. This α proton also shows no interactions with any other protons, so it was confirmed that it was indeed a secondary shift for the G3 residue. There was also a secondary set of peaks for the C2 residue for each of the protons. There were some weak intensity peaks for the amide and α protons interacting with one another, and that α proton was interacting with two β protons again. These β protons did not have as high an intensity as the protons for the other conformer, but they were present in the "α region". These α and β protons also showed no interactions with any other protons, such as a γ proton, therefore it was assumed that these peaks belonged to a secondary C2 spin system. The E1 spin system also exhibited a secondary set of peaks as well, but some of the interactions were more intense than expected. The amide and α protons show a weak interaction more upfield than the previous conformer and this α proton again interacts with four different protons. These four protons represent the β and γ protons, and they are seen interacting with the alternate α proton of E1. Using the same methods of deduction as the primary conformer, the signals that appeared further downfield belong to the γ protons and the remaining signals are for the β protons. Just as with the first conformer, a set of three different spin systems and the individual protons were assigned to each of the three amino acids using the COSY and TOCSY spectra.

Table 2.3.3: ^{1}H NMR chemical shift assignments for a secondary conformer of the resin-bound peptide GSH referenced internally to the solvent peaks of DMF-d_7

Secondary	NH	Hα	Hβ	Hγ	Other
E1	8.26	4.27	2.29 2.34	2.63 2.70	-
C2	8.43	4.62	2.88	-	-
G3	8.58	4.01	-	-	-

Assignment of the protons in the individual spins systems was complete and the next step was to use these shifts and examine the ROESY spectrum to determine the structure of the peptide. The use of the ROESY spectrum not only provides insight to the structure, but it also assists in determining if the synthesis of glutathione was done correctly. The ROESY spectrum (Figure 2.3.3) produced 20 interactions between protons through-space. These interactions shown in Table 2.3.4, are broken down into 14 that are within the same spin system and six which are between different amino acids. Of these interactions, two were used to establish the secondary structure of GSH.

Table 2.3.4: Tabulated list of NOE interactions for the resin-bound peptide GSH. The interactions listed in bold red are those between different amino acids. The ROESY spectrum was referenced internally to the solvent peaks of DMF-d_7.

Interaction #	Chemical Shift F1 (ppm)	Chemical Shift F2 (ppm)	Amino Acid 1	Amino Acid 2
1	**8.57**	**4.75**	**G3 NH (2)**	**E1 Hα**
2	**8.55**	**4.63**	**G3 NH (2)**	**C2 Hα (2)**
3	**8.31**	**4.50**	**G3 NH**	**C2 Hα**
4	8.50	4.50	C2 NH	C2 Hα
5	8.45	4.74	E1 NH	E1 Hα
6	8.31	3.94	G3 NH	G3 Hα
7	8.57	4.02	G3 NH (2)	G3 Hα (2)
8	8.45	2.98	E1 NH	E1 Hγ
9	8.45	2.71	E1 NH	E1 Hβ
10	**8.31**	**2.62**	**G3 NH**	**C2 Hβ**
11	8.50	2.64	C2 NH	C2 Hβ
12	4.74	2.99	E1 Hα	E1 Hγ
13	4.62	2.88	C2 Hα (2)	C2 Hβ (2)
14	4.71	2.84	E1 Hα	E1 Hβ
15	4.75	2.79	E1 Hα	E1 Hβ
16	4.75	2.74	E1 Hα	E1 Hβ
17	**4.74**	**2.69**	**E1 Hα**	**C2 Hβ**
18	4.49	2.67	C2 Hα	C2 Hβ
19	4.51	2.59	C2 Hα	C2 Hβ
20	**4.74**	**2.58**	**E1 Hα**	**C2 Hβ**

The 14 interactions that are within the same spin system were already seen in the previous COSY and TOCSY experiments and do not provide any additional information beyond what is already known. Each of the amide protons and α protons of the three peptides were seen interacting within their respective spin systems, C2 had interactions between the α and β protons, and E1 had many interactions within the residue as well between the α, β and γ protons. Interactions between sequential amino acids will help confirm that the synthesis was done correctly. Interactions 3 and 10 confirm that the synthesis between the first two amino acids

were successful. Interaction 3 is between the amide proton of G3 and the α proton of C2 and interaction 10 is between the G3 amide proton and the C2 β proton. The distance between these protons interacting is very minimal and is not surprising to see in the ROESY spectrum. These interactions provide good evidence that the first two amino acids were coupled correctly. Interactions 17 and 20 are between the E1 and C2 residues, but both are between the E1 α proton and the C2 β proton. If E1 and C2 we joined through the α amino and carboxyl group (such as the case for C2 and G3), these two protons interacting with one another would not be unusual, but because the linkage is between the γ carboxyl group of E1 and the α amino group of C2, interactions 17 and 20 become quite far in space. This interaction does suggest that the synthesis was done correctly, but it is an interaction for this specific linkage that is greater than the normal acceptable distance for ROESY interactions. Having this interaction present suggests that the E1 end of the peptide must be much closer in space than if it were in an extended state. Based on the through-space couplings from the ROESY spectrum, the secondary conformer of the peptide must have some sort of folding motif that allows the E1 α end to be closer to the middle of the peptide.

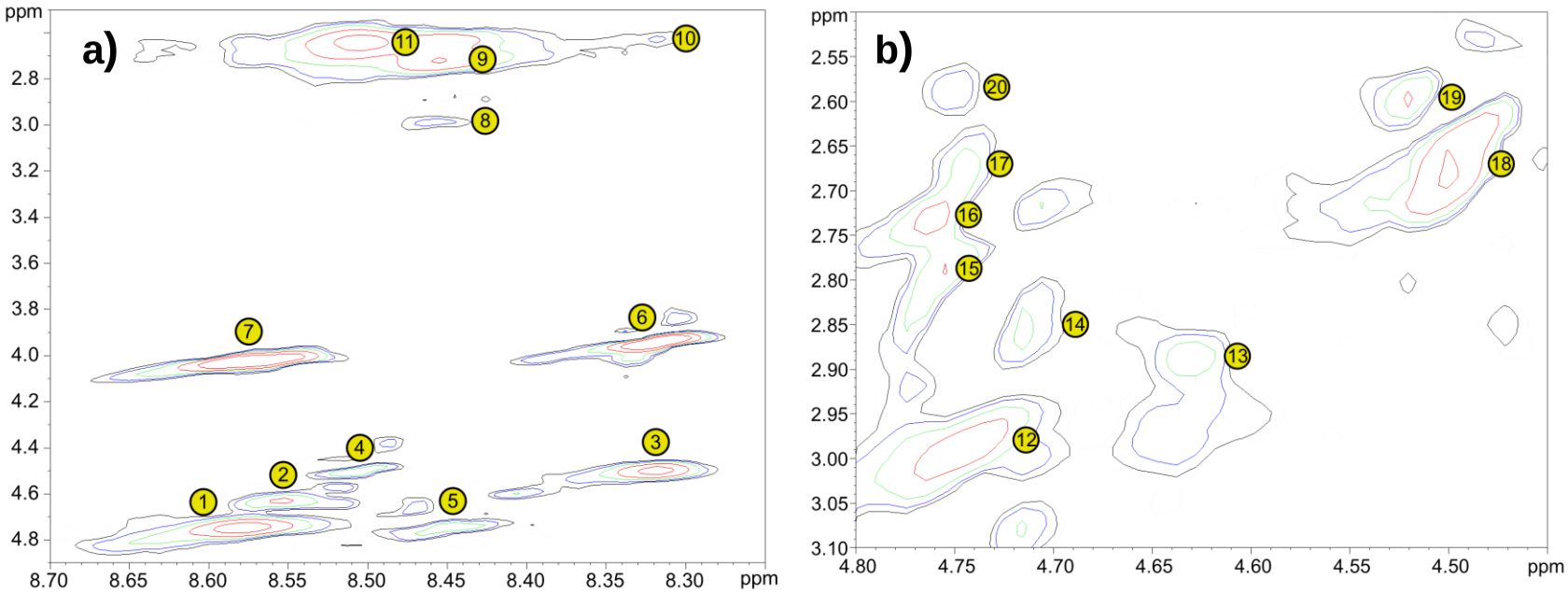


Figure 2.3.3: Portions of the ROESY spectrum for the tripeptide GSH. The numbers in the yellow circles correspond to the interactions in Table 2.3.4. A) The amide region of the ROESY spectrum shows 11 different interactions between the residues. B) The α region of the spectrum shows 9 different interactions. In total 20 interactions were observed with 4 of them being between different amino acids. These interactions were used to confirm that the synthesis was done correctly and to hypothesize a structure for this resin-bound tripeptide. The full ROESY spectrum can be seen in the Appendix.

By combining all of the ROESY interactions, a three-dimensional structure of GSH can be proposed. Interactions between the α proton of E1 and the β protons of C2 suggest that this end of the tripeptide is folded in towards the middle. The G3 and C2 interaction just suggests that the two residues are close to one another and potentially exist in a linear state. The ROESY results with the two key interactions do suggest that there is some folding involved with the structure of the peptide. The lack of a G3-E1 interaction could suggest the peptide is in a semi-folded state, but it's possible that interaction was not seen in the ROESY. A folded or semi-folded state is consistent with what Zhang *et al.* suggested in their NMR and computational study.[43] The effect of the linker and DMF solvent play a role in the folding of this peptide that makes it differ from the solution results. A combination of computational calculations (DFT and molecular dynamics) can give insight into the structure that would be formed while the tripeptide is on the linker.

There were also three interactions that were between protons for the secondary conformer; interactions 2, 7 and 13. Interactions 7 and 13 are within the same spin system, where 7 was between the amide and α protons of G3 and 13 was between the α and β protons of C2. Interaction 2 was between sequential amino acids for this secondary conformer, G3 amide interacting with the C2 α. This single interaction does not give any insight into a possible secondary conformer of this tripeptide, it merely suggests that there first two amino acids are in the proper order. It is impossible to propose a structure for this GSH conformer without more interactions. An unusual interaction between a proton in the secondary set of shifts and the main set was observed. It is labelled as interaction 1 and is between the secondary amide shift of G3 and the α proton of E1 in the main set of shifts. It is odd such an interaction would arise, but this still would not give insight into the conformation of the secondary conformer or even the first.

2.3.3 Infrared Multiple Photo Dissociation Analysis of Glutathione

Upon acquiring the IRMPD spectrum of protonated GSH, a large number of protonated isomers were created at the various numbered positions as listed in Figure 2.3.4. The initial structure and all isomers were calculated using the B3LYP/6-31g(d) level of theory.

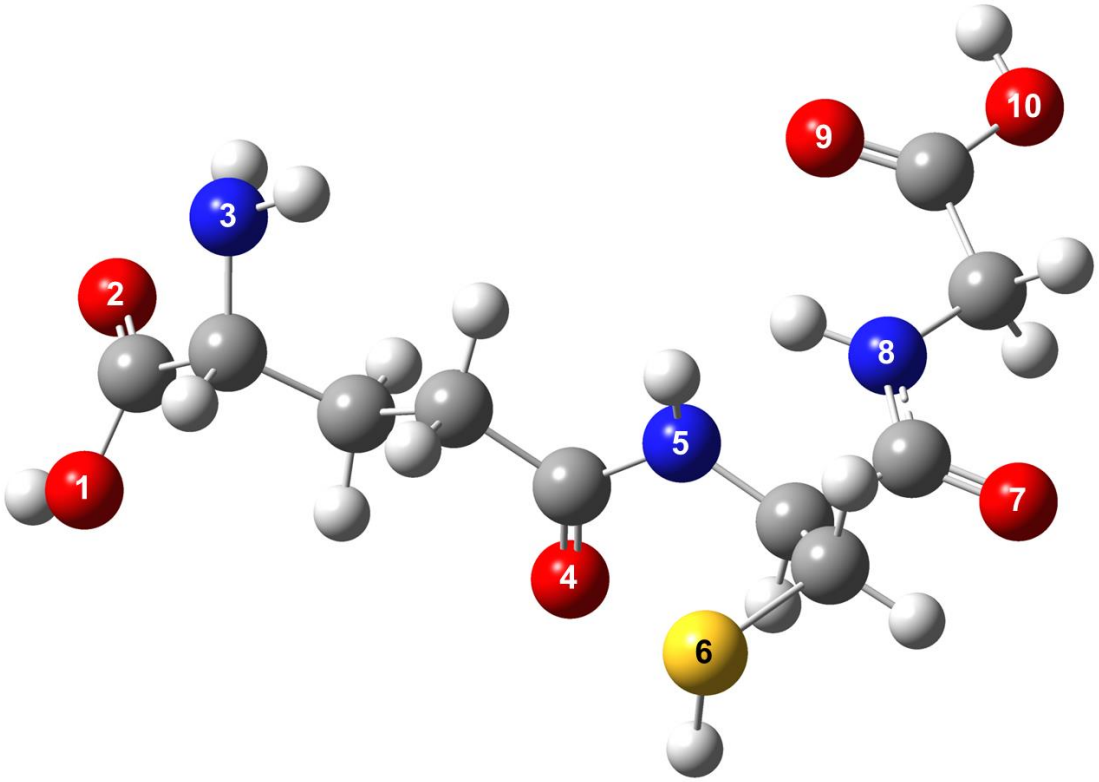


Figure 2.3.4: Optimized structure of gas-phase GSH using B3LYP/6-31g(d) level of theory. The possible sites of protonation are numbered and are referred in the interpretation and analysis of spectra. The carbon, oxygen, nitrogen, sulphur and hydrogen atoms are coloured in grey, red, blue, yellow and white, respectively.

Of all the calculated structures, five isomers had relatively similar calculated spectra that matched the experimental spectrum (Figure 2.3.5). The calculated spectra compared to the experimental spectrum can be seen in the Appendix. There were only three different sites of protonation for these isomers, but for three of the sites the orientation of the proton differed between isomers (Figure 2.3.6). The corresponding calculated energies are a sum of the electronic and thermal free energies and are shown along with the relative energies in Table 2.3.5.

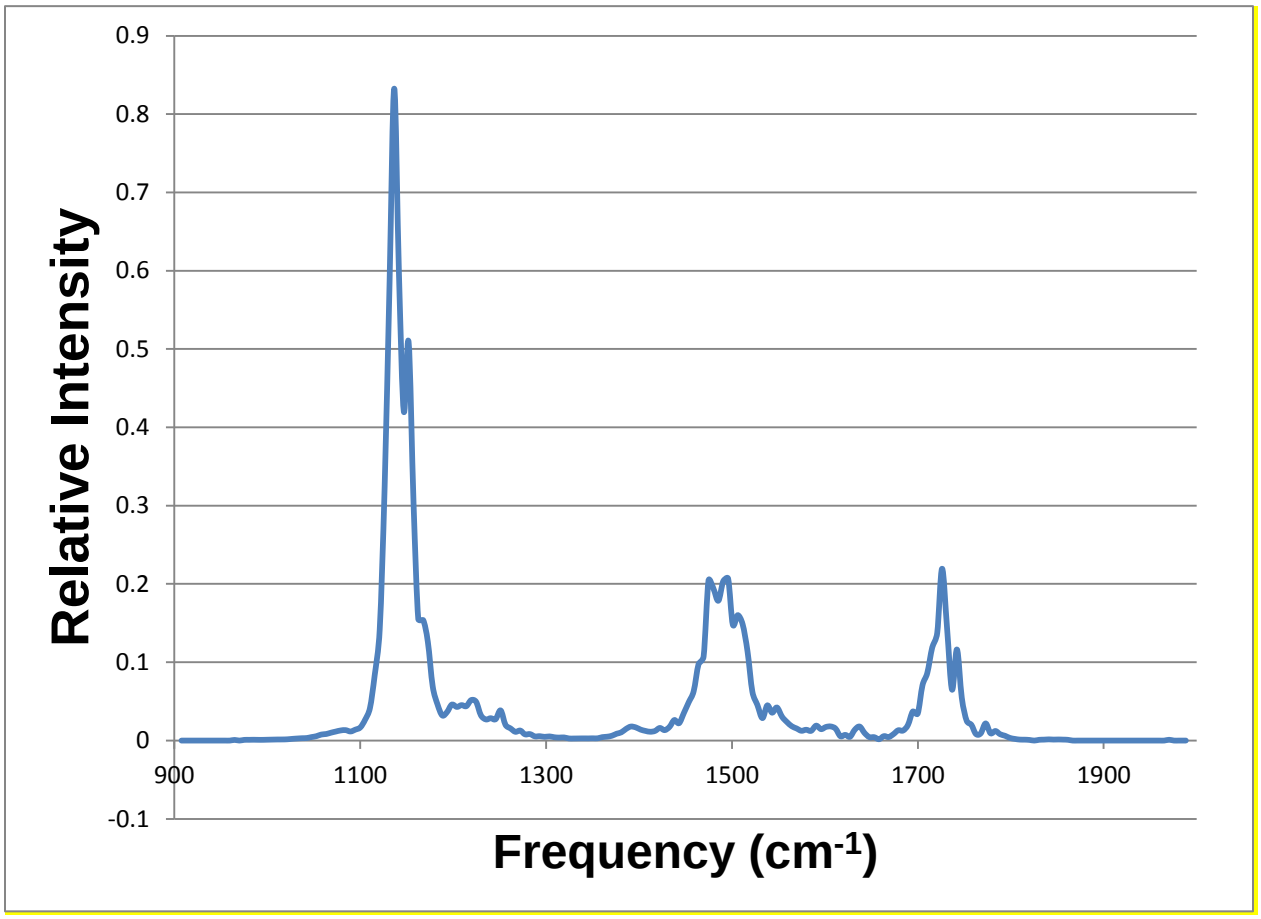


Figure 2.3.5: Experimental IRMPD spectrum of protonated glutathione. The intensities are reported as IRMPD efficiency versus the frequency of the fragment (in cm-1).

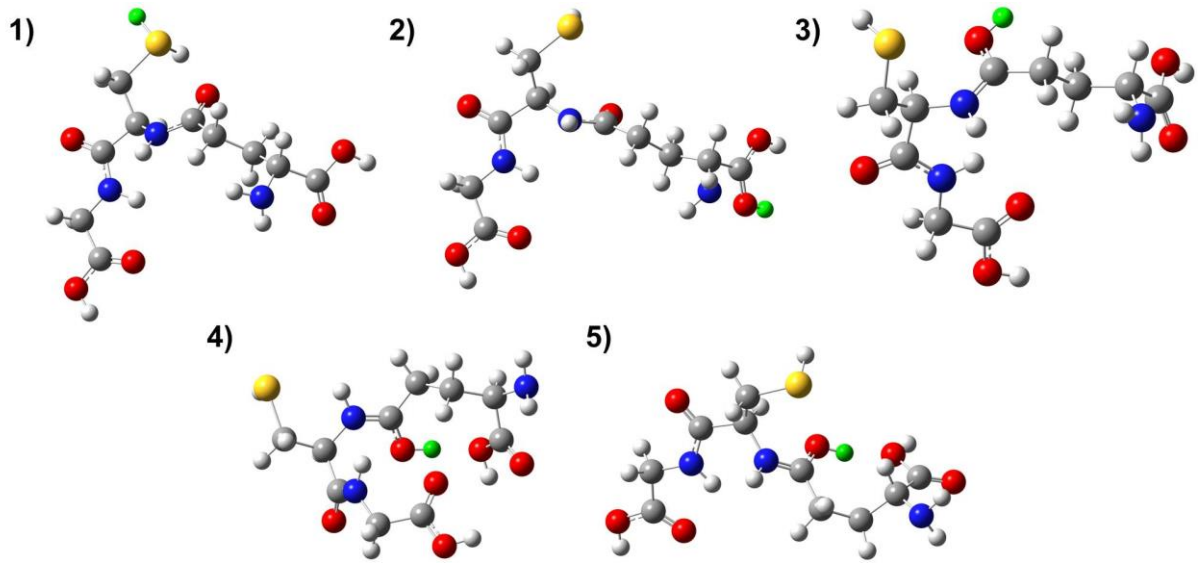


Figure 2.3.6: Five possible pronated species (labelled 1-5) of GSH that have calculated spectra that most resemble the experimental spectrum. The candidates were chosen based on the similarity between the calculated spectra and the experimental spectrum. Comparative studies are done to determine which protonated species would best represent the experimental spectrum. The hydrogen, carbon, oxygen, nitrogen and sulphur atoms are represented in white, grey, red, blue and yellow respectively. The extra proton that is added for each isomer is coloured green for easier visualization.

Table 2.3.5: Calculated free energies (kJ mol^{-1}) and relative free energies of the different isomer candidates for the GSH molecule. All structures were calculated using the B3LYP/6-31g(d) level of theory.

Structure Number	Σ Electronic and Thermal Free Energies (kJ mol^{-1})	Relative Free Energies (kJ mol^{-1})
4	-3690131	0
5	-3690128	+3
3	-3690090	+41
1	-3690027	+104
2	-3690005	+126

These five isomers were chosen based on the similarities between the experimental spectrum and the calculated spectra. The carbonyl frequencies of the experimental and calculated spectra are listed in Table 2.3.6. The frequencies that best match the experimental spectrum come from isomers 1, 2, and 5. The protonation on isomer 1 was on atom 6, for isomer 2 it was on atom 2 and for isomer 5 it was on atom 4. The mode of vibration for isomers 1 and 2 are the E1 γ carbonyl stretch and for isomer 5it is the C2 α carbonyl stretch. When comparing the differences in carbonyl frequencies to the experimental, isomers 1, 2, and 5 are only slightly red-shifted. The bond lengths between the carbon and the oxygen for these C2 carbonyls are fairly typical of carbonyl groups (1.22 Å); therefore a large discrepancy in the shifts is not expected. The vibrational mode for isomers 3 and 4 are a C2 carbonyl stretch and G3 carbonyl stretch, respectively. Protonation for isomers 3 and 4 both occurred on atom 4. The differences in frequencies for these two isomers compared to the experimental are very blue-shifted. A possible explanation for this large blue-shift is that the bond lengths between these atoms are slightly shorter than normal, which requires more energy to stretch. Therefore the frequency in which these modes occurred is going to appear higher than expected. From the carbonyl frequencies, isomers 1, 2, and 5 are the best matches to the experimental spectrum so far, but examination of the amide region and low frequency region will provide additional support to determine which isomer is the best fit for the experimental spectrum.

Table 2.3.6: Calculated carbonyl frequencies for the five possible isomers and the deviation from the experimental shift.

Isomer Number	Carbonyl Stretching Frequency (cm^{-1})	Vibrational Mode	Δ in Frequency from Experimental
Experimental	1726	-	-
1	1723	E1 γ Carbonyl	-3
2	1723	E1 γ Carbonyl	-3
3	1740	C2 α Carbonyl	+14
4	1752	G3 α Carbonyl	+26
5	1723	C2 α Carbonyl	-3

The carbonyl peaks in the calculated and experimental spectra were used to pick the most likely candidates for the protonated structure, but the set of peaks between 1800 cm^{-1} and 1700 cm^{-1} were used to provide additional evidence to identify the fitting isomer. Shifts for the peaks in the vicinity of 1700 cm^{-1} are reported in Table 2.3.7. For each of the isomers, the calculated spectra had at least one peak that matched very well to the experimental spectrum. Isomer 1 had two different vibrational modes for the two frequencies that matched best which were a C2 amide and G3 amide bend. Isomer 2 also had a C2 amide bend that matched exactly with the experimental spectrum, but the frequency at 1488 cm^{-1} was unusual. This mode was a concerted bend of the G3 and C2 amide groups, E1 α carboxyl, and protonated carbonyl group. This peak at 1488 cm^{-1} was red-shifted by a small amount, but it was the largest difference between frequencies for all the peaks that had matched well. Isomers 3, 4, and 5 all exhibited a G3 amide bend. Isomers 3 and 4 both had frequencies that matched very well with the experimental spectrum, but isomer 5 was slightly blue-shifted from the experimental. A possible explanation for this isomer to having a slightly blue-shifted peak could be a result of the slightly longer bond length between the G3 amide and C2carbonyl causes the energy required to stretch the

bond to be increased. Many of the frequencies in this region of the calculated spectra for all isomers matched well with the experimental spectrum, therefore eliminating certain structures from being the potential experimental protonated structure is difficult at this point. Analysis of the low frequency region of the different spectra can be used to narrow down the search for the best structure.

Table 2.3.7: Calculated amide/amino frequencies for the five possible isomers and the deviation from the experimental shift. There were multiple peaks in the amide and amino range, therefore the peak that was closest for the calculated spectra dictate the difference from the experimental.

Isomer Number	Amide/Amine Frequency (cm^{-1})	Vibrational Mode	Δ in Frequency from Experimental
Experimental	1474	-	-
	1495	-	-
	1506	-	-
1	1471	C2 NH Bend	-3
	1497	G3 NH Bend	+2
2	1488	Various Bends	-7
	1506	C2 NH Bend	0
3	1497	G3 NH Bend	+2
4	1505	G3 NH Bend	-1
5	1480	G3 NH Bend	+6

The final set of experimental peaks, that will help characterize the which isomer is the potential best fit, are in the range of approximately 1100 cm^{-1} to 1200 cm^{-1}. This range shows C-O stretching and bending motions, along with C-C stretch and bend, but in these calculated peptides two modes were found: C-OH bend and C-H twist. The data for this area in the calculated spectra are summarized in Table 2.3.8. Each of the isomers was either 5 or 6 cm^{-1} blue-shifted from the respective experimental peaks. Isomers 1 and 3 experienced a concerted twisting motion of the E1 side chain. This motion caused the β and γ carbons to twist along with

the C2 residue and α groups of E1. These shifts were only blue-shifted by 5 cm^{-1} and were very close to the experimental value. Isomers 4 and 5 had a vibrational mode that represented a G3 carboxyl bend. Both of these shifts were only blue-shifted by 6 cm^{-1}, which also was very similar to the experimental spectrum. Isomer 2 experienced a different vibrational motion at 1141 cm^{-1} than isomers 1 and 3. The motion was a COH bend of the E1 α carboxyl group for the OH group and the protonated carbonyl. As opposed to being the E1 side chain twisting that was seen at this frequency for other isomers, the protonated carbonyl and carboxyl group are both bending at the same time at this frequency. Isomer 2 also experienced the G3 amide bend at 1158 cm^{-1} just as isomers 4 and 5 did as well. There were no peaks that were deviated from the experimental spectrum by a large amount for this region either, so determination of which isomer would be represent the protonated GSH could come from the analysis of the calculated structures and energies.

Table 2.3.8: Calculated vibrational modes for experimental and calculated GSH species in the 1100 cm^{-1} to 1200 cm^{-1} range. All modes may not match up equally, therefore the deviation is calculated based on the peak it closest resembles.

Isomer Number	Frequency (cm^{-1})	Vibrational Mode	Δ in Frequency from Experimental
Experimental	1136	-	-
	1152	-	-
1	1141	E1 Twisting	+5
2	1141	E1 COH Bend x2	+5
	1158	G3 COH Bend	+6
3	1141	E1 Twisting	+5
4	1158	G3 COH Bend	+6
5	1158	G3 COH Bend	+6

Comparison of the relative energies of the different isomers was given in Table 2.3.5. Isomer 4 has the lowest energy of all the structures and was used to calculate the relative energies. Multiple intramolecular hydrogen bonds within isomer 4 helped stabilize the calculated structure, namely in the E1 residue. Protonation on the E1 γ carbonyl has allowed it to form a hydrogen bond with the E1 α carboxyl group, bringing the α and γ ends closer in space. There is also a hydrogen bond across the chain between the β proton of E1 and the G3 carbonyl. A β proton is very weakly acidic compared to an α proton, but this interaction seems to be occurring in the peptide chain and could provide a small stability factor for the peptide. Isomer 5 is the second most stable in terms of relative energy, but there is a major difference in structure between 4 and 5 even though the sites of protonation are the same. The E1 residue still forms a hydrogen bond within itself, but the other end of the chain does not form the hydrogen bond as seen in isomer 4. The G3 end of the peptide is in an extended conformation and is facing away from the E1 residue. This major change in the calculated structure may be responsible for the higher relative energy of this conformer. A large gap in the relative energies occurs between isomers 5 and 3. Isomer 3 is the third most stable of all the conformers. This isomer does not contain any clear intramolecular hydrogen bonds that distort the structure in the same manner that isomers 4 or 5 did. This peptide appears to be in a semi-folded state, where the G3 end of the chain is almost linear, but then the C2 α carbon the chain takes a turn before extending off again. In each of the ends, the atoms appear to be in a staggered conformation, thus lowering the overall energy, but not enough to be one of the lowest. Isomer 1has the protonation occurring on atom 6, but again this proton does not seem to form any intramolcular hydrogen bonds. It also appears to have a semi-folded conformation like isomer 3, but the E1 side chain is more linear for isomer 3. The group of atoms from the C2 amide to the E1 β

carbon has a dihedral angle that is closer to 180°. Since there are not many hydrogen bonds stabilizing this structure and is not as folded, the energy of this peptide is much higher than that of isomer 3. The relative energy of isomer 2 was +126 kJ mol^{-1}, which was the highest of all five candidates. This isomer also has no intramolecular hydrogen bonds that would add extra stability for this protonated structure. The structure is also in a semi-folded conformation and is even more linear in the E1 side chain than isomer 1. The isomer also has been protonated on atom 2, which does remove some of the stability of a regular carboxyl group. When compared to a neutral carboxyl group, there is resonance between the carbonyl and hydroxyl group, but due to the added proton at on the carbonyl for this isomer, the resonance structures are the same for both oxygen atoms. This increases the energy of the overall system and makes this isomer very unstable when compared to isomer 4.

Based on the similarities in spectra, stability and energy of the structure, certain isomers can be removed as the possible protonated species. Isomer 5 is the best candidate of the five structures. The differences between the experimental spectrum and calculated spectrum are minimal for all three regions of the spectrum. The structure of this protonated species has an intramolecular hydrogen bond stabilizing the E1 end of the peptide, which makes the calculated energy of this isomer the second lowest of all five. Isomer 4 is also a very good candidate to represent the experimental spectrum. The amide and low frequency regions of the spectra match well with one another, but the calculated carbonyl frequency is blue-shifted compared to the experimental. Even though there are some slight discrepancies in the shifts for the spectra, the energy of the protonated peptide is the lowest of all five, making it the most stable. The structure of this protonated species is very stable due to a variety of intramolecular hydrogen bonds that are being formed within the entire peptide, giving it a folded structure. Since both of

these isomers have the same site of protonation, it is possible that both of these isomers exist as the experimental protonated species. Evidence of this is comes from the experimental spectrum; there are two peaks at 1726 and 1742 cm^{-1}. Isomer 5 had a calculated frequency very close to 1726 cm^{-1} and isomer 4 had a calculated frequency very close to 1742 cm^{-1}. These two calculated peaks combined could be the cause for the experimental spectrum having two peaks such as this. This would suggest that there could be a statistical mixture of the two isomers for this protonated sample of GSH. Isomer 5 is probably the strongest candidate of these five isomers that would represent the protonated GSH molecule and isomer 4 is also a good structure to represent the molecule, but there is a strong possibility that both of these structures exist in a mixture of species.

2.3.4 Computational Modelling of Glutathione

The result of the DFT calculation of GSH in DMF resulted in the structure shown in Figure 2.3.7. The structure of the GSH was calculated without the influence of the polymer resin, but still gives a good representation of the peptide's monomeric conformation in DMF.

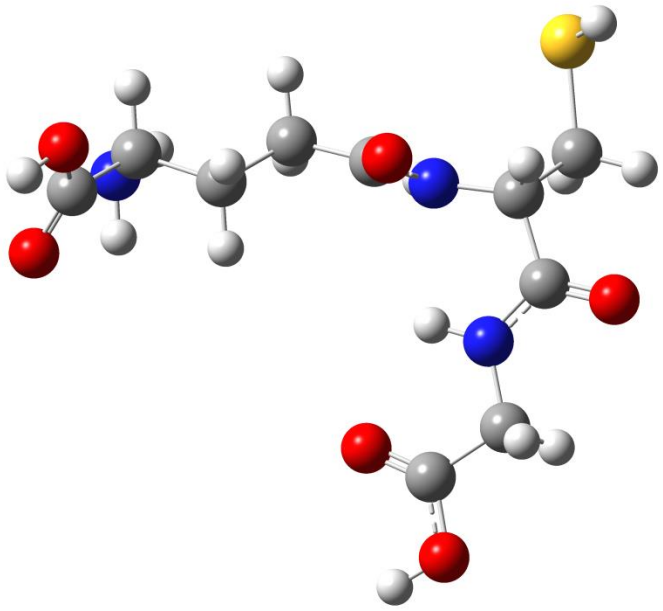


Figure 2.3.7: Optimized structure of GSH using B3LYP/6-31g(d) level of theory. The PCM model of solvation was used to simulate a DMF environment around the GSH molecule. The carbon, oxygen, nitrogen, sulphur and hydrogen atoms are coloured in grey, red, blue, yellow and white, respectively.

Starting with the structure in Figure 2.3.7, a variety of GSH conformers can be explored. An AMBER force field for the GSH molecule needed to be generated before the different structures of the initial product could be created. Using Gaussian 09W with the ChelpG algorithm, the partial charges for the individual atoms in the molecule were determined (see Appendix). Combining the AMBER library of known values between atoms and the calculated partial charges, the AMBER force field for the specific GSH molecule was created to observe the different conformations. Simple molecular mechanics calculations were performed using the AMBER force field and in total, 20 000 structures were generated from the MM calculations. The output files of each individual calculation contained a geometrically optimized structure

and an energy in Hartrees. The collection of 20 000 structures were then tabulated based on their output number and the energy that accompanies that structure (Figure 2.3.8).

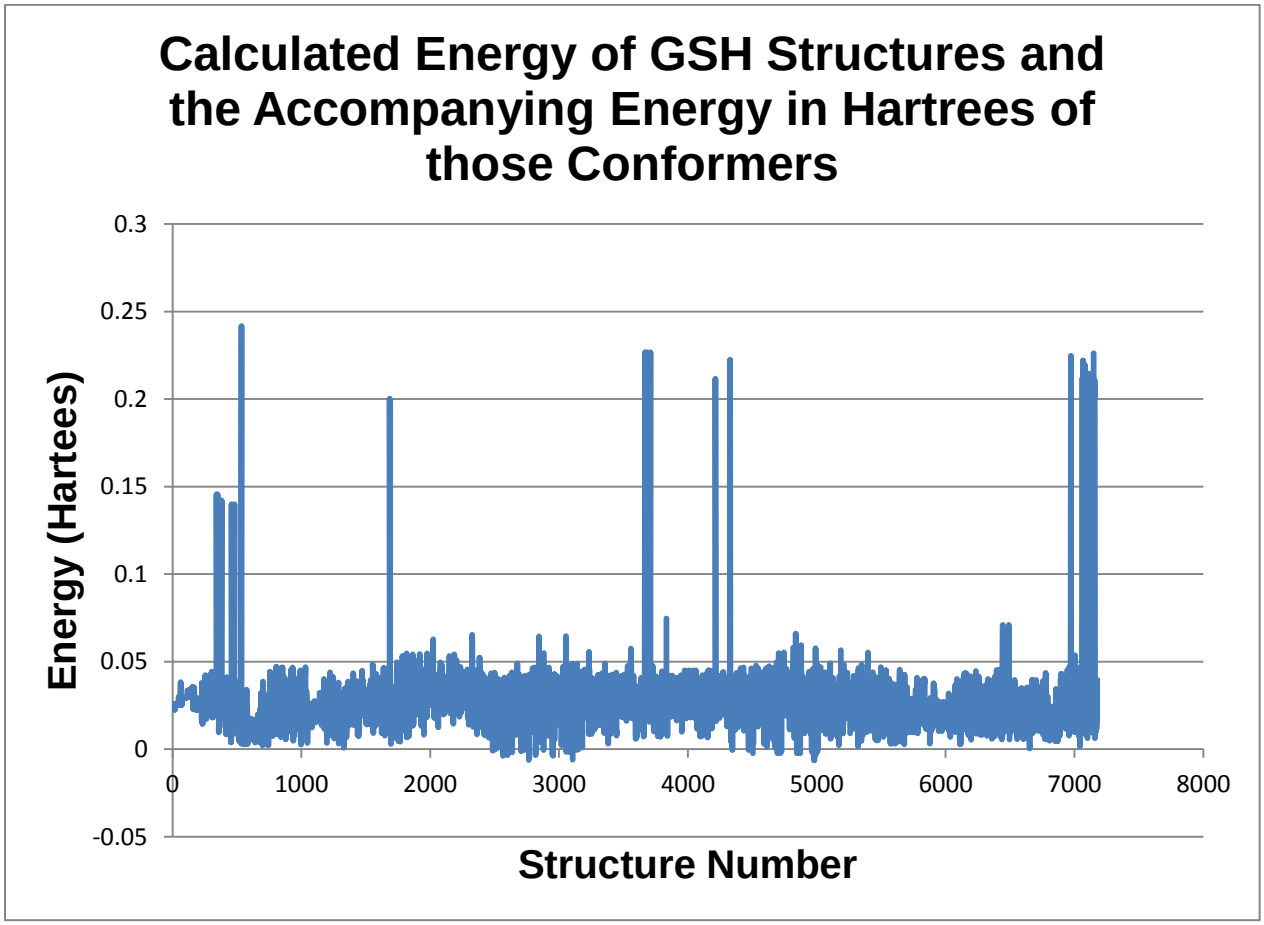


Figure 2.3.8: Plot depicting the energy of the individually calculated structures with respect to the output number of the GSH molecule.

The next step was organizing the calculated structures from lowest in energy to highest (Figure 2.3.9). The lowest energy structure in this case is the most stable of all the structures, but it is not a good assumption that the lowest energy structure after the MM calculations will be the same as the DFT structure. The reasoning behind this is that MM calculations are less precise than the DFT calculations and will sometimes give deviations between structures. The change in

energy with the change in structure can be observed graphically by looking at the change in the slope of the graph. The points on the graph where the slope changes dramatically is indicative of a large change in geometry and a large change in energy. The front end of the graph has a large slope for the low energy conformations representing some changes in the more stable structures. As the plot progresses, not much change in the energy or the structures are seen through the middle of the set of structures, but towards the end large changes are seen by the sharp spike for the high energy conformations. The points on the graph where the slope changes the most (red squares in Figure 2.3.9) are the structures that are the most interesting to observe since they will be unique to one another due to the large changes in geometry. The points in the middle of the graph where the slope does not change as much are still important. This plot was not specifically used to determine which structures should be observed, but was used to generate the next plot which ultimately decided which specific structures were chosen to analyze.

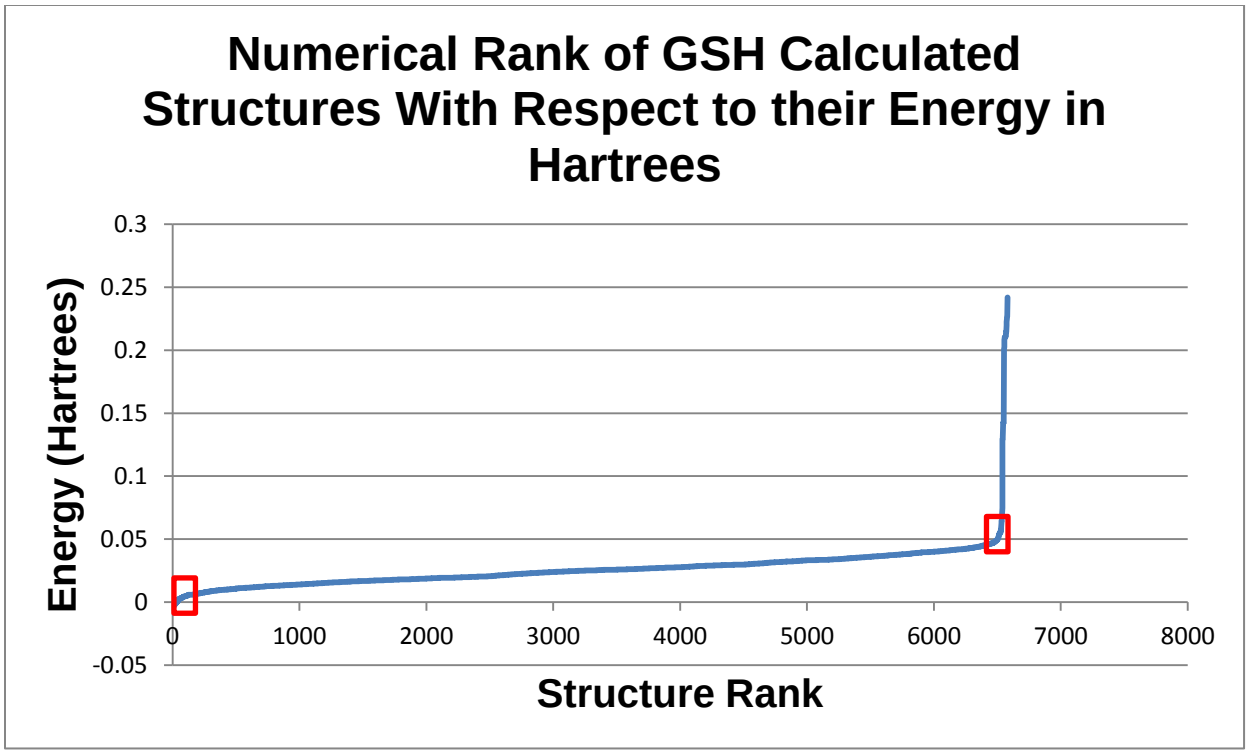


Figure 2.3.9: The GSH molecules after they were optimized individually were reorganized based on their energies (in Hartrees) from lowest to highest. The red squares indicate the areas of the plot with the biggest change that would possess the most unique structures to analyze.

Choosing the structures that are unique and to be further optimized was done by creating a "threshold" graph (Figure 2.3.10). The threshold graph is a derivative plot of Figure 2.3.9 and by setting a limit that requires a change in energy of 1.00^{-5} Hartrees between successive structures, the most unique structures can be chosen. A change in energy less than 1.00^{-5} between successive structures was marked a zero when tabulated and suggests there was not a major difference between conformers. Structures were chosen by finding the largest changes in the points of the derivative plot. Each line in Figure 2.3.10 represents a single structure and large energy changes between structures are noted by a drop from one line to another. If there is very little change in energy (i.e. the threshold of 1.00^{-5} was not met) there will be a gap in the

plot and the structure before the gap is the most unique structure for that region. The plot shows a great degree of change at the extremes indicating that the structures at these points are the most interesting to analyze.

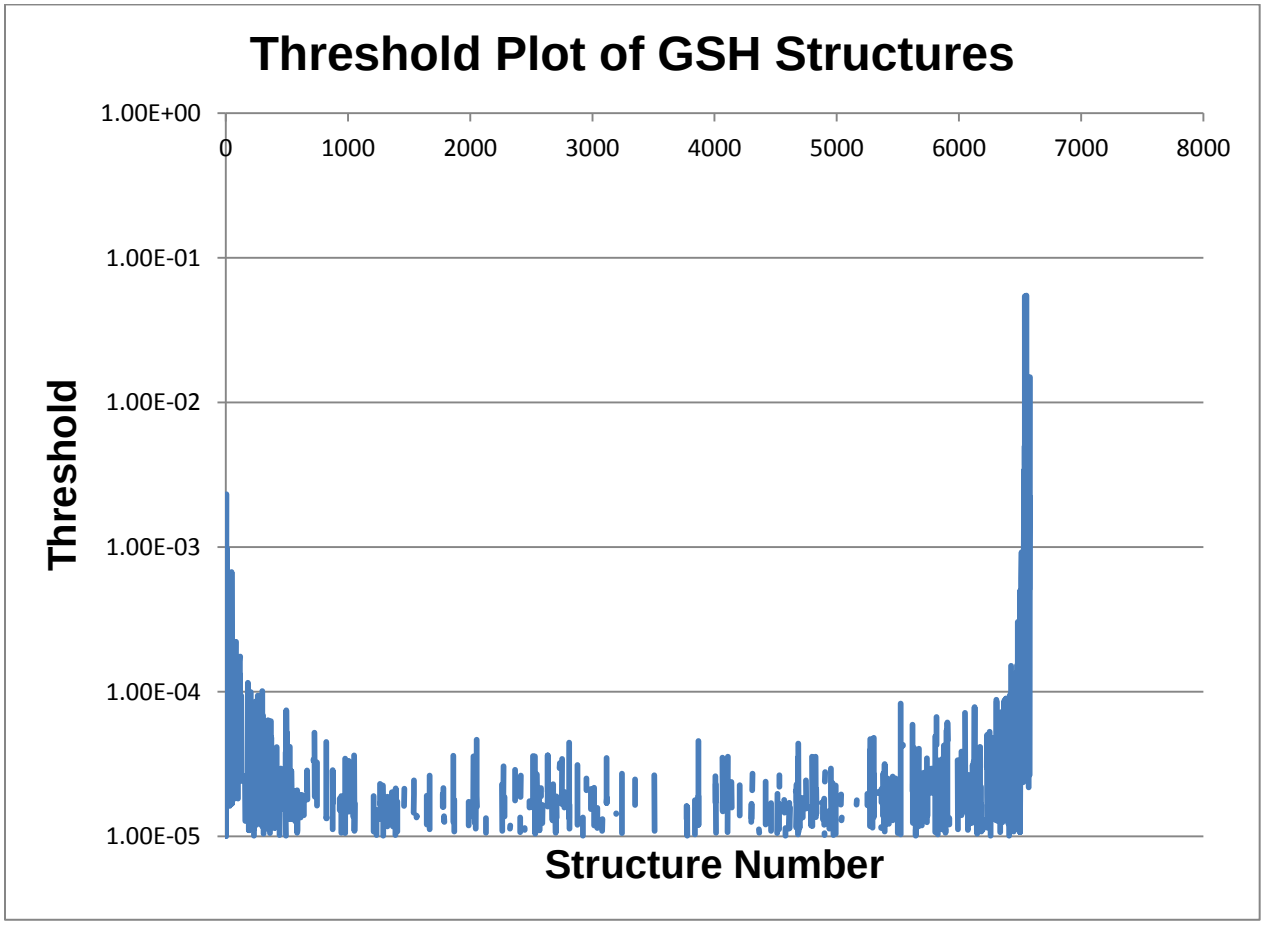


Figure 2.3.10: Threshold plot depicting the difference in energies between structures. If the threshold of 1.00 x 10^{-5} in energy difference between structures was not met, a value of zero was input indicating little to no change in structure conformation.

Using the threshold plot, 19 structures were chosen that were considered unique and important in the understanding of the different conformations of GSH. Some of the structures that were chosen were the lowest and highest energy conformer, along with some more

structures at the extremes of the plot. A few conformers from the middle of the plot were chosen as well to cover most of that region. The lowest energy structure (Figure 2.3.11) has a folded conformation, where the E1 end is close in space to the G3 end and the C2 system is the middle point of the folded tripeptide. The folding of the tripeptide in this conformation is favoured due to the hydrogen bonding between the α carbonyl of E1 and the proton on the carboxyl group of G3. Also an interaction between the amide proton of G3 and the carboxyl backbone of E1 causes the structure to be slightly folded in as well. This peptide is also very flexible at the C- and N-terminus which does not make the tripeptide planar. Even though this calculated structure was done at a relatively low level of theory, it does show some similarities to the results obtained from the ROESY spectrum. The sequential amino acid interactions are seen in this peptide with an intramolecular distance of less than 6 Å for each interaction. One of the most important interactions that gave the most information about structure between the α proton of E1 and β proton of C2 is very close in space for this conformer. Again, since the structure was calculated using molecular mechanics, the structures obtained should be viewed as rough estimations and not considered as entirely accurate when compared to the higher levels structures obtained by DFT calculations. There are large differences in the structure between the lowest energy MM structure and the DFT structure, since the MM structure is very folded and the DFT structure seems to be only semi-folded. Also, comparisons to the ROESY NMR should be viewed with caution as the structure provided from the MM might not be the one seen in the NMR experiments.

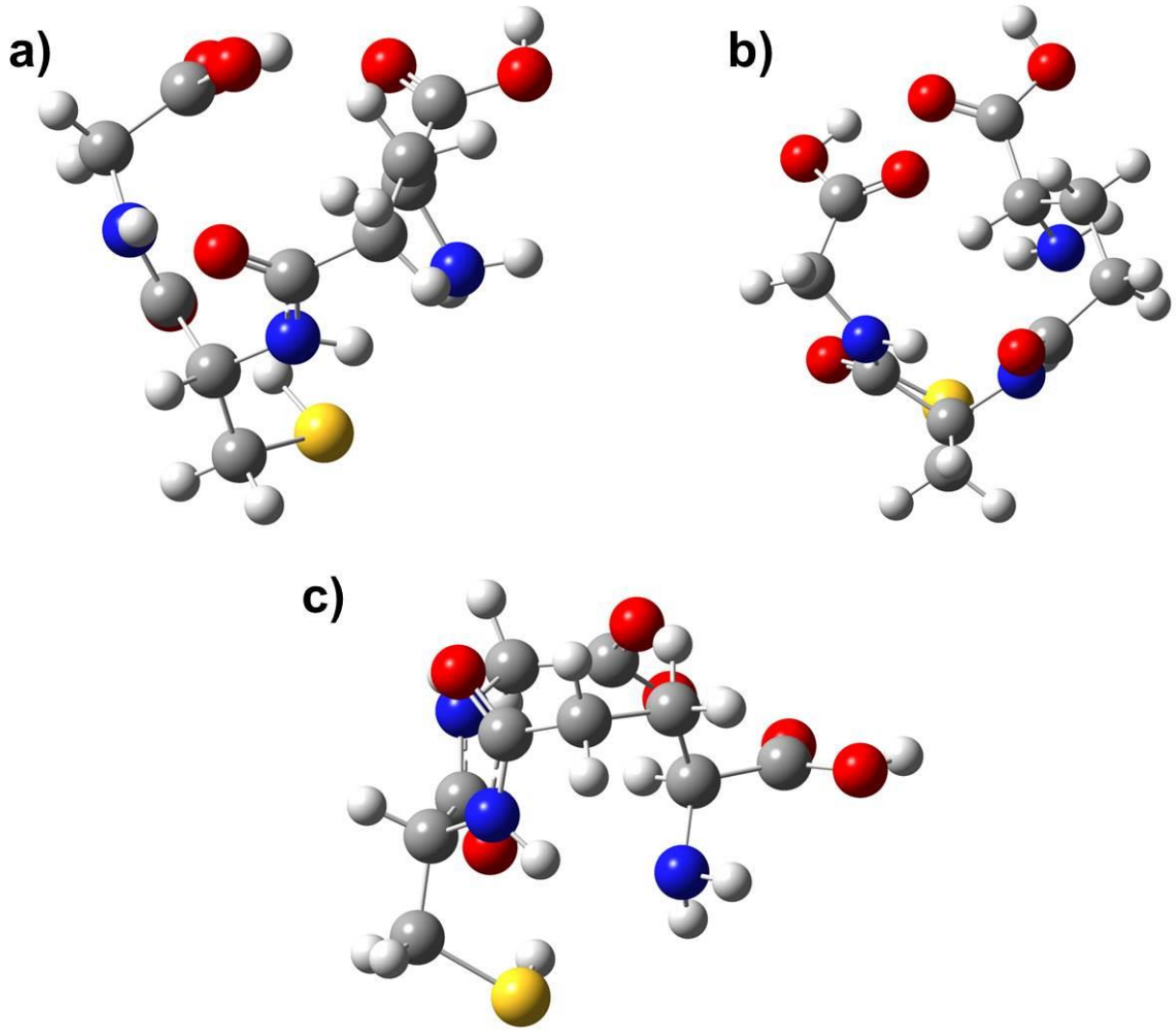


Figure 2.3.11: Lowest energy conformation of a free GSH calculated using molecular mechanics and a custom AMBER force field viewed from various angles. Carbon, hydrogen, oxygen, nitrogen and sulphur atoms are shown in grey, white, red, blue and yellow, respectively. Parts a) and b) show the folded shape of the tripeptide and an overhead view of some of the hydrogen bond interactions bringing the chain close together; c) depicts the flexible ends of the GSH molecule.

On the other hand, the structure with the highest energy (Figure 2.3.12) after the MM calculations does not have the same conformation at all. The G3 end in this conformation points

away from the E1 end, but the E1 amino group is oriented so that it is pointing towards the G3 end. The E1 end is forced into this position due a hydrogen bonding interaction between the α amino group of E1 and the γ carboxyl group of E1. The G3 end of the peptide does not have any interactions with any other part of the peptide chain, which leads to that portion to be extending away from the rest of the peptide chain. This peptide has a semi-folded conformation, much like one of the forms it has been known to take from Zhang's NMR and computational studies.[43] Using the low level of theory calculations, a group of structures of glutathione have been generated and can be used for more complex calculations to establish an idea of the different conformations this peptide can adopt.

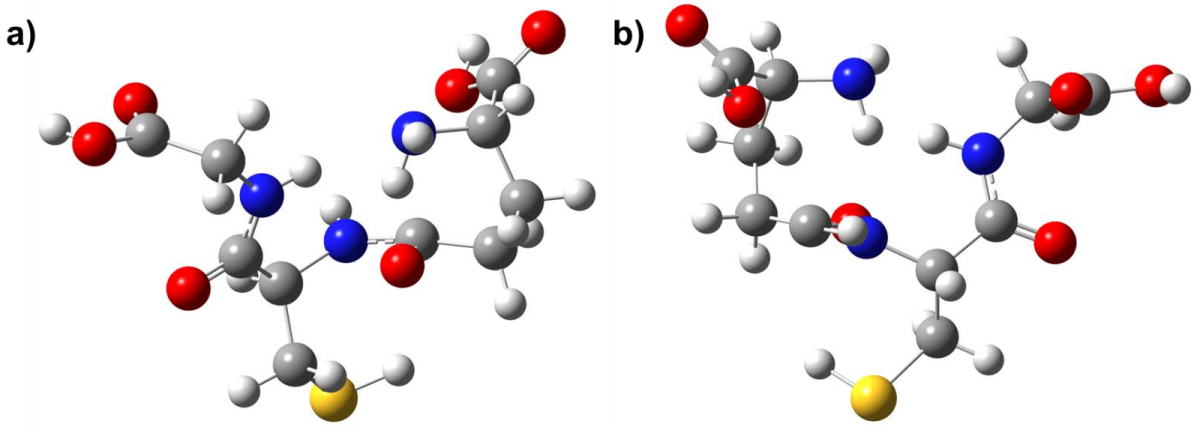


Figure 2.3.12: Highest energy conformation of the tripepetide GSH shown in the semi-folded conformation. Carbon, hydrogen, oxygen, nitrogen and sulphur atoms are shown in grey, white, red, blue and yellow, respectively. The E1 end of the peptide is folded in towards the middle of the peptide, whereas the G3 end is extending away from the middle.

The next step was to attach the lowest energy conformers to a linker and then subject this linker-bound GSH to different molecular dynamics simulations and higher level

calculations (DFT level of theory). The linker-bound peptide calculations should provide a better interpretation of how the peptide folds while being bound to the resin.

The molecular dynamics simulations provided the probabilities for each of the different dihedral angles that could be present. Figure 2.3.13 shows the results of the simulation and values that each dihedral angle is most likely to adopt. Each curve represents a group of sequential atoms and the most populated dihedral angle it adopted in the simulation. For example, the dihedral angle between E1 carbonyl oxygen/E1 carbonyl carbon/E1 Cα/E1 Cβ (denoted as O-C-CT-CT in the histogram) is most likely to have a dihedral angle of -79°. Table 2.3.9 summarizes the angles in which all of the angles the peptide favours the most.

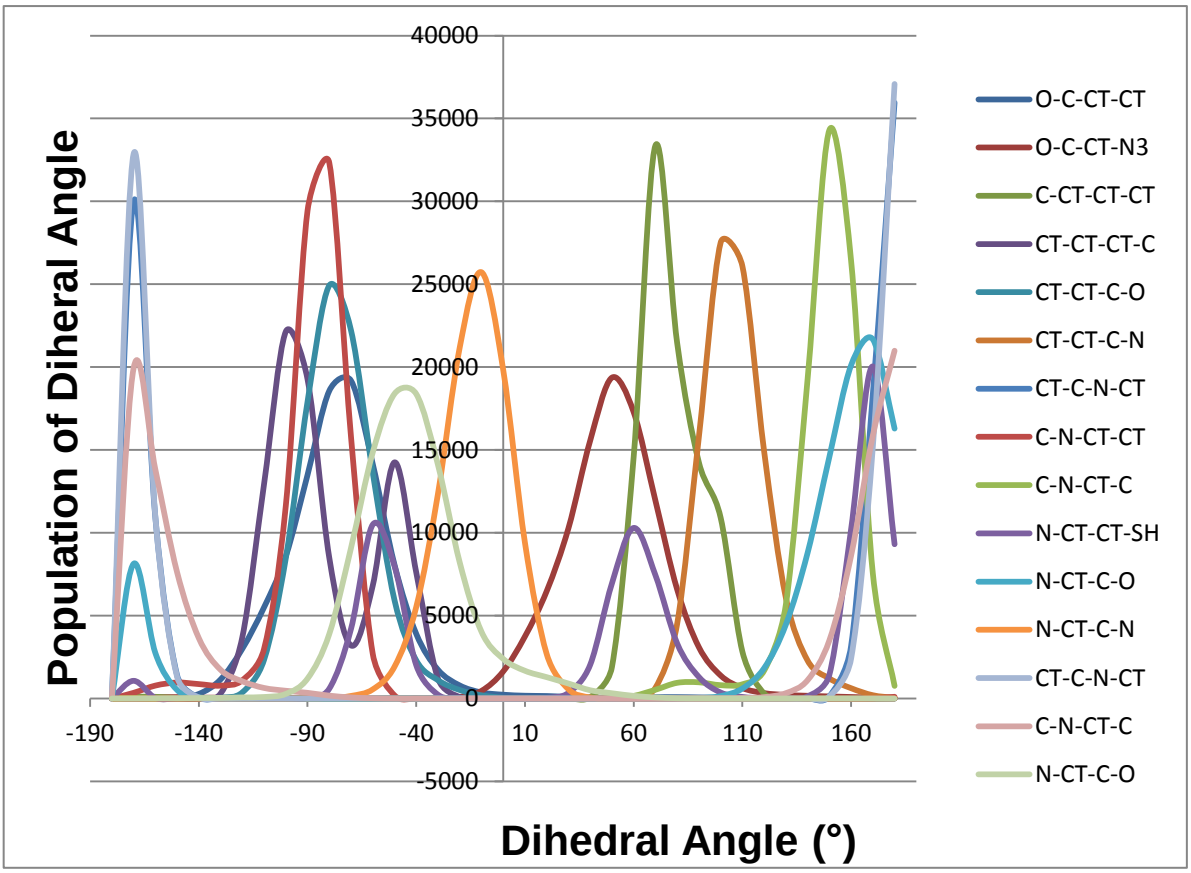


Figure 2.3.13: Histogram of the most favoured angles in the glutathione peptide while bound to the linker. The data for this chart was collected using MMTK to simulate and predict the peptide's most favoured conformations at 298.15K. Each coloured line represents a different dihedral angle in the GSH peptide. The data from this figure is summarized in Table 2.3.9.

Table 2.3.9: From Figure 2.3.12, the most populated angles for each of the dihedral angles is presented in the second column. C=O is representative of the carbonyl carbon and O represents the oxygen atom in the carbonyl group participating in the dihedral angle.

Angle	Dihedral Group	Angle (°)	2nd Most Populated Angle (°)
1	E1 O/E1 C=O/E1 Cα/E1 Cβ	-79	-89
2	E1 O/E1 C=O/E1 Cα/E1 NH_2	50	60
3	E1 C=O/E1 Cα/E1 Cβ/E1 Cγ	70	-
4	E1 Cα/E1 Cβ/E1 Cγ/E1 γ C=O	-109	-99
5	E1 Cβ/E1 Cγ/E1 γ C=O/E1 γ O	-89	-79
6	E1 Cβ/E1 Cγ/E1 γ C=O/C2 NH	100	110
7	E1 Cγ/E1 γ C=O/C2 NH/C2 Cα	180	-
8	E1 γ C=O/C2 NH/C2 Cα/C2 Cβ	-89	-99
9	E1 γ C=O/C2 NH/C2 Cα/C2 C=O	150	-
10	C2 NH/C2 Cα/C2 Cβ/C2 SH	170	-
11	C2 NH/C2 Cα/C2 C=O/C2 O	170	160
12	C2 NH/C2 Cα/C2 C=O/G3 NH	-19	-29
13	C2 Cα/C2 C=O/G3 NH/G3 Cα	180	-
14	C2 C=O/G3 NH/G3 Cα/G3 C=O	180	-
15	G3 NH/G3 Cα/G3 C=O/G3 O	-69	-59

These simulations are very much like the molecular mechanics calculations and are only approximations. While the most populated angles are being reported, some of the other angles are populated as well, but not to the same extent as the preferred angle. In some cases, the difference in population was only less than 3000 structures, of the possible 100000. This secondary angle is shown in the fourth column of Table 2.3.8 and was only used if the difference between the primary and secondary angle population was less than 3000. With these results, the methods of molecular dynamics can be compared to the DFT structure.

The free GSH peptide from Figure 2.3.7 was used to do a further calculation using the DFT level of theory. The peptide was coupled to a linker and then subjected to the higher level calculation to determine the structure of the peptide and evidence of the effects (if any) introduced by the linker. The resulting structure is shown in Figure 2.3.14. Assuming the structure exists in a semi-folded state, this structure fits fairly well with the NMR predictions based on the proximity of the E1 α proton to the C2 β proton. The structure is slightly folded, but the linkage between the G3 residue and the linker influence the secondary structure formation of this peptide. The linker almost keeps the G3 and C2 residues linear, but then the semi-folded structure begins at the C2 α group. The dihedral angles of the backbone between G3 α-G3 NH-C2 C=O-C2 α is approximately -179°, but the G3 NH-C2 C=O-C2 α-C2 NH dihedral is about 1.5°. The structure quickly goes from being almost a sheet structure to having some sort of bend or fold. The rest of the backbone after this fold has almost a sheet structure as well. This slight bend in the peptide chain brings the E1 end closer to the C2 portion of the peptide allowing for the key NOE interaction to be observed. There are only slight differences between the DFT structure of the free GSH and the linker bound structure, but those are minor differences in the angles. It is clear that, there are many similarities in the structures between the two peptides. This suggests that the linker does not affect the structure of the peptide very much for small peptides such as GSH. This comparison also supports the notion that the linker-resin still allows the peptide to adopt the native conformation that the peptide prefers.

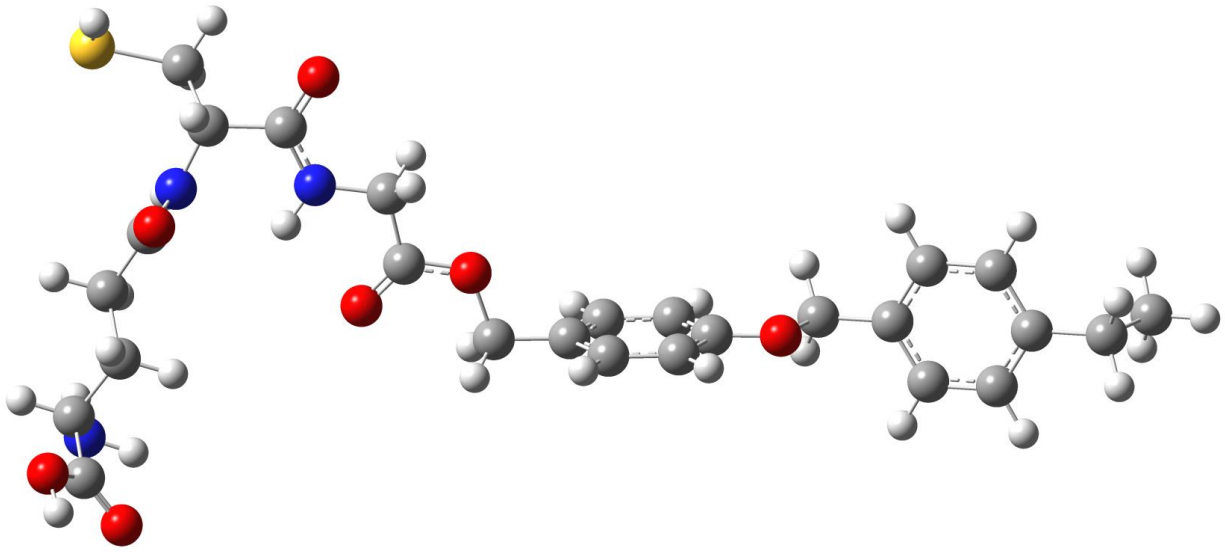


Figure 2.3.14: DFT calculated structure of GSH while bound to the resin via a linker group. The peptide that was calculated matches the NMR spectra fairly well. The semi-folded structure brings the E1 α proton close enough to the C2 β proton to see that key interaction, consistent with what was observed in the ROESY spectrum. The differences between this structure and Figure 2.3.7 are minimal suggesting that the linker does not affect the native secondary structure of the peptide on a large scale. Small differences in the torsion angles and bond angles were observed, but nothing drastic. The carbon, oxygen, nitrogen, hydrogen, and sulphur atoms are coloured in grey, red, blue, white and yellow, respectively.

By combining all of the results from the different methods there are many similarities in the results. The NMR results suggest that the structure has some sort of folded structure, semi-folded or folded, from interactions 17 and 20 between the E1 Hα and C2 Hβ protons. These interactions suggest that the peptide must be folded or semi-folded to bring the E1 end closer to the middle part of the peptide. From the IRMPD, the two possible best candidates that best fit the experimental spectrum were isomers 4 and 5 in Figure 2.3.6. Isomer 5 had a structure that was folded at one end of the peptide stabilized by an intramolecular hydrogen bond. The isomer

matched well in all regions of the spectra and had the second lowest energy of all five isomers. Isomer 4 had a spectrum that matched fairly well to the experimental and the energy was the lowest for all five structures and also had a folded structure stabilized by a variety of hydrogen bonds. These structures are protonated, which is different from the resin-bound peptide, so the behaviour of the protonated peptide may be different than the neutral structure. Therefore, analyzing the IRMPD structures on their own, separate from the NMR and structure calculations was done and not used to draw upon the final structure of the resin-bound GSH. The molecular mechanics and dynamics simulations suggest that the peptide is in more of a folded state, but the DFT calculation on the GSH-linker structure suggest a semi-folded state. Both the molecular mechanics and dynamics experiments are very rough approximations and cannot be considered entirely accurate on their own. Information from the NMR and DFT structures has more value as experimental observations. The combination of NMR and DFT calculations suggest a semi-folded conformation when GSH is bound to the linker. However, to simulate the full effect of the linker and solvent, the molecule should be treated using the explicit model of solvation where GSH interacts with molecules of DMF as opposed to being in a box simulating the dielectric constant of DMF.

Using a variety of spectroscopic and computational methods, it is likely that the peptide has adopted a semi-folded state. The molecular mechanics and dynamics experiments are only approximations and cannot be interpreted as fully accurate and were not used to draw conclusions about the GSH structure. The lack of an E1-G3 interaction in the ROESY spectrum eliminates the possibility of a folded conformation for the resin-bound peptide. The DFT calculations also suggest that the peptide has adopted a semi-folded conformation while bound to the resin further eliminating the possibility of a folded structure. The IRMPD results were

analyzed separately and did not factor into the analysis of the resin-bound peptide conformation since the protonated species would behave differently than the synthesized peptide. The protonated species likely has a folded conformation, where protonation of the E1 γ carbonyl allows for the peptide to form many intramolecular hydrogen bonds to create a stable structure. The computational calculations at all levels could be improved by adding more complex situations. These situations could be using different basis sets, creating a larger set of structures for the peptide and linker and using the explicit model of solvation. NMR results can be improved by attaining a sample that has higher molecular substitution and a ROESY with more sensitivity. These results on the study of glutathione, suggest that the peptide has a semi-folded conformation due to the agreement between the NMR and DFT results.

2.4 Discussion

Synthesis of the tripeptide, glutathione, was done successfully achieving a final substitution of 0.3987 mmol/g (49.8% substitution of the total available sites). The synthesis was performed using standard solid-phase peptide synthesis methods. A Wang resin was used to anchor the peptide, and the protecting groups on the residues required mild conditions to be cleaved. The C2 residue had a Mmt protecting group that was removed with 2% hydrazine in DMF and the E1 residue had an ODmab group that was removed with 1% TFA in DCM. Coupling of the individual amino acids was successful for G3 and E1, but with C2 there was some difficulty coupling the residue. The methods that were used are "old-fashioned" and have been replaced by modern methods that do not require as many materials, and techniques that are more efficient. The initial coupling, even though an 85% coupling was achieved, could have been improved with the use of DMAP or MSNT. These two starting materials are more effective than DCBC as they allow for multiple attempts without capping the reaction sites. They do, however, require more materials and take longer, but the outcome of multiple attempts could yield better results than DCBC.[12] The HBTU/HOBt method is also out-dated as HOBt has been classified as an explosive and cannot be transported.[27] Newer methods simply use HCTU as a coupling reagent and only need a two hour reaction time per attempt. By replacing these two flaws in the original synthesis, the substitution could be increased for this tripeptide so that more than 50% of the total sites could be occupied.

Both 1D and 2D HRMAS NMR experiments were used to analyze the resin-bound sample of glutathione. Using COSY and TOCSY, the three different amino acid spin systems were identified along with their respective R-groups. The E1 residue showed a signal at 10.28 ppm which was not seen in the COSY and TOCSY spectrum, but was seen in the 1D spectrum.

Each of the residues also gave rise to a set of secondary shifts, but the degree of detail was not as high as the first set of peaks. This secondary set of shifts could be the result of a secondary conformer formed by the peptide. With the spin systems defined, ROESY was performed to analyze the through-space interactions of the peptide, and these were used in an attempt to hypothesize a structure. In total, 20 different interactions were observed where six were through-space. The key interactions from this experiment were between the E1 α proton and each of the two C2 β protons. Having these protons in close proximity suggests that the peptide does adopt somewhat of a semi-folded conformation. The peptide is not completely folded because there were no E1 and G3 interactions observed, and the peptide is not fully extended based on interactions noted above, which would not be seen by ROESY in the extended state. The secondary set of shifts also saw few interactions, but not enough to hypothesize a structure since those interactions were just within the same spin system or on sequential backbone residues. The semi-folded conformation is a form that Zhang *et al.* had proposed in their computational and NMR study[43], so it is not surprising that this peptide adopted one conformation from this study.

IRMPD studies on this peptide were also done with the aid of Dr. Terry McMahon and Dr. Rick Marta. An IRMPD spectrum was obtained from a sample of GSH and was compared to five different calculated spectra of protonated species. The regions that were used to characterize and compare the spectra were the carbonyl region, amide region and low frequency region. The protonated species that had the calculated spectrum that matched the best to the experimental was protonated isomer 5. Even though it was not an exact match, the small differences in the frequencies gave isomer 5 the closest match in terms of peak positioning. This structure had the second lowest energy of all five structures and had an intramolcular hydrogen

bond forming between the protonated carbonyl and the E1 residue. Isomer 4 had the lowest electronic and thermal free energies of the five possible candidates. There was only one peak that was very blue-shifted for the calculated spectrum, but the peaks in the other regions matched very well. The structure of this isomer was in a folded state, where a variety of hydrogen bonds brought the two ends of the chain together in close proximity and stabilized this structure. Taking the comparison of the calculated and experimental spectra, free energies, and conformation of the calculated structures into account, isomers 4 and 5 are both possible candidates that fit the experimental spectrum the best and it is possible that a mixture of these species exists.

Computational studies on GSH were also conducted at the AMBER and DFT levels of theory to predict the most favourable conformation of the peptide. Initially free GSH was drawn into Gaussian 09W and then optimized using DFT level of theory with the 6-31g+(d,p) basis set in DMF using the PCM model of solvation. Using the DFT structure as a starting point, the structure was modified to create thousands of new structures and subjected to molecular mechanic calculations using a custom AMBER force field. The results were organized and the most favourable conformation of GSH was extracted and used in further studies. GSH was then attached to a linker and molecular dynamic experiments were done to predict the most favourable dihedral angles for the peptide. Another GSH-linker peptide was put into Gaussian and calculated at the DFT level of theory to produce a realistic interpretation of the GSH peptide while bound to the linker in DMF. This structure best represents the peptide that was synthesized and the results of the calculations share similarities to the NMR results. The molecular dynamics results did not agree though, as this technique is not as accurate and only provides rough approximations. Therefore, the structure from the computer based calculations is

a semi-folded conformation and shares similarities to the NMR results for this resin-bound peptide.

Using SPPS, a sample of GSH was bound to a resin and analyzed by HRMAS NMR spectroscopy. An IRMPD spectrum of GSH was also obtained and different protonated structures and their corresponding spectra, were generated to draw comparisons. Computational calculations were done with free GSH and the linker to generate a structure that best resembles the peptide while bound to the resin. All of these methods were then compared to one another and similarities in the findings were reported. The structure seems to favour a semi-folded conformation. These initial studies are similar to Zhang's recent work[43] where the peptide exists in a semi-folded state, but further work needs to be done. The synthesis of GSH can be improved to achieve a higher substitution on the resin in a number of ways. The HMBA-AM resin used for the FK-13 synthesis (see next chapter) can be used along with a different first amino acid coupling method (not DCBC). Also, Fmoc-Glu-O-2-PhiPr and Fmoc-Cys(Trt)-OH would be better amino acids to be used in conjunction with the HMBA-AM resin. These two protecting groups are acid labile and would not affect the removal of the peptide from the resin.[13, 50] The computational calculations on GSH can be further improved by using an explicit model of solvation as opposed to the implicit model to understand the way the peptide interacts with the DMF molecules. Also the structures that were chosen after the random sampling could be all subjected to more complex calculations, which were not done due to time constraints in this study. There are a lot of areas for improvement and future work for these peptides, but the work done in this thesis can easily be built upon for further exploration and understanding.

3.0 FK-13

3.1 Introduction

3.2 Experimental Procedures

3.3 Results

3.4 Discussion

3.1 Introduction

FK-13 is a part of the larger LL-37 peptide, which plays a vital role in the first line of defense against bacterial infections. It has also been shown to have anti-cancerous and anti-HIV properties, but these characteristics specifically come from the FK-13 region of the peptide. Some of the first experiments to determine the structure of this peptide were done using circular dichroism methods, but now NMR has been manipulated to determine its structure, function and mechanism. Understanding the structure and role of this peptide is key has many biological and biochemical implications to the prevention of infections and other harmful agents.

The peptide LL-37, a portion of a larger protein hCAP-18, has been identified as part of the cathelicidin family of proteins and also possess antibacterial properties.[51] The gene encoding this peptide is highly expressed in bone marrow and the testes, organs that are usually free from infectious bacteria. The structure of the LL-37 peptide is an α-helix, which plays a large role in the antibacterial function and enhanced capabilities of this peptide.[52] Solid phase synthesis and subsequent analysis by circular dichroism provides evidence that the LL-37 peptide favours an α-helical conformation, where a plane across the centre of the helix separates the hydrophobic and hydrophilic residues.[53] Furthermore, by placing the LL-37 peptide in an aqueous environment, the α-helix formation was induced more favourably and was seen to have enhanced activity against *E. coli*. An aqueous salt environment was also used to test the formation of the peptide, and suggested that this medium also favoured helix formation.[53] A variety of factors in the environment dictates how much α-helix formation occurs with this sequence. Helix formation could be dependent on salt, pH, and concentration of the peptide sequence. Increased concentrations of the peptide and salt ions favour the formation of the helix, and environments that have neutral pH will favour the formation of the helix over low

pH.[52] The helix conformation is not only the preferred conformer of this peptide, but is a necessity for antibacterial activity. The helical structures displayed greater activity against *E. coli* than conformers that were random structures.[52]

The process in which the LL-37 peptide functions is believed to bein a manner where oligomers of the peptide attack and disrupt the bacterial membrane. When the oligomers bind to the outer bacterial membrane they will begin to disturb its order and are able to enter the inner membrane. Once inside, the oligomers dissociate into monomers and cover the inner membrane of the bacteria. Finally, after a certain concentration of LL-37 has been reached inside the bacterial membrane, the peptide begins to breakdown the inner membrane, killing the bacterium in the process.[53] Oren and his group hypothesized that this was the mechanism of LL-37 to fight against bacterial infections. ^{15}N and ^{31}P NMR studies also have been conducted to determine the mechanism in which LL-37 penetrates and disrupts bacterial membranes. ^{15}N solid-phase NMR studies revealed that the action of LL-37 was similar to what Oren's group had hypothesized, but the ^{31}P results of this study disagreed with this model. The ^{31}P NMR and differential scanning calorimetry results suggested that the LL-37 peptide induces a positive curvature strain. Bacterial membranes normally have a negative curve in their structure, but the presence of LL-37 warps this negative conformation and strain the lipid bilayer to curve to a positive one. With larger concentrations of LL-37, the strain in the membrane allows for penetration and the leakage of ions and molecules.[54] LL-37 possess unique properties for a helical antimicrobial peptide, which make it stand apart from other helical peptides. The LL-37 peptide easily oligomerizes in solution, protecting it from enzymatic degradation compared to native peptides that are susceptible to this degradation. The LL-37 peptide can bind effectively to zwitterionic or negatively charged phospholipid vesicles, whereas native peptides can only bind to

negatively charged vesicles. When the peptide is bound to these vesicles, it is highly resistant to proteolytic degradation whether the vesicle is zwitterionic or negatively charged. This peptide is also able to self-associate when bound to zwitterionic phospholipid vesicles, but will dissociate when bound to negative vesicles.[53] These properties of LL-37 make it a powerful molecule to fight bacterial infections. LL-37 also has properties that fight against cystic fibrosis[56], sepsis in rats[57], and anti-HIV and anti-cancerous properties[58].

The portion of the LL-37 peptide that is responsible for the anti-HIV properties is FK-13. FK-13, called that for the first two residues and the number of residues that make up this portion of the entire LL-37, has the sequence Phe-Lys-Arg-Ile-Val-Gln-Arg-Ile-Lys-Asp-Phe-Leu-Arg (FKRIVQRIKDFLR).[51,53] This section of the peptide is the shortest the LL-37 can be cleaved to and still retain antimicrobial properties. Li had tested the shortest length that the LL-37 peptide could be cleaved to and still possess antimicrobial and anticancer activity. The process was to cleave amino acids from the C- and N-termini and evaluate peptide's activity; the results revealed that residues 17-29 were the minimal portion of the α-helical peptide to maintain its antimicrobial and anticancer properties.[59] Further studies revealed that the F17 residue was essential for the anti-HIV activity. Upon deletion of this residue, the peptide (KR-12) lost all anti-HIV activity, indicating that the function of the peptide was dependent on the FK-13 segment being intact and F17 being present in the sequence.[58]

The three-dimensional structures of native LL-37 and FK-13 have been predicted with the use of NMR. Solution NMR studies on the hCAP-18 protein have been conducted in the past to determine the structure of this protein. Using ^{1}H-^{1}H TOCSY experiments, the backbone of the peptide was able to be assigned and with NOESY experiments, the conformation of the peptide was hypothesized. Using the NOE interactions and calculating the best possible

structures (with the proper restraints in place), the protein was determined to have an α-helical structure.[60]

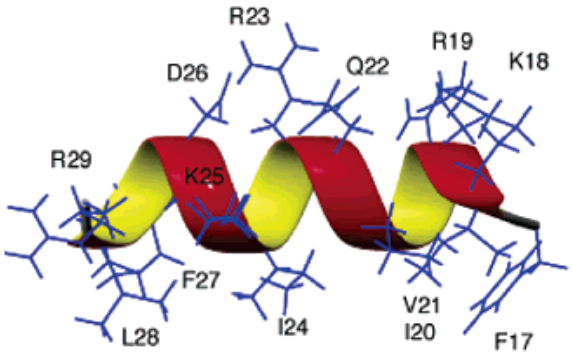


Figure 3.1.1: The native structure of FK-13, which is a part of the LL-37 peptide, is an α-helix. The α-helical structure plays a large role in the peptide's antibacterial function and mechanism of this peptide to fight against bacterial infections. Adapted from Li *et al*.[59]

Solid-state NMR was used by Henzler to investigate the structure and mechanism in which LL-37 operates. In one of their first studies, the structure was determined using ^{13}C CPMAS NMR of a powder sample of the LL-37 peptide. The carbonyl peaks were used to establish the structure of the peptide based on characteristic shifts for secondary structures. The frequencies at which the carbonyls in the peptide resonated were indicative of an α-helix. ^{15}N and ^{31}P NMR experiments were used in the same study to determine the mechanism of the peptide and the way it orients itself to disturb bacterial membranes.[55] Further studies done by Henzler utilized solid-state ^{2}H NMR to further understand where and how the LL-37 peptide inserts itself into lipid membranes. Their results examined the method by which the LL-37 peptide penetrates hydrophobic/hydrophilic bilayers. They hypothesized that the hydrophobic half of the helix inserts itself approximately 5-6 Å into the hydrophobic portion of the bilayer. They observed this phenomena by using de-Paked ^{2}H NMR experiments, and noticed a decrease

in quadrupolar splitting in the spectra, which indicated that there was some disorder in the lipid bilayer.[61] Two-dimensional heteronuclear experiments also played a critical role in the analysis of the LL-37 peptide. ^{15}N-^{1}H NMR experiments were used to justify that a sample of LL-37 was successfully cloned and purified for further examination. The conclusion was made based on characteristic Asn and Gln cross-peaks in the HSQC spectrum that were consistent with the peptide sequence. Although this study did not fully explore the structure of the peptide, it demonstrated the versatility of two-dimensional heteronuclear NMR in characterizing LL-37.[62] A study that uses both solid-state and solution-state NMR to investigate the structure of LL-37 was done by Porcelli. First, the resonances for the individual residues were established using two-dimensional ^{1}H-^{1}H TOCSY and NOESY experiments. By using all the NOE interactions, H/D exchange results, and computational simulations, the peptide was determined to have an α-helix-break-helix conformation. NOESY was used again to determine the structure of the peptide when it interacts with micelles. The results showed that the hydrophobic residues will point towards the centre of the micelle and the hydrophilic residues will be facing the solvent.[63] The structure of the peptide in micelles again was studied by Wang, but they used a different detergent in the micelles, and three-dimensional NMR. An isotopically enriched (^{15}N and ^{13}C) sample of LL-37 was analyzed using triple-resonance experiments that were ^{1}H-^{1}H-X (where X is either ^{15}N or ^{13}C). The results of this study revealed that there were some differences in the manner in which LL-37 interacts with micelles compared to Porcelli's study. They found that the peptide was not mobile at the N-terminus, but flexible at the C-terminus, the structure was a helix throughout (rather than helix-break-helix), and that the starting and ending residues are different between the different micelles.[51] Wang also attempted ^{13}C natural abundance experiments on this peptide and used ^{1}H-^{13}C and ^{1}H-^{15}N HSQC experiments to look at the

interactions between the peptide and micelles again. The heteronuclear experiments were used to establish dihedral angles of the peptide for better structure calculations, ^{13}C results revealed what portions of the peptide were more structured and the use of HSQC was used to determine specific residues interacting with the micelle.[64] Aside from the many studies that look at micelle and LL-37 interactions using NMR, some have proceeded to understand how the structure plays a role in cancer treatment.[65,66]

3.2 FK-13 Experimental

3.2.1 Synthesis of FK-13

Synthesis of the 13-residue cyclic peptide, FK-13, was done using methods outlined and adapted from Kates and Albericio[12] and Chan and White[13]. The amino acids used in this synthesis were: Fmoc-Asp-O-2-PhiPr-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ile-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Val-OH, Fmoc-Phe-OH and Fmoc-Leu-OH. All of the amino acids that were used in this experiment are considered to be part of the "standard 20 amino acids" (with the exception of Fmoc-Asp-O-2-PhiPr-OH) due to the strong acidic cleavage conditions all of the protecting groups share. In this synthesis, HOBt and HBTU were not used together as coupling reagents- instead HCTU replaced both these reagents. Synthesis was performed in a Merrifield vessel connected to a 250 mL filter flask for easy removal of excess reagents and solvents.

Another change implemented for this synthesis from that of glutathione was the use of an HMBA-AM resin (Figure 3.2.1). The HMBA-AM resin contains a free hydroxyl group for attachment of the first amino acid, much like the Wang OH resin, but this type of resin is base labile requiring strongly basic conditions to cleave the peptide from the resin.[21] Initially, 0.5153 g of HMBA-AM resin (copoly(styrene-1% DVB), 200-400 mesh, 0.70 – 1.20 mmol/g substitution) was weighed into the Merrifield vessel and was allowed to swell for 30 minutes in distilled DCM and then 30 minutes in DMF. The initial molar substitution of 0.88 mmol/g was used, as indicated on the bottle, to determine the weight of the first amino acid and volume of pyridine and DCBC needed for the first coupling.

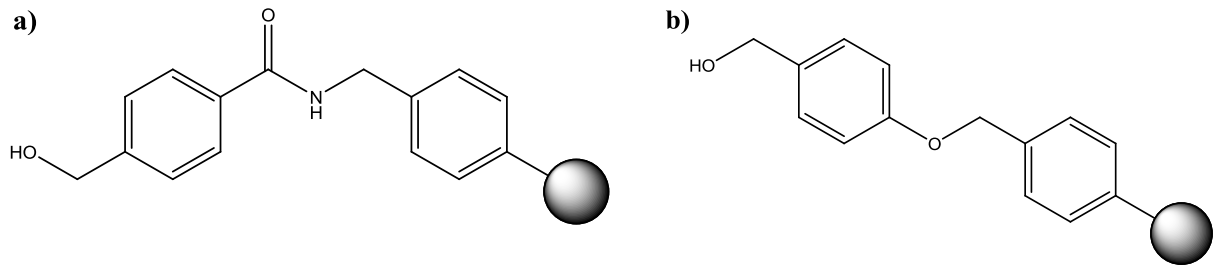


Figure 3.2.1: Two different resin-linker combinations used in solid-phase peptide synthesis: a) HMBA-AM resin, contains a free hydroxyl group and is base labile- the peptide can be cleaved from the resin using strong basic conditions;[21] b) Wang resin with a hydroxyl functionality cleavage of peptides from this resin can be accomplished using strong acidic conditions.

The coupling of Fmoc-Asp-O-2-PhiPr-OH (D10), required 1.074 g of the Fmoc protected amino acid, 302 μL of pyridine, 325 μL of DCBC, and approximately 10 mL of DMF. These reagents were added directly to the vessel and allowed to shake for a minimum of 18 hours per attempt. After each attempt, the resin was washed 5 times with DMF, and a small portion of the resin was removed and dried in a fume hood for the substitution check. Multiple attempts were needed to couple the D10 residue to the resin, but after 5 attempts, the molar substitution (measured using methods outlined in section 1.5 and used for all other substitution checks done for this synthesis) was 0.5605 mmol/g (63.9% substitution of the previously available sites). Even though this is below the benchmark set earlier in the experiment, the 3rd, 4th and 5th substitution checks did not show any change in the substitution and the synthesis of the peptide was continued. Based on this substitution, 397 μL of DIPEA and 214 μL of acetic anhydride was added to the vessel and shaken for an hour to cap all of the remaining unreacted sites of the resin. A substitution check was done after this capping process, and no change in the molar substitution was observed. The resin-bound amino acid was treated with approximately 10 mL of 20% piperidine in DMF to cleave the Fmoc protecting group. After numerous

piperidine treatments, the substitution check for Fmoc removal gave a negative value, which indicated that no Fmoc remained, allowing the next amino acid to be added to the chain.

In addition to the change of resin being employed, the HOBt and HBTU method has been replaced with simply HCTU. Using a single reagent instead of two is a cleaner and quicker process and has the same function as the HOBt/HBTU method. The coupling of the second amino acid, Fmoc-Lys(Boc)-OH (K9) required 0.6767 g of the Fmoc protected amino acid, 0.5975 g of HCTU and 506 μL of DIPEA. These reagents were weighed and then added directly into the Merrifield vessel along with approximately 10 mL of DMF to dissolve the reagents. The reaction mixture only required 2 hours of shaking per coupling attempt which allowed for more attempts each day. Following every attempt, the resin was washed 5 times with DMF and then a small portion of the resin was removed for the substitution check. In total, 7 attempts to couple the K9 were needed to give a molar substitution of 0.4512 mmol/g (80.5% substitution of the previously available sites). After the substitution was deemed sufficient for the coupling, the Fmoc protecting group was cleaved off K9 with 20% piperidine in DMF allowing the next amino acid to be added to the chain.

Synthesis of the peptide continued with the addition of Fmoc-Ile-OH (I8) to elongate the chain further. This step required 0.4108 g of I8, 0.4809 g of HCTU and 407 μL of DIPEA for the reaction. The reagents were added into the Merrifield vessel with approximately 10 mL of DMF and allowed to shake for a minimum of 2 hours per attempt. After each attempt, the resin was washed 5 times with DMF, and a small amount of the resin was removed to do a substitution check. After 7 attempts, the coupling showed great success with a molar substitution of 0.4368 mmol/g (96.8% substitution of the previously available sites). After a sufficient coupling for this amino acid, the peptide was treated with 20% piperidine in DMF to

cleave the protecting group. Once a negative value was obtained for the absorbance reading, the next amino acid was added to the chain.

The Fmoc-Arg(Pbf)-OH (R7) residue was the next amino acid to be added to elongate the chain. This step in the synthesis required 0.7301 g of R7, 0.4656 g of HCTU and 394 μL of DIPEA which was all added to 10 mL of DMF in the Merrifield vessel. The vessel was allowed to shake for a minimum of 2 hours per attempt and after each attempt, the resin was washed through 5 times with DMF. A small amount of resin was taken out of the vessel and dried for a substitution check. The coupling showed quite good success after 9 attempts, achieving a molar substitution of 0.3631 mmol/g (83.1% substitution of the previously available sites). The substitution level was sufficient for the next amino acid to be added, so the Fmoc protecting group was cleaved with 20% piperidine in DMF allowing for the next amino acid to be coupled.

Chain elongation continued with the addition of Fmoc-Gln(Trt)-OH (Q6) to the peptide. The addition of this residue required 0.5713 g of the Fmoc protected amino acid, 0.3870 g of HCTU and 328 μL of DIPEA for the reaction. All of these reagents were added into the Merrifield vessel along with approximately 10 mL of DMF as the reaction solvent. The vessel then was shaken for at least 2 hours in each attempt and followed by washing with DMF 5 times. A small amount of the resin was removed for substitution checks after each attempt, and after 9 attempts, the molar substitution was measured to be 0.2917 mmol/g (80.3% substitution of the previously available sites). This was sufficient for the next amino acid addition, so the resin was treated with 20% piperidine in DMF to remove the Fmoc protecting group. Fmoc-Val-OH (V5) was the next amino acid in the sequence of the peptide being synthesized. This step required 0.2551 g of V5, 0.3109 g of HCTU and 263 μL of DIPEA - all dissolved in 10 mL of DMF which was placed directly into the Merrifield vessel. All of these reagents required a

mixing time of at least 2 hours for each coupling, and after each shaking period, the resin was washed 5 times with DMF, and then a small amount of the resin was removed to dry for a substitution check. This amino acid coupling took 4 attempts to achieve a molar substitution of 0.2745 mmol/g (94.1% substitution of the previously available sites). This substitution level was more than acceptable and the next step was cleaving the Fmoc group from the V5 residue so that the next amino acid could be coupled.

The next amino acid that was to be added was another Fmoc-Ile-OH (I4) residue. This step in the synthesis required 0.2500 g of I4, 0.2926 g of HCTU and 248 μL of DIPEA and all reagents were added directly into the Merrifield vessel with 10 mL of DMF. The reagents in the Merrifield vessel were shaken for a minimum of 2 hours for each attempt and were washed 5 times with DMF after each reaction period. A small amount of resin was then removed from the vessel and dried in a fume hood for a substitution check. This coupling attempt took only 3 tries to reach a molar substitution of 0.2677 mmol/g (97.5% substitution of the previously available sites). This substitution was enough to proceed with the cleavage of the Fmoc protecting group (using 20% piperidine in DMF) and the addition of the next amino acid.

Another Fmoc-Arg(Pbf)-OH (R3) residue was added and was the 8^{th} amino acid to be added to the sequence. This step in the synthesis needed 0.4475 g of R3, 0.2853 g of HCTU and 242 μL of DIPEA all placed into the Merrifield vessel along with 10 mL of DMF. The reaction vessel was then shaken for a minimum of 2 hours and followed by washing with DMF 5 times. Each substitution check required a small amount of the resin to be removed and dried after the DMF washings. The coupling of the R3 residue took 7 tries to obtain a molar substitution of 0.2352 mmol/g (87.8% substitution of the previously available sites). The coupling was

sufficient for this step and the resin was treated with 20% piperidine in DMF to remove the Fmoc protecting group, allowing the next amino acid to be coupled.

The next amino acid in the sequence was another Fmoc-Lys(Boc)-OH (K2) residue. For this coupling, 0.2839 g of K2, 0.2507 g of HCTU and 212 μL of DIPEA were added into the Merrifield vessel containing the resin-bound peptide and 10 mL of DMF. Again, a minimum of 2 hours for each reaction was needed to allow the reagents to react and after each 2 hour time period, the resin was washed with DMF 5 times. After each set of washings, a small amount of resin was removed from the vessel and dried for a substitution check. After 5 attempts, the degree of substitution was measured to be 0.2041 mmol/g (86.7% substitution of the previously available sites). This substitution was good and the resin-bound peptide was subjected to 20% piperidine in DMF to remove the Fmoc protecting group to allow the next amino acid to be coupled.

The 10^{th} amino acid added in this synthesis was Fmoc-Phe-OH (F1) and required 0.2307 g of the Fmoc protected amino acid, 0.2175 g HCTU and 184 μL of DIPEA. All of these reagents were weighed out and then added into the Merrifield vessel along with 10 mL of DMF. The vessel was then shaken for at least 2 hours for each coupling attempt and subsequently washed 5 times with DMF. After the washings, a small portion of the resin was removed from the vessel and dried in a fume hood for the coupling check. The coupling of F1 took 5 tries to get a molar substitution of 0.1904 mmol/g (93.3% of the previously available sites). This substitution level was enough for the next amino acid to be coupled so the peptide was treated with 20% piperidine in DMF to remove the Fmoc protecting group.

The next amino acid in the sequence was another Fmoc-Arg(Pbf)-OH (R13) residue. This amino acid needed 0.3138 g of R13, 0.2029 g HCTU and 172 μL of DIPEA per coupling to proceed. With each attempt, the reaction vessel was shaken for 2 hours and then washed with DMF 5 times after the reaction. A small portion of the resin was then removed from the vessel to dry in a fume hood for the substitution check. After 8 couplings, the R13 residue coupled with a molar substitution of 0.1711 mmol/g (89.9% substitution of the previously available sites). With a molar substitution of this level, the Fmoc group on R13 could be cleaved, using 20% piperidine in DMF, and then the next amino acid could be added.

The 12th amino acid to be added in this peptide was Fmoc-Leu-OH (L12). This synthesis step required 0.1558 g of L12, 0.1824 g of HCTU and 154 μL of DIPEA. All of the reagents were weighed out before being added into the Merrifield vessel containing the resin-bound peptide and approximately 10 mL of DMF. Each coupling attempt required a 2 hour shaking period and was followed by washing with DMF 5 times. Next, a small portion of the resin was removed, dried and used in a substitution check. Five coupling attempts were needed for the L12 coupling to achieve a 0.1592 mmol/g molar substitution (92.4% substitution of the previously available sites). This substitution was good and so the resin was then treated with 20% piperidine in DMF to cleave the Fmoc protecting group from the L12 residue allowing for the final amino acid to be added.

The final amino acid to be added for this peptide sequence was another Fmoc-Phe-OH (F11) residue. Each coupling attempt required 0.1579 g of the Fmoc protected amino acid, 0.1686 g of HCTU, and 143 μL of DIPEA. With each attempt, a 2 hour shaking period was needed for the reaction to proceed. After the shaking time, the resin was washed 5 times with DMF and a small portion of the resin was removed after each attempt to perform a substitution

check. In total 6 coupling attempts were done to reach a final molar substitution of 0.1471 mmol/g (93.0% substitution of the previously available sites). With this amino acid being coupled, all of the residues in the sequence had been added to the peptide chain while still bound to the resin.

The next step was to cyclize the peptide connecting the F11 α amino group to the D10 α carboxyl group. The first amino acid in this sequence, D10, had the O-2-PhiPr protecting group on it, which needed to be removed first. Cleavage of the O-2-PhiPr protecting group was carried out by treating the resin-bound peptide with 1% TFA in DCM. Approximately 5 mL of this mixture was placed into the Merrifield vessel and allowed to shake for 2 minutes. After the shaking, the resin was washed 3 times with DMF to remove any excess solvent, reagents, and any of the O-2-PhiPr that may have been cleaved off. This cleavage of the O-2-PhiPr protecting group was done 10 times to leave a free α carboxyl group. The Fmoc protecting group on F11 was the next to be cleaved; this required approximately 8 mL of 20% piperidine in DMF. The reaction vessel was shaken for approximately 30 minutes to remove all the Fmoc protecting group from the peptide. After a few attempts, the Fmoc check showed a negative absorbance for the UV measurement, indicating no Fmoc remained on the peptide. With a free α amino and α carboxyl group on the peptide, the "head-to-tail" cyclization of the peptide could be carried out. Based on the final substitution on the F11 residue, the amounts of HCTU and DIPEA for the cyclization were determined to be 0.1568 g and 133 μL, respectively. This cyclization process was treated in the same manner as the coupling of an amino acid, so the after the reagents were added into the vessel, it was shaken for 2 hours and then washed 5 times with DMF. There was no way to check the substitution of the peptide, so the reaction was carried out 4 times to attempt to complete the cyclization of the peptide.

After assuming the peptide was cyclized, the final step was to cleave all of the protecting groups from all the remaining residues. The remaining protecting groups (Pbf, Boc and Trt) all required strongly acidic conditions to be cleaved. To remove these protecting groups, the resin was treated once with a mixture of 95:5 TFA:TIS. The vessel was allowed to shake for 30 minutes and was then washed 3 times with TFA, then DCM, followed by 10% DIPEA in DCM, and finally DCM again. The resin was then swollen in DMF for 24 hours, filtered and then allowed to dry in a fume hood until it was completely dry. The final product was the 13 amino acid cyclic peptide FK-13 free from the effects of protecting groups. The dry product was then transferred to a vial and placed in a desiccator.

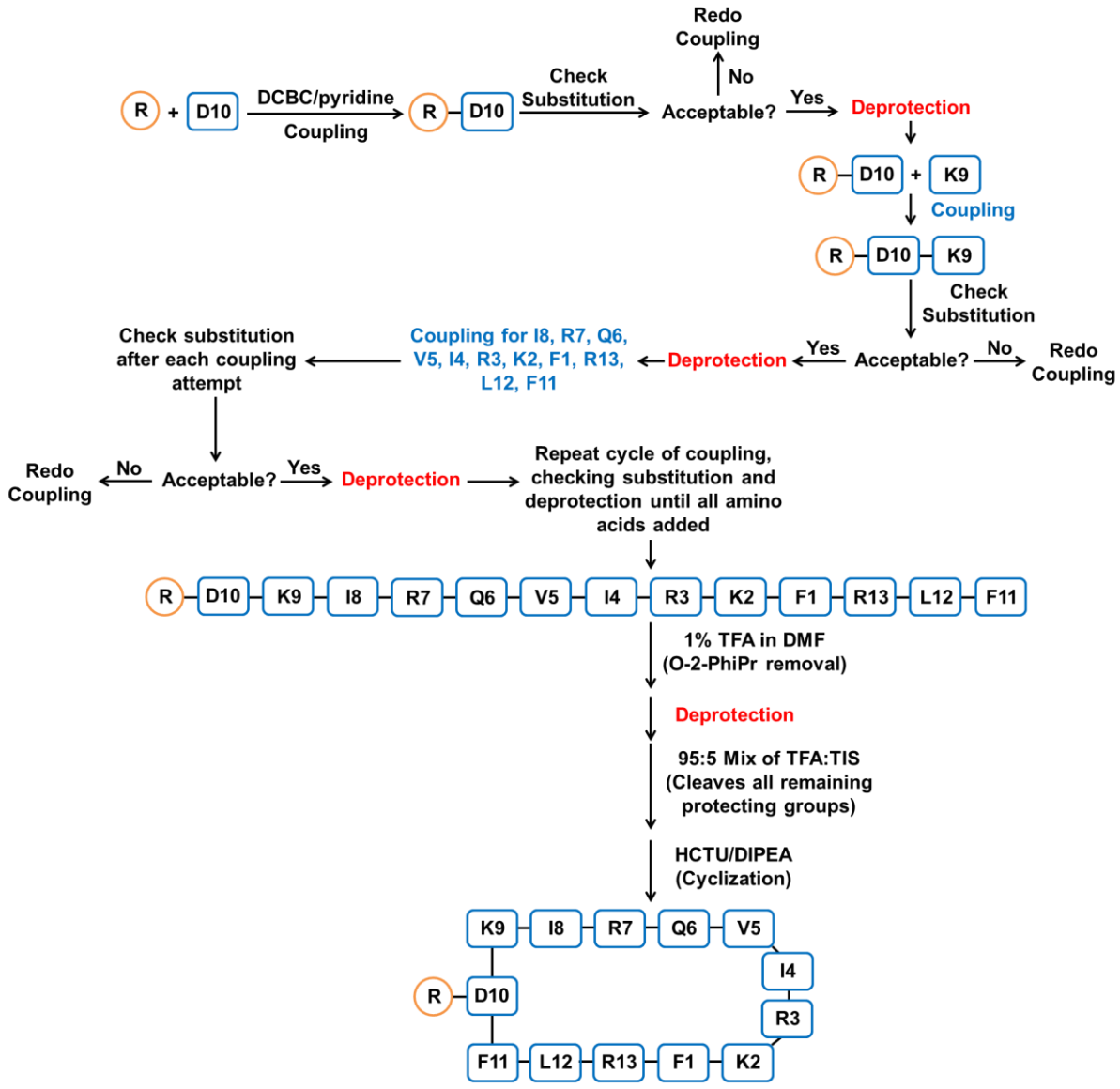


Figure 3.2.2: Flow chart for the synthesis of the cyclic peptide FK-13, using methods adapted from Kates and Alberts[12] and Chan and White[13]. HCTU replaced HOBt and HCTU as the coupling reagents, and each coupling time period only required 2 hours of mixing. In the end, the linear sequence was cyclized on resin free from the effects of protecting groups.

3.2.2 NMR of FK-13

All NMR experiments conducted on the FK-13 samples were done using a Bruker Avance 500 MHz NMR spectrometer at the University of Waterloo's NMR facility. The magnet was equipped with a Doty Scientific triple-tuned DSI-XC4 HRMAS probe equipped with a ^{13}C-^{2}H LF Trap Wand, ^{13}C-^{2}H MF Tune/Voltage Division Wand, ^{13}C-^{2}H LF Tune/Voltage Division Wand and a ^{13}C-^{1}H capacitor. All experiments were conducted at ambient temperature (~295 K). A small amount of the dry resin-bound FK-13 was loaded into a XC4 40 μL Kel-f sealing cell and filled to approximately to half the volume; the remaining volume was filled with DMF-d_7 and was capped with a XC4 Teflon plug. The sample needed to have no air bubbles in the cell; if there was air present, it would have to be reopened and filled again until no bubbles remained. The cell was loaded into a Doty Scientific XC4 4-mm silicon nitride thin wall rotor and capped with Torlon short turbine caps, which was then placed into the probe and spun at a speed of 4.4-4.7 kHz at the magic angle of 54.7°.

One-dimensional ^{1}H NMR spectra, plotted as spectral intensity versus frequency, were obtained using the dipsi21D pulse sequence. The spectrum was collected after 1024 transients, with an acquisition time of 0.62 seconds and a spectral width of 6614 Hz. A 90° ^{1}H pulse length of 7.81 μs was used at low power (0.00 dB). The resulting 8k data points were zero-filled to 16k data points for Fourier Transformation and apodized using the exponential multiplication window function with a Lorentzian broadening constant of 0.30 Hz, sine bell shift of 0, and a Gaussian broadening factor of 0. The spectrum was referenced internally to the chemical shifts of the deuterated DMF solvent, and applied to all other 1D and 2D ^{1}H experiments. All spectra were processed using the Bruker XWIN-NMR 3.5 pl6 software.

Two-dimensional ^{1}H-^{1}H COSY experiments used the cosyqf90 pulse sequence and collected a set of 256 x 1024 data points in the t_1 and t_2 dimensions, respectively. Each increment consisted of 136 transients with 4 dummy scans over a spectral width of 5000 Hz in both dimensions. A 90° ^{1}H pulse length of 7.81 μs was applied at a power level of 0.00 dB with an acquisition time of 0.1025 seconds. A relaxation delay of 2 seconds with delay increments of 0.166 ms and a dwell time of 83.2 μs were used in the pulse sequence as well. After acquisition and Fourier transform was applied, the data sets underwent apodization using the SINE multiplication function with a Lorentzian broadening constant of 0.30 and 1.00 Hz in the t_1 and t_2 dimensions, respectively. A sine bell shift of 0 was used in both dimensions and a Gaussian broadening factor of 0.1 was used in the t_1 dimension, and 0 in the t_2 dimension. The spectrum was calibrated using the solvent reference value obtained from the 1D ^{1}H experiments and applied to both dimensions. The spectrum was processed using the Bruker XWIN-NMR 3.5 pl6 software.

Two-dimensional ^{1}H-^{1}H TOCSY experiments utilized the DIPSI-2 pulse sequence and collected a set of 256 x 1024 data points in the t_1 and t_2 dimensions, respectively. A total of 128 transients and 16 dummy scans per increment were used over a spectral width of 5500 Hz in both dimensions. A 90° ^{1}H pulse length of 7.69 μsec was used at low power (0.00 dB) and a low power 90° pulse of 60 μs was used at a power level of 18.41 dB for the TOCSY-spin lock. A TOCSY mixing time of 35 msec and an acquisition time of 0.09353 seconds were used in the experiments. Delay increments of 0.18177 ms, a dwell time of 91.2 μs and a relaxation delay of 5 seconds were also implemented in the pulse sequence. Upon completion of acquisition, the data sets were Fourier transformed and apodized using the QSINE multiplication function with a Lorentzian broadening factor of 0.30 and 1.00 Hz in the t_1 and t_2 dimensions, respectively. In

both dimensions a sine bell shift of 2 was applied and Gaussian broadening factor was set to 0.1 in the t_1 dimension and 0 in the t_2 dimension. The TOCSY spectrum was referenced to the value obtained from the 1D ^{1}H experiments conducted before this experiment. The spectrum was processed using the Bruker XWIN-NMR 3.5 pl6 software.

Two-dimensional ^{1}H-^{1}H ROESY experiments used the roesyph pulse sequence and collected a set of 256 x 1024 data points in the t_1 and t_2 dimensions respectively. A total of 128 transients and 16 dummy scans per increment were used over a spectral width of 5500 Hz in both dimensions. A 90° ^{1}H pulse length of 7.69 μs was used at a power level of 0.00 dB along with a 65 ms pulse for the ROESY spin-lock at a power level of 28.30 dB. An acquisition time of 0.09353 seconds was used along with a 20 μs power switching delay between low and high power pulses. After acquisition and Fourier transform, the data sets were apodized using the QSINE multiplication function with a Lorentzian broadening factor of 0.30 and 2.00 Hz in the t_1 and t_2 dimensions respectively. A sine bell shift of 2 was applied in both dimensions and a Gaussian broadening factor of 0.1 in the t_1 and 0.05 in the t_2 dimensions. The spectrum was referenced to the value obtained from the 1D ^{1}H experiment conducted earlier. The spectrum was processed using the Bruker XWIN-NMR 3.5 pl6 software.

3.3 FK-13 Results

3.3.1 Results of Synthesis

The results of the solid-phase synthesis of the 13-residue peptide are summarized in Table 3.3.1. Coupling of the first amino acid Fmoc-Asp(O-2-PhiPr)-OH (D10) to the HMBA-AM resin was done using the DCBC method in DMF. The use of this amino acid lends itself well for cyclic synthesis as the α-carboxyl group is protected and the γ-carboxyl group is free to react with the base-labile resin. Linkage through the side chain allows for the elongation to occur on the α-amino group of D10 and eventually cyclizing through the α-carboxyl group once all the residues have been added. The initial functionality of the resin was assumed to be 0.88 mmol/g and after five attempts, the coupling was measured to be 0.5605 mmol/g. This coupling was well below the ideal percentage of 80%, and this could be attributed to the use of DCBC in the reaction. The acetyl chloride group of 2,6-dicholorbenzoyl chloride acts in the same manner as acetic anhydride for the capping process. As the coupling reaction occurs, the unreacted sites are simultaneously being capped as well. Any attempts after the first or second try usually won't show a change in the substitution since the sites have either reacted on already or been capped by the acetyl chloride. The use of the 4-dimethylaminopyridine (DMAP) or 1-(mesitylene-2-sulphony)-3-nitro-1*H*-1,2,4-triazole (MSNT) are better reagents to use for the first coupling as they will not cap the unreacted sites as well. The downside to the DMAP method is that enantiomerization and dipeptide formation can occur, but this can be limited by using a minimal amount of DMAP.[13] MSNT is favourable when a hydroxyl functionality is being used to anchor the first amino acid or when enantiomerization is possible, but this method requires all of the equipment being used to be free of moisture for this coupling.[13] Either of these two methods

should be considered in future synthesis of peptides on resin, especially for longer sequences such as FK-13.

Table 3.3.1: The coupling conditions and success of each coupling attempt in the solid-phase synthesis of the 13-residue peptide, FK-13. The percentage of total available sites was based off the initial molar substitution of 0.88 mmol/g

Amino Acid	Coupling Method	Solvent	# of Coupling Attempts	Resin Substitution (mmol/g)	% Subs of Previous Available Sites	% Subs of Total Available Sites
1st: Asp (O-2-PhiPr)	DCBC	DMF	5	0.5605	63.9%	63.9%
2nd: Lys (Boc)	HCTU	DMF	7	0.4512	80.5%	51.2%
3rd: Ile	HCTU	DMF	7	0.4368	96.8%	49.6%
4th: Arg (Pbf)	HCTU	DMF	9	0.3631	83.1%	41.3%
5th: Gln (Trt)	HCTU	DMF	9	0.2917	80.3%	33.1%
6th: Val	HCTU	DMF	4	0.2745	94.1%	31.2%
7th: Ile	HCTU	DMF	3	0.2677	97.5%	30.4%
8th: Arg (Pbf)	HCTU	DMF	7	0.2352	87.8%	26.7%
9th: Lys (Boc)	HCTU	DMF	5	0.2041	86.7%	23.2%
10th: Phe	HCTU	DMF	5	0.1904	93.3%	21.6%
11th: Arg (Pbf)	HCTU	DMF	8	0.1711	89.9%	19.4%
12th: Leu	HCTU	DMF	5	0.1582	92.4%	18.0%
13th: Phe	HCTU	DMF	6	0.1471	93.0%	16.7%

Table 3.3.1 shows that every step in the synthesis was successful and reached the desired 80% substitution of the previously available sites. The shorter and simpler amino acids such as Ile, Val, Phe, and Leu were all 90% coupling at each of the steps. This could suggest that amino acids that do not have protecting groups can couple easier and require fewer attempts to achieve a high substitution. On the other hand, while not being lower than the desired 80%, the longer amino acids with protecting groups had the lowest substitution percentages. Opposite to what the short non-polar amino acids suggest, it is possible that longer polar amino acids could somehow interfere with the coupling process due to their flexibility and bulkiness of the protecting groups. Overall the chain elongation of FK-13 was successful and the cyclization of the peptide was carried out.

The D10 residue was protected with an O-2-PhiPr group that needed to be removed with 1% TFA in DCM. Mild acidic conditions were used to ensure that the other protecting groups (Pbf, Trt and Boc) would remain and the functionalities they were protecting would not interfere with the cyclization process. The resin-bound peptide was treated with the mild acid solution 10 times and allowed to shake for 2 minutes each time. After each treatment, the resin was washed with DMF 3 times to remove any excess TFA and the cleaved O-2-PhiPr group. A substitution check was done after to measure the coupling of the peptide to ensure nothing had changed when the cleavage was carried out. This check gave a substitution of 0.1375 mmol/g, which is 93.4% of the final substitution and was sufficient to carry on with the cyclization. This difference in the substitution could be attributed to the fact that only one check was done as opposed to three due to the limited amount of resin-bound peptide available or that a portion of the resin that was removed could have had less peptide than other parts of the resin. Removal of the Fmoc protecting group from F11 was carried out using 20% piperidine in DMF and was

done 3 times to ensure as much of the Fmoc group would be removed. A substitution check was done after these attempts and the absorbance reading was negative after the final two attempts indicating that no Fmoc remained. The cyclization process was carried out by joining the α carboxyl group of D10 to the α amino group of F11 using 5 molar equivalents of HCTU and 10 molar equivalents of DIPEA and reacted for 2 hours on the shaker. This process was repeated 4 times and after each attempt, the resin was washed 4 times with DMF. The final step after cyclization was to remove all the protecting groups from the peptide with a 95% TFA 5% TIS solution. This mixture was added to the Merrifield containing the resin-bound peptide and allowed to shake for 1 hour. After the resin was washed 3 times with TFA and the filtrate was collected a put aside in case the peptide was accidentally cleaved from the resin. The resin was then washed 3 times with DCM, 10% DIPEA in DCM and finally in DCM again. The resin was then allowed to dry under suction filtration for 3 hours then placed in the fume hood to dry. This synthesis gave a cyclic resin-bound sample of FK-13 that could be analyzed by NMR spectroscopy.

3.3.2 NMR Analysis of FK-13

Assignment of all the protons in the 13-residue peptide was done using 1D ^{1}H and 2D ^{1}H-^{1}H experiments and the results are tabulated in Table 3.3.2. The dipsi21D pulse sequence was used to obtain the 1D spectrum (Figure 3.3.1) and was a better choice than the regular zg pulse sequence that was used for the GSH experiments. This pulse sequence suppresses the signal from the resin and the signals from the peptide are not buried under the intense resin signals.

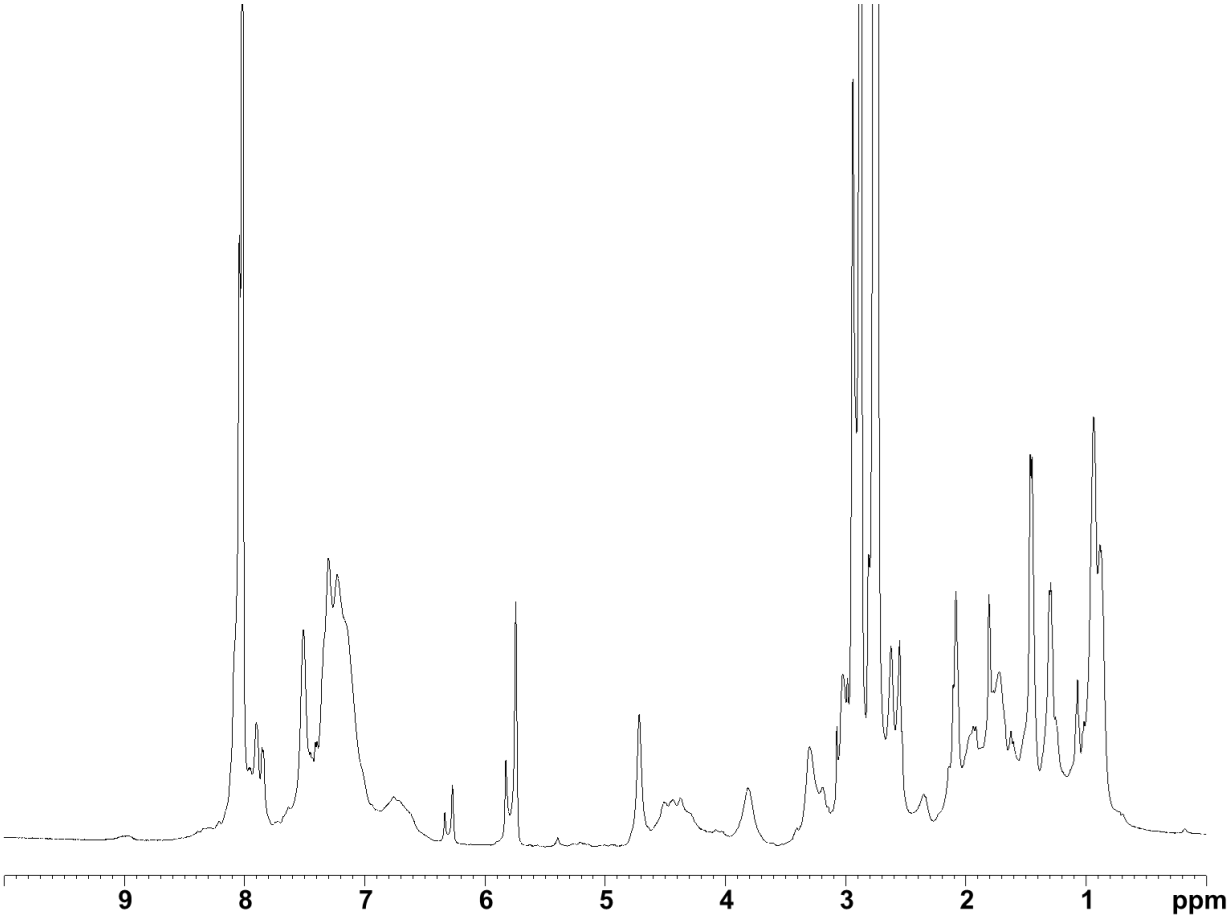


Figure 3.3.1: 1D spectrum of the resin-bound FK-13 where chemical shift is plotted against spectral intensity. The dipsi21D pulse sequence was used to supress the signal from the resin so that the peptide frequencies would be easier to visualize.

Using the COSY and TOCSY spectrum (Figure 3.3.2) 13 different spins systems were identified where an amide proton interacts with an α proton.

Table 3.3.2: ^{1}H chemical shifts of the 13-residue cyclic peptide, FK-13. The chemical shifts of these protons were referenced internally to the solvent peaks of DMF-d_7

Amino Acid	NH	Hα	Hβ	Hγ	Hδ	Other
D10	8.030	4.322	3.288	-	-	-
K9	8.303	4.702	3.114	1.504	1.727	3.387 (ε)
I8	8.330	4.432	2.326 1.738 (β')	1.614	1.337	-
R7	8.158	4.378	1.940	1.309	3.285	-
Q6	8.129	4.525	3.123	1.952	-	-
V5	8.557	4.291	1.906	0.929	-	-
I4	7.754	3.842	2.875 1.172 (β')	2.020	0.937	-
R3	8.975	4.534	1.712	1.043	3.146	-
K2	7.851	4.506	1.923	1.445	1.723	3.157 (ε)
F1	8.022	4.757	2.767	-	-	-
R13	8.411	3.831	2.015	1.179	2.863	-
L12	7.936	4.283	2.141	1.867	0.935	-
F11	8.022	4.757	2.767	-	-	-

Identifying the first four residues, (D10, K9, I8 and R7) were simple since a sample of the peptide was analyzed after these four residues were coupled to the resin. In these four residues, the D10 system was easy to identify as it was the only one that had amide, α and β interactions, but no γ or further interactions. The amide proton (8.030 ppm) and the α proton (4.322 ppm) showed a strong correlation and the β proton (3.288 ppm) showed a strong correlation to the α proton. Differentiating between the K9, I8 and R7 residues was quite challenging since these residues are all quite long and have many groups of protons. One residue was observed to have an interaction between the α proton and an ε proton and was determined to belong to K9. Four interactions that lined up with the α proton (4.702 ppm) in the "α region" belonged to K9 and were between the β (3.114 ppm), γ (1.504 ppm), δ (1.727 ppm), and ε (3.387 ppm) protons. Differentiating between I8 and R7 came down to the various interactions between the α proton

and the β, γ, and δ protons. One α proton was seen to interact with a set of protons at 2.326, 1.738, 1.614, and 1.337 ppm. Another α proton was seen to interact with a set of protons at 1.940, 1.309 and 3.285 ppm. The major difference between these two sets of protons is that the second set contains a proton at 3.285 ppm which would not occur in a non-polar aliphatic amino acid such as isoleucine. Therefore, the α proton (4.378 ppm) that interacts with protons at 3.285 ppm, must belong to the R7 residue and in particular the δ protons of this residue. The β protons at 1.940 ppm and γ protons at 1.309 ppm also have interactions with this α proton. I8 is a non-polar aliphatic residue and it is expected that the frequency for all these protons appear less than 3 ppm. The α proton (4.432 ppm) is interacting with a β proton (2.326 ppm), γ proton (1.614 ppm) and a δ proton (1.337 ppm). There is also another interaction at 1.738 ppm and this belongs to the methyl group coming off the β carbon and will be denoted as β'. Identification of the remaining amino acids was also challenging – the spectrum was cluttered with peaks that represented the other 9 amino acids with peaks possibly overlapping one another. Distinguishing between certain residues was simpler than others. Phenylalanine was the only remaining residue that had an amide, α and β protons (along with aromatic protons) and so assigning shifts to this residue were simple. An amide proton (8.022 ppm) was seen interacting with an α proton (4.757 ppm) in the amide region of the TOCSY spectrum. This α proton was also seen interacting with a β proton at 2.767 ppm, but no other residues. This set of shifts was assigned to F1 and F11 under the assumption that these residues are close and that the peaks are stacked on top of one another. Valine was another residue that was readily assigned as it was a short residue with an aliphatic side chain. The α proton (4.291 ppm) was seen to be interacting with a β proton (1.906 ppm) and another group far upfield below 1 ppm. This signal was determined to belong to the two methyl groups of valine and represents the γ protons (0.929

ppm) of the V5 residue. Q6 was another residue that was simple to assign as it only has amide, α, β, and γ protons, where one of the proton groups will be shifted downfield more due to the proximity of the amide in the R group. The α proton at 4.525 ppm was seen interacting with a signal at 3.123 ppm and 1.952 ppm. The fact that these protons were shifted this high suggests that these signal do not belong to a residue such as isoleucine or leucine. Therefore, the peak at 1.952 ppm belongs to the β protons and the peak at 3.123 ppm belongs to the γ protons. Two residues that were difficult to identify from one another were I4 and L12, as they are simply isomers of one another. The key to separating the two in terms of assignment was understanding that L12 would only contain 5 sets of signals (amide, α, β, γ, and δ) while I4 would have 6 sets of signals (amide, α, β, β', γ, and δ). An α proton (4.283 ppm) that was seen interacting with an amide proton (7.936 ppm) was also seen interacting with only 3 other signals, 2.141 ppm, 1.867 ppm and 0.935 ppm. This set of 5 peaks definitely belongs to L12 and was assigned as such. These three signals likely represent the β, γ and the two δ methyl groups in L12. As for I4, an amide proton at 7.754 ppm was interacting with an α proton at 3.842 ppm. This α proton is interacting with a β proton (2.875 ppm), β' proton (1.172 ppm), a γ proton (2.020 ppm) and a δ proton (0.937 ppm). These four α cross-peaks for I4 helped distinguish it from the L12 residue (which only had three α cross-peaks). The remaining residues to be assigned are R3, K2 and R13, where K2 will have 6 signals as opposed to R3/R13 having 5. The K2 residue was identified after the amide proton at 7.851 ppm was seen interacting with an α proton at 4.506 ppm. Analysis of the "α region" using the shift for the α proton revealed 4 other interactions to this proton. The β proton was seen at 1.923 ppm, the γ proton at 1.445 ppm, the δ proton at 1.723 ppm and the ε proton at 3.157 ppm. The δ and ε protons appear shifted downfield more than the γ proton due to the amino functionality in the K2 side chain. The final two arginine

residues, R3 and R13, were identified on the 5 different proton signals for these residues. R3 had an interaction between an amide proton at 8.975 ppm with an α proton at 4.534 ppm. This α proton was then seen to interact with β (1.712 ppm), γ (1.043 ppm) and δ (3.146 ppm) protons. These shifts were not very different from the R7 residue that was assigned earlier. The R13 residue had an amide proton at 8.411 ppm interacting with an α proton at 3.831 ppm. This α proton was interacting with a β proton at 2.015 ppm, a γ proton at 1.179 ppm and a δ proton at 2.863 ppm. These protons seem to be shielded and appear at a lower chemical shift than any of the other arginine residues, but this could be due to the nearby F1 residue and the large aromatic ring that is next to this residue. The orientation of the aromatic ring could induce a shielding effect on the protons of the R13 residue which make it seem to be at a lower shift than expected. With all 13 amino acid residues identified, the structure of the peptide on the resin can be proposed using the ROESY spectrum.

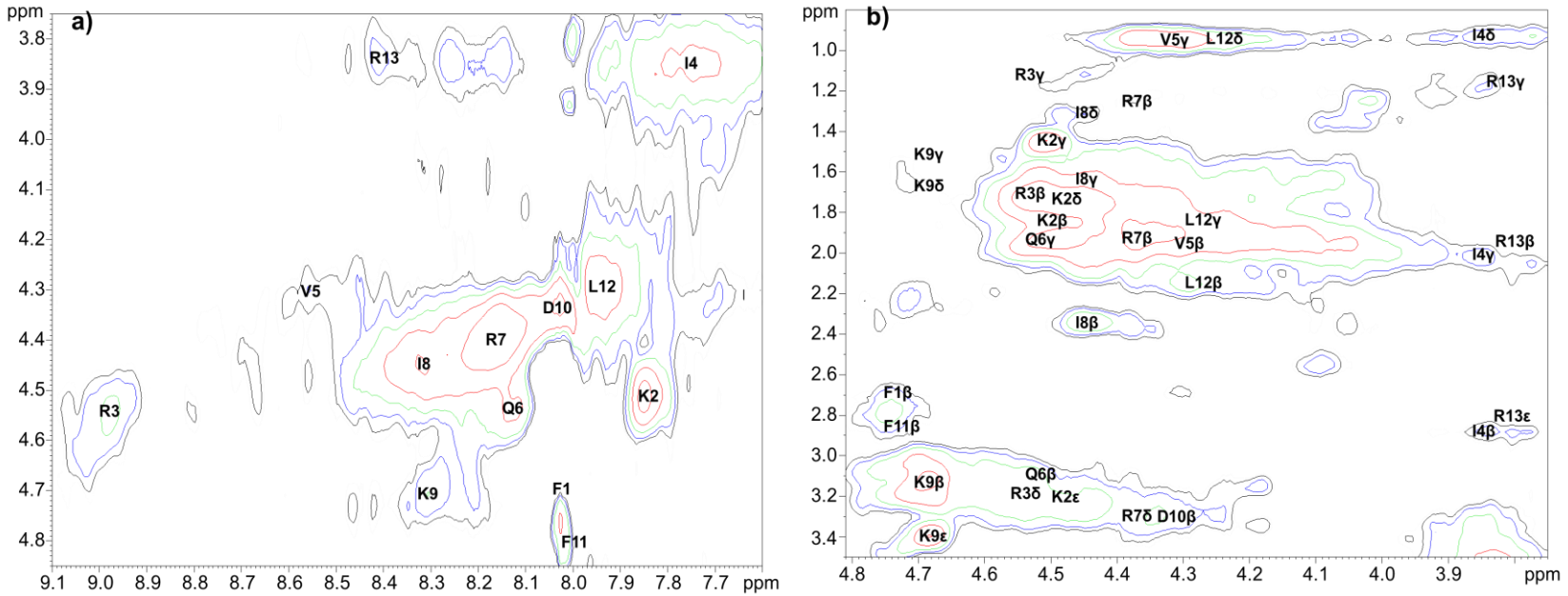


Figure 3.2.2: Expanded views of the amide and α regions in the TOCSY spectrum for FK-13. In total, 13 spin systems were identified and the protons in their respective R groups were assigned ^{1}H chemical shifts using this experiment. A) The amide region where all the backbone interactions were assigned. The peaks are interactions are between the NH and Hα of the labelled amino acid. For example, a peak labelled D10 would be an interaction between the amide proton of D10 and the α proton. B) The α region was used to assign the shifts of the R groups for the different spin systems. Each label correlates to the α proton interacting with the indicated proton, i.e. a peak labelled D10β, is an interaction between the α proton of D10 and its β proton.

The ROESY spectrum had 14 different cross-peaks, representing 14 different through-space interactions. These interactions are summarized in Table 3.3.3. Of these interactions, two were within the same spin system, five were between sequential amino acids and the remaining seven were through-space interactions between different amino acids. The two interactions within the same spin system, interactions 6 and 7, are for I8 and R3, respectively. In both cases the amide proton and the α proton of these residues are interacting, but do not provide much evidence to the structure of the peptide. The five interactions between sequential amino acids can provide some insight to the structure of the peptide. Unfortunately only 5 of the possible 13 interactions were seen, which might be attributed to the very low substitution of the peptide on the resin. Regardless, these five interactions (1,2,5,9 and 10) are still useful in the understanding of the peptide structure. All of these interactions were between the amide proton and the α proton of another residue. Interaction 6 between the K9 amide proton and the I8 α proton indicates that these two amino acids that were joined at the beginning of the synthesis was done correctly and that up to this point the synthesis was going well. The interaction between the I8 amide proton and the R7 α proton (interaction 11) was another indication that the synthesis was going smoothly. With the continuation of the synthesis the R7 amide proton was seen interacting with the α proton of Q6 suggesting that after the four amino acids were analyzed by NMR, the synthesis continued successfully. The last sequential interaction was interaction 1 and involved the amide proton of Q6 and the α proton of V5, which shows the valine residue was successfully coupled to the glutamine residue that came before it. These five interactions suggest that the first 5 amino acids were coupled successfully, but no other sequential interactions can suggest that the peptide was fully synthesized.

Table 3.3.3: The through-space NOE interactions observed in the ROESY spectrum from the FK-13 peptide. Interactions listed in bold red are those between different amino acids. The ROESY spectrum was referenced internally to the solvent peaks of DMF-d_7

Interaction	Chemical Shift F1 (ppm)	Chemical Shift F2 (ppm)	Amino Acid 1	Amino Acid 2
1	**8.115**	**4.297**	**Q6 NH**	**V5 Hα**
2	**8.156**	**4.456**	**R7 NH**	**I8 Hα**
3	**8.197**	**4.572**	**R7 NH**	**R3 Hα**
4	**8.105**	**4.497**	**Q6 NH**	**I8 Hα**
5	**8.280**	**4.451**	**K9 NH**	**I8 Hα**
6	8.325	4.468	I8 NH	I8 Hα
7	9.011	4.614	R3 NH	R3 Hα
8	**8.067**	**4.509**	**D10 NH**	**K2 Hα**
9	**8.153**	**4.542**	**R7 NH**	**Q6 Hα**
10	**8.219**	**4.405**	**I8 NH**	**R7 Hα**
11	**8.404**	**4.477**	**R13 NH**	**K2 Hα**
12	**8.538**	**4.359**	**V5 NH**	**R7 Hα**
13	**8.495**	**4.437**	**R13 NH**	**I8 Hα**
14	**8.592**	**4.494**	**V5 NH**	**K2 Hα**

The interpretation of the through-space interactions can verify if the synthesis was done correctly and if the 13-residue peptide has a unique structure. The seven through-space interactions can give some insight into the structure of the peptide and whether the synthesis was done successfully. Interactions 3, 4, 8, 11, 12, 13, and 14 were between amino acids there were not adjacent to one another. Interactions 4 and 12 are between amino acids that are separated by two residues. These interactions are between an amide proton of one residue and an α proton that is two residues away. Two of the interactions: interaction 4 which was between the amide proton of Q6 and the α proton of I8 and interaction 12 between the amide of V5 and the α of R7. These interactions are possible if the backbone begins to turn at these residues bringing the residues into close contact with one another. Since these 4 amino acids are sequential, it could be that this portion of the peptide has an arc conformation or a curve that is

part of the whole peptide. Interaction 11 is also between amino acids that are two residues apart from one another. The amide proton of R13 and the α proton of K2 were seen interacting with one another. These two amino acids were at the end of the synthesis of this peptide and could be coupled if the peptide were also turned at this point near the end of the synthesis. The coupling of the F11 residue to the D10 residue to form the cyclic peptide could influence the two residues to come in closer contact. If this peptide were linear, these interactions would be too far in space to be seen in ROESY. An explanation for these couplings appearing would be that the peptide is not linear and has some sort of turning conformation, whether that conformation is completely cyclic, a simple β turn or an α-helix cannot be determined using just these two interactions. The other four interactions (3, 8, 13 and 14) are through-space interactions that are between residues that are not very close in space to one another if the peptide had adopted a linear conformation. Interaction 3 is between the amide of R7 and R3 α proton, suggesting that this loop must be tightly packed in order for these backbone protons to be interacting. Depending on the actual conformation of the peptide, this interaction is hypothesized to be one that is bringing one portion of the chain in close contact to the other. Interaction 14 is an interaction between the amide proton of V5 and the α proton of K2. Knowing that interactions 4 and 12 potentially make up a portion of the peptide that is starting to make a turn, it could be that K2 is at the end of this turn portion of the peptide as it extends to the D10 residue. If this is the case, then the V5 and K2 residues would be somewhat close in space and would suggest that the residues I4 and R3 would be the last bit of the turn before K2. A very interesting interaction that seems to bring opposite regions of the peptide in close contact with one another is interaction 8. The amide proton of D10 and the α proton of K2 are seen interacting through-space with one another. This interaction may be more useful in hypothesizing a structure for this peptide, since the first

amino acid added in the synthesis is seen interacting to one that was added near the end. If these two residues are indeed interacting, it suggests that the chain has to be coming back around to form the peptide bond between F11 and D10 to form a cyclic peptide. This interaction also rules out the possibility of the peptide forming the natural α-helix since this interaction is beyond the *n*+3 limit of ROESY peaks for helices.[67]Interaction 13 is another interesting interaction and also suggests that the peptide has adopted a cyclic conformation. This interaction is between the R13 amide proton and the I8 α proton. The I8 residue was one of the first ones added in the synthesis and the R13 was one of the last, so seeing this interaction highly suggests that the peptide has come around full loop to form the peptide bond between the D10 residue and the F11 residue. Assuming that the peptide was linear, none of these interactions would be possible, especially interactions 8 and 13, because the distance between them would be too far for ROESY to detect. Also, if the peptide were just an α-helix, some of the interactions violate the *n*+3 interaction restraint that is seen in helices.[67] There is a possibility that the peptide has formed a β-turn as opposed to being cyclic, but this is unlikely. β-turns usually have a stabilizing factor between the opposite ends of the chain that bring the two ends closer together and have a repetitive sequence that forms the turning point. The FK-13 sequence does not have a repetitive portion that would create the turn that a β-turn structure needs. Also the many hydrophobic residues would likely be pushed into the core of this peptide, leaving all of the hydrophilic (and polar) residues on the outside of the chain. For a β-turn the polar residue side chains facing in towards the centre of the peptide would be ideal so that hydrogen bonds could be formed and stabilize the β-turn structure. It would be expected that many interactions with the polar side chains be observed in the spectrum, but none were observed in the ROESY spectrum. Although the ROESY was not sensitive enough to detect interactions between the R groups of the different

residues, due to the low substitution of the peptide, the evidence from the backbone interactions suggest that the peptide does adopt a cyclic conformation. This cyclic conformation is based on interactions 8 and 13 from Table 3.3.3 and these interactions eliminate the possibility of other conformers from being formed.

Further studies into the structure should be considered for examining this peptide. The synthesis of the cyclic product could be done with higher efficiency using a different starting material to ensure the substitution is higher. A higher substitution should give more signals in the ROESY spectrum that could be used to better deduce the structure. More through-space interactions in the ROESY could show the desired D10-F11 interaction along with other interactions to understand the orientation of the specific R groups and backbone conformation. Also the proximity of the side chains within the core of the cyclic peptide could be revealed by analyzing the α region of the peptide. No interactions were seen in this ROESY spectrum, possibly due to poor sensitivity. These initial studies could be expanded upon to probe the success of the synthesis and whether the peptide cyclized or not through different NMR techniques and computational studies.

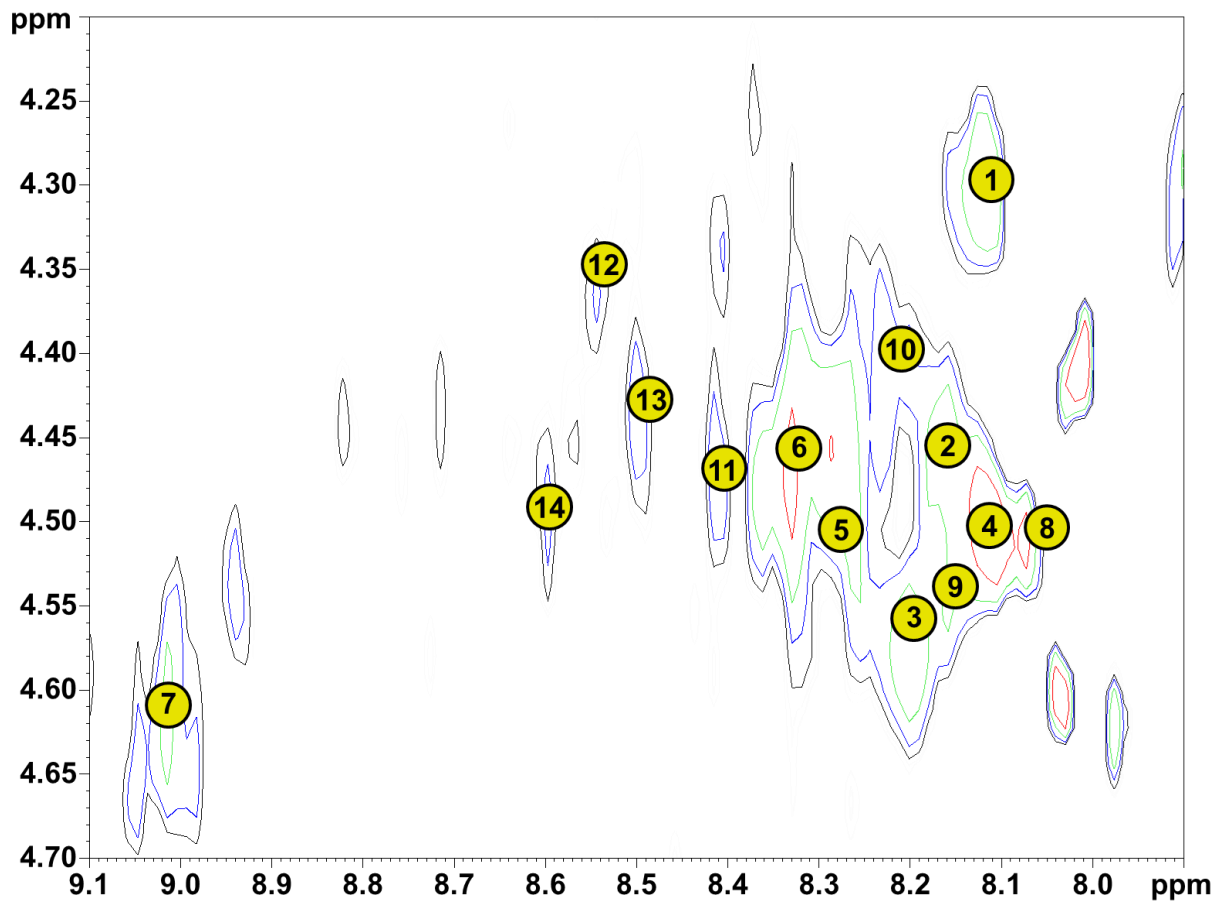


Figure 3.3.3: Amide region of the ROESY spectrum for the resin-bound FK-13 peptide. Each interaction labelled corresponds to those in Table 3.3.3. In total 14 interactions were observed where 2 were within the same spin system and the remaining 12 were between different amino acids. Using the 12 interactions, the three-dimensional structure of FK-13 could be hypothesized.

3.4 Discussion

The synthesis of FK-13 was done using a modified solid-phase peptide synthesis technique. The use of an HMBA-AM resin allowed for amino acids that contained acid-labile protecting groups to be used. Each of the amino acids that had a protecting group, required harsh acidic conditions (95% TFA) to be cleaved from the residues (with the exception of the D10 residue). Using the HMBA-AM resin avoids the use of protecting groups that require different cleavage conditions since the resin requires strong basic conditions to cleave the peptide from the resin. The final substitution on the resin was 0.1471 mmol/g (16.7% of the total available sites assuming the initial molar substitution was 0.88 mmol/g). Compared to other peptides synthesized in the past by the Power group, the substitution is low. Depending on the coupling efficiency of each amino acid throughout the synthesis, the overall coupling begins to significantly drop after 6-12 residues. Peptides containing more than 6 residues are difficult to synthesize manually while trying to achieve a high level of substitution. The reaction sites of the resin become less accessible after numerous reaction cycles, especially if β-branched amino acids are used (such as Ile and Val).[68] Even though the overall substitution was low, the individual steps were done with great success. The non-polar residues all coupled with over 90% efficiency, and those that had protecting groups coupled over 80% as well. The primary issue with the synthesis came in the first step, where D10 was added to the resin using the DCBC method. Future peptide syntheses should not use the DCBC method as it caps the reaction sites as it couples. The use of DMAP or MSNT as the coupling reagent should be favoured, but both of these methods require more materials and time to achieve a successful coupling.[12] Especially in the case for longer peptides that may contain many polar protected

groups, using a coupling reagent that will allow for the highest possible substitution will be favoured over DCBC.

Analysis of the resin-bound FK-13 peptide was carried out using a variety of 1D and 2D ^{1}H NMR experiments. The use of COSY and TOCSY were used to identify the 13 different spin systems for the peptide. Identification of the first four residues was based on results obtained earlier in the study; the remaining nine were identified based on characteristic shifts of the individual amino acids. With the 13 amino acids characterized, ROESY was used to hypothesize a three-dimensional structure of the resin-bound peptide. The intent of this study was to synthesize a cyclic FK-13 so that comparative studies between the Power (NMR) and McMahon (IRMPD) group could be done. In total, 14 through-space interactions were observed in the ROESY spectrum allowing for some structure elucidation. There were two key interactions in this spectrum: interaction 8 which was between the D10 NH and the K2 α, and interaction 13 which was between R13 NH and I8 Hα. These interactions are between amino acids at the start of the synthesis (D10 and I8) and one at the end of the synthesis (K2 and R13). These interactions suggest that the synthesis did work accordingly, and that the peptide did create a head-to-tail peptide bond to form the cyclic peptide. The other interactions in Table 3.3.3 also eliminate the possibility of the peptide being an α-helix (the natural conformation of this sequence), β-sheet or β-turn. Although it is two interactions in the amide region of the ROESY spectrum, early indication does suggest that the peptide achieved a cyclic structure. In addition to performing the synthesis again with a different starting reagent, other NMR experiments should be explored and further elucidate the structure of the resin-bound peptide.

The study of the antibacterial peptide FK-13 was done with the aid of solid-phase peptide synthesis and HRMAS NMR spectroscopy. Further work should be done to improve the

synthesis of the cyclic peptide to achieve higher substitution. With a higher substitution, the NMR experiments that were conducted will have better sensitivity, and more cross-peaks in the ROESY should be observed. More interactions should give better insight to the structure and evidence that the peptide is cyclic. Better NMR results will provide a better model to compare to the IRMPD results that were obtained by the McMahon group for this cyclic peptide. Computational calculations to determine the most favoured conformation of FK-13 based on the spectroscopy results can be performed. All of these methods combined will provide greater insight to the structure of the peptide and the results can be compared to one another for greater confidence in predicting peptide structure.

4.0 Concluding Remarks

4.0 Concluding Remarks

The analysis of resin-bound GSH and FK-13 was accomplished using a variety of techniques. GSH was a small tripeptide that was synthesized with methods used by this group in the past. The peptide was analyzed using HRMAS NMR spectroscopy and key through-space interactions between the E1 α proton and two C2 β protons suggested that the peptide had adopted a semi-folded conformation. The proposed semi-folded structure was confirmed by the use of DFT calculations on the resin-bound peptide. Molecular dynamics simulations were conducted, but did not agree with the NMR or DFT calculations. The IRMPD results on the protonated GSH molecule suggested that the extra proton added gave rise to intramolecular hydrogen bonds that causes the peptide to have a folded conformation. A mixture of isomers for the protonated species is possible as the lowest energy structure was not the one that possessed the best matching spectrum, but the other isomer presented shared many similarities to the experimental spectrum and had a low calculated energy. From these studies, it is evident that GSH is a very flexible molecule and can adopt different conformations depending on the situations it is placed in and how it is analyzed.

FK-13 was a 13 residue peptide that was synthesized differently from the GSH peptide. It involved more modern techniques that allowed for the synthesis to be completed with effectiveness. Replacing the Wang resin with an HMBA-AM resin and HBTU/HOBt with HCTU made the synthesis easier and go to completion. Analysis on the resin-bound peptide was done using HRMAS NMR spectroscopy. Analysis of the 13 residue peptide was challenging since the TOCSY spectrum contained many overlapping peaks in the α region, but the amide region had fairly good resolution that assisted in the initial assignment of the peptide. Two key interactions in the ROESY spectrum between the residues D10/K2 and R13/I8, suggested that

the peptide was synthesized correctly and had adopted a cyclic conformation. The work that was presented for this peptide has plenty of room for growth, but all of this work done would be useful in future analysis. This work has suggested that manual synthesis of long peptide can be accomplished with great effectiveness using the proper procedures and materials. Also, analysis of this peptide by NMR spectroscopy shows that even with low sensitivity there are interactions that provide insight into the structure of the peptide and should be seen if this cyclic peptide were to be examined again. Overall, the study on FK-13 was much more difficult than GSH due to the size of the peptide, which affected the synthesis and analysis of the NMR spectra, but it is possible to achieve good results if some modifications are made to these methods.

Possible future work on both of these peptides can be done to improve the results for both systems. Solid-phase peptide synthesis of the two peptides can be improved so that a higher substitution can be achieved. The use of the HMBA-AM resin combined with the "standard 20" amino acids work well in tandem. The HMBA-AM resin is a base-labile resin, meaning that strong basic conditions are required to cleave the synthesized peptide from the resin.[21] The "standard 20" amino acids all have protecting groups that are acid labile, i.e. the protecting groups all require strong acidic conditions to be cleaved.[12] Synthesis of peptides should consider the use of these materials throughout the synthesis to minimize the number of different cleavage conditions for protecting groups. The use of HCTU as an aminium salt for the coupling reactions is an exceptional substitute for HBTU and HOBt. Using just HCTU will reduce the reaction time per coupling to 2 hours or less, and is a cleaner method since the reagents can be weighed directly into the Merrifield vessel. Also using different starting reagents such as MSNT or DMAP will increase the first amino acid coupling since this reagent will not cap the reaction sites as it couples and a higher substitution can be achieved.[12] If a

higher starting substitution is achieved, the final substitution can be higher if there are no poor couplings throughout the synthesis. Manual synthesis is not necessarily the best approach to peptide synthesis; other methods such as automated synthesis[21], chemical ligation[69], and microwave assistance[70] are approaches that improve the molecular substitution of the peptide. In the case of the FK-13 synthesis, a higher final substitution of the peptide would result in better sensitivity in the ROESY spectrum. The sample would be more concentrated and more interactions would be seen in the spectrum to provide more support for the proposition of a cyclic structure.

The continuation of the FK-13 work involves analysis of the peptide using NMR spectroscopy and then interpreting at the IRMPD results obtained by Dr. McMahon's group as well. The analysis of the NMR spectra will provide insight into the structure of the resin-bound peptide and can be compared to the results obtained in this thesis. Many of the interactions that were seen in this thesis should still be observed in future work along with additional interactions, namely in the α region, that would be useful in supporting the proposition of the cyclic structure. Analysis of the IRMPD structures needs to be done to evaluate the structure of the protonated peptide. Generating the protonated structures using DFT calculations would be the first step and then comparing the spectra, calculated structure and energies of the species would need to be done. The methods and steps for analyzing the IRMPD results for FK-13 can be done in the same manner as GSH, i.e. choosing candidate structures, then narrowing down the selection based on similarities in spectra, calculated energy and structure.

Additional work for the structure calculations on GSH and FK-13 can be further improved. The GSH results were very good and agreed with the NMR results, but the semi-folded conformation is not the one that exists for GSH. The exploration of the different

structures that exists for GSH can be accomplished using a method known as "basin-hopping".[71] This technique can be used to explore the potential energy surface of a molecule and determine the structures that occupy the different energy minima. This technique uses an unbiased algorithm to randomly sample different conformations of the peptide, but also uses lower level calculations to complete the analysis.[71] After obtaining the structures from the low level calculations, the structures that occur the most are indicative of the structures that occupy the different energy minima of the potential surface and can be further optimized using DFT calculations. This will expand upon the analysis of the GSH conformation by not only finding the most favourable conformation, but also by looking for other structures that are energetically favourable. Analysis of the potential energy surface of the resin-bound GSH and FK-13 peptide can be accomplished using this technique to provide more insight into what structures these peptides prefer.

Analysis of the DFT structures for GSH was accomplished using the PCM model of solvation. While this technique gives a good interpretation of how the resin-bound peptide behaves while in a solution of DMF, it is not the most accurate way to interpret the structure. The PCM model of solvation is an implicit model of solvation and puts a "dielectric box" of solvent around the molecule to simulate the solvent interactions. This is not entirely accurate because it assumes the small scale interactions are the same as the large scale interactions.[72] Using the explicit model of solvation would improve the analysis of the resin-bound peptide structures because this method introduces actual solvent molecules into the system. This allows for the solvent molecules to interact with the peptide being calculated as opposed to simulating the effects of the solvent.[72] Even though the GSH peptide showed a semi-folded conformation in the DFT calculation performed, the results could be more accurate by using the explicit

model of solvation. For future analysis of the FK-13 peptide, the cyclic structure on its own should be calculated first, followed by analysis of the resin-bound peptide. After this the introduction of actual solvent molecules into the system can be done to fully analyze the effect that solvent molecules would have on the resin-bound cyclic structure of FK-13.

There are a lot of opportunities to further improve upon the results obtained in the GSH work. The sample could be resynthesized if necessary to achieve a higher substitution and subsequently could improve upon the NMR results. The main area of improvement for the GSH study would be analysis of the different conformations that GSH could adopt with the use of basin-hopping. Also, obtaining a more accurate structure for the DFT structure could be accomplished using the explicit model of solvation. The work done for FK-13 in this thesis can be used a good starting point for future work. The methods outlined to improve the solid-phase synthesis of this peptide can be used to achieve a higher final substitution. Having a higher substitution on the resin will help improve the sensitivity of the ROESY spectrum, which will give rise to interactions in the α region so that more evidence for the cyclic structure can be determined. Analysis of the resin-bound FK-13 can also be explored using structure calculations to determine the most favourable conformation of this peptide when interacting with DMF molecules. Finally, the IRMPD structures of FK-13 can be determined from the experimental spectrum obtained by Dr. McMahon's group using the same procedure as the GSH IRMPD analysis. The work presented in this thesis has given a lot of valuable information about the structure of GSH and FK-13 and is an excellent starting point for future work in the analysis of either of these peptides.

References

1. Wade, L. G., (2006). *Organic Chemistry 6th ed.*, Upper Saddle River, New Jersey: Pearson Prentice Hall
2. Pauling, L. and Corey, R. B., *Proceedings of the National Academy of Sciences*, **1951**, *37*, 729-740
3. Voet D. and Voet, J., (2004). *Biochemistry 3rd ed.*, New Jersey: John Wiley and Sons Inc.
4. Harris, D. C., (2007). *Quantitative Chemical Analysis 7th ed.*, New York: W.H. Freeman and Company
5. Silverstein, R. M., Webster, F. X., and Kiemle, D. J., (2007). *Spectrometric Identification of Organic Compounds 7th ed.*, Hoboken, New Jersey: John Wiley and Sons Inc.
6. Polfer, N. C, *Chemical Society Reviews*, **2011**, *40*, 2211-2221
7. Roberts, G. C. K., (1993). *NMR of Macromolecules: A Practical Approach*, New York: IRL Press
8. Merrifield, R. B., *Journal of the American Chemical Society*, **1963**, *85*, 2149-2154
9. Sheehan, J. C., *Journal of the American Chemical Society*, **1955**, *77*, 1067-1068
10. Merrifield, R. B., *Journal of the American Chemical Society*, **1964**, *86*, 304-305
11. Carpino, L. A., and Han, G. Y., *Journal of Organic Chemistry*, **1972**, *37*, 3404-3409
12. Kates, S. A. and Albericio, F. (2000). *Solid-Phase Synthesis A Practical Guide*, New York, New York: Marcel Dekker Inc.
13. Chan, W. C. and White, P. D. (2000). *Fmoc Solid Phase Peptide Synthesis A Practical Approach*, New York, New York: Oxford University Press

14. Fridkin, M., Patchornik, A., and Katchalski, E., *Journal of the American Chemical Society*, **1965**, *87*, 4646-4648

15. Lebl, M. and Hruby, V. J., *Tetrahedron Letters*, **1984**, *25*, 2067-2068

16. Schiller, P. W., Nguyen, T., and Miller, J., *International Journal of Peptide and Protein Research*, **1985**, *25*, 171-177

17. Sklyarov, L. Y. and Shashkova, I. V., *Academy of Sciences of the USSR*, **1969**, *39*, 2778-2783

18. Isied, S. S., Kuehn, C. G., and Lyon, J. M., *Journal of the American Chemical Society*, **1970**, *104*, 2632-2634

19. Carpino, L. A., Beyermann, M., Wenschuh, H., and Bienert, M., *Accounts of Chemical Research*, **1996**, *29*, 268-274

20. Atherton, E., Holder, J. L., Meldal, M., Sheppard, R. C., and Valerio, R. M., *Journal of the Chemical Society, Perkin Transactions 1*, **1988**, *10*, 2887-2894

21. Artherton, E. and Sheppard, R. C. (1989) *Solid Phase Peptide Synthesis A Practical Approach*, Oxford, England: Oxford University Press

22. Siber, P., *Tetrahedron Letters*, **1987**, *28*, 6147-6150

23. Blankemeyer-Menge, B., Nimtz, M. and Frank, R., *Tetrahedron Letters*, **1990**, *31*, 1701-1704

24. Knorr, R., Trzeciak, A., Bannwarth, W., and Gillessen, D., *Tetrahedron Letters*, **1989**, *30*, 1927-1930

25. Carpino, L. A., *Journal of the American Chemical Society*, **1993**, *115*, 4397-4398

26. Kaiser, E., Colescott, R. L., Bossinger, C. D., and Cook P. I., *Analytical Biochemistry*, **1970**, *34*, 595-598

27. Krchňák, V., Vágner, J, and Lebl, M., *International Journal of Peptide and Protein Research*, **1988**, *32*, 415-416

28. United Nations (2008). *Amendment to the Proper Shipping Name of UN 3474 for inclusion of 1-HOBt Monohydrate Transmitted by the International Council of Chemical Associations (ICCA)*, New York

29. Levitt, M. H. (2002). *Spun Dynamics Basics of Nuclear Magnetic Resonance*, New York, New York: John Wiley and Sons, Ltd.

30. Andrew, E. R. (2010) Magic Angle Spinning. In A. E. McDermott and T. Polenova (Eds.), *Solid-State NMR Studies of Biopolymers* (pp. 83-96). United Kingdom: Jon Wiley and Sons Ltd.

31. Hiller, S., Wasmer, C., Wider, G., and Wüthrich, K., *Journal of the American Chemical Society*, **2007**, *129*, 10823-10828

32. Braun, S., Kalinowski, H.-O., and Berger, S. (1998) *150 and More Basic NMR Experiments*, Weinheim, Germany: Wiley-VCH

33. Njålsson, R. and Norgren, S., *Acta Paediatrica*, **2005**, *94*, 132-137 (originally 7)

34. Hayes, J. D. and McLellan, L. I., *Free Radical Research.*, **1999**, *31*, 273-300

35. Pompella, A., Visvikis, A., Paolicchi, A., De Tata, V., and Casini, A. F., *Biochemical Pharmacology*, **2003**, *66*, 1499-1503

36. Coles, B., F. and Kadlubar, F. F., *Biofactors*, **2003**, *17*, 115-130

37. Fujiwara, S., Formicka-Kozlowska, G., and Kozlowski, H., *Bullet of the Chemical Society of Japan*, **1977**, *50*, 3131-3135

38. York, M. J., Beilharz, G. R., and Kuchel, P. W., *International Journal of Peptide and Protein Research*, **1987**, *29*, 638-646

39. Fuhr, B. J. and Rabenstein, D. L., *Journal of the American Chemical Society*, **1973**, *95*, 6944-6950

40. Rabenstein, D. L., *Journal of the American Chemical Society*, **1973**, *95*, 2797-2803

41. Sardo, M., Sigel, R., Santos, S. M., Rocha, J., Gomes, J. R. B., and Mafra, L., *Journal of Physical Chemistry A*, **2012**, *116*, 6711-6719

42. Lampela, O, Juffer, A. H., and Rauk, A., *Journal of Physical Chemistry A*, **2003**, *107*, 920-9220

43. Zhang, R. and Wu, W., *Journal of Molecular Liquids*, **2011**, *162*, 20-25

44. Vila-Viçosa, Teixeria, V. H., Santos, H. A. F., and Machuqueiro, M., *Journal of Physical Chemistry B*, **2012**, *117*, 7507-7517

45. Dai, Z., Chu, Y., Wu, L., and Ding, C., *Acta Pharmacol Sin*, **2008**, *29*, 759-771

46. Zhao, J., Siu, K.. W. M., and Hopkinson, A. C., *Organic and Biomolecular Chemistry*, **2011**, *9*, 7384-7392

47. Osburn, S., Berden, G., Oomens, J., O'Hair, R. A. J., and Ryzhov, V., *Journal of the American Society for Mass Spectrometry*, **2012**, *23*, 1019-1023

48. Osburn, S., Berden, G., Oomens, J., Gulyuz, K., Polfer, N. C., O'Hair, A. J., and Ryzhov, V., *Chem Plus Chem*, **2013**, *78*, 970-978

49. Scuderi, D., Bakker, J. M., Durand, S., Maitre, P., Sharma, A., Martens, J. K., Nicol, E., Clavaguéra, C., and Ohanessian, G., *International Journal of Mass Spectrometry*, **2011**, *308*, 338-347

50. Dick, F. (1998) *Peptides 1996, Proc. Of 24th European Peptide Symposium*, Birmingham, Mayflower Scientific Ltd.

51. Wang, G., *Journal of Biological Chemistry*, **2008**, *283*, 32637-32643

52. Johansson, J., Gudmundsson, G. H., Rottenberg, M. E., Berndt, K. D., and Agerberth, B., *Journal of Biological Chemistry*, **1998**, *273*, 3718-3724

53. Agerberth, B., Gunne, H., Odeberg, J., Kogner, P., Boman, H. G., and Gudmundsson, G. H., *Proceedings of National Academy of Sciences*, **1995**, *92*, 195-199

54. Oren, Z., Lerman, J. C., Gudmundsson, G. H., Agerberth, B., and Shai, Y., *Biochemistry Journal*, **1999**, *341*, 501-513

55. Wildman, K. A. H., Lee, D., and Ramamoorthy, A., *Biochemistry*, **2002**, *42*, 6545-6558

56. Bals, R., Weiner, D. J., Meegalla, R. L., and Wilson, J. M., *Journal of Clinical Investigation*, **1999**, *103*, 1113-1117

57. Cirioni, O., Giacometti, A., Ghiselli, R., Bergnach, C., Orlando, F., Silvestri, C., Mocchegiani, F., Licci, A., Skerlavaj, B., Rocchi, M., Saba, V., Zanetti, M., and Scalise, G., *Antimicrobial Agents and Chemotherapy*, **2006**, *50*, 1672-1679

58. Wang, G., Watson, K. M., and Buckheit, R. W., *Antimicrobial Agents and Chemotherapy*, **2008**, *52*, 3438-3440

59. Li, X., Li, Y., Han, H., Miller, D. W., and Wang, G., *Journal of the American Chemical Society*, **2006**, *128*, 5776-5785

60. Chen, C., Brock, R., Luh, F., Chou, P., Larrick, J. W., Huang, R. and Huang, T., *FEBS Letters*, **1995**, *370*, 46-52

61. Wildman, K. A. H., Martinez, G. V., Brown, M. F., and Ramamoorthy, A., *Biochemistry*, **2004**, *43*, 8459-8469

62. Li, Y., Li, X., and Wang, G., *Protein Expression and Purification*, **2006**, *47*, 498-505

63. Porcelli, F., Verardi, R., Shi, L., Wildman, K. A. W., Ramamoorthy, A., and Veglia, G., *Biochemistry*, **2008**, *47*, 5565-5572

64. Wang, G., *Biochimica et Biophysica Acta*, **2010**, *1798*, 114-121

65. Wu, W. K. K., Wang, G., Coffelt, S. B., Betancourt, A. M., Lee, C. W., Fan, D., Wu, K., Yu, J., Sung, J. J. Y., and Cho, C. H., *International Journal of Cancer*, **2010**, *127*, 1741-1747

66. Gudmundsson, G. H., Agerberth, B., Odeberg, J., Bergman, T., Olsson, B., and Salcedo, R., *European Journal of Biochemistry*, **1996**, *238*, 325-332

67. Wüthrich, K., Billeter, M., and Braun, W., *Journal of Molecular Biology*, **1984**, *180*, 715-740

68. Kent, S. B. H. (1985) *Peptides, structure and function; Proceedings of the 9th American Peptide Symposium*, Rockford, IL, Pierce Chemical Co.

69. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. H., *Science*, **1994**, *266*, 776-779

70. Palasek, S. A., Cox, Z. J., and Collins, J. M., *Journal of Peptide Science*, **2006**, *13*, 143-148

71. Wales, D. J. and Doye, J. P. K., *Journal of Physical Chemistry A*, **1997**, *101*, 5111-5116

72. Lee, M. S., Salsbury, F. R., and Olson, M. A., *Journal of Computational Chemistry*, **2004**, *25*, 1967-1978

Appendix

Table A: Custom AMBER force field used to generate a wide variety of structures to examine the best conformation for the GSH peptide. The 'Atom Type' is the AMBER atom type needed to run the molecular mechanics calculations. The 'Partial Charge' is the result from the ChelpG calculation. The X, Y and Z columns are the coordinates of the first GSH peptide ran using the AMBER force field.

Atom Type	Partial Charge	X	Y	Z
C-C	0.721629	-6.72142441	-0.05211683	-1.95029936
O-O	-0.622561	-7.38574698	-0.44712557	-1.01027255
O-OH	-0.704190	-7.26912658	0.50191347	-3.05077773
H-HO	0.488775	-8.23487438	0.54009544	-2.93047890
C-CT	0.486977	-5.20042778	-0.06967449	-2.01437781
H-H1	0.026662	-4.91878404	-0.25417687	-3.05699220
N-N3	-1.196722	-4.71491297	-1.18957114	-1.20951521
H-H	0.408767	-4.95537244	-1.05511750	-0.22968409
H-H	0.451086	-3.70336480	-1.26007140	-1.27066936
C-CT	0.018362	-4.68883013	1.34200174	-1.61496087
H-HC	-0.020618	-5.18602589	2.09454306	-2.23447356
H-HC	0.006749	-4.96802951	1.53696546	-0.57186950
C-CT	-0.177405	-3.17061949	1.49041312	-1.78457440
H-HC	0.033001	-2.63267874	0.79909288	-1.12883654
H-HC	0.036978	-2.88787554	1.24385494	-2.81682479
C-C	0.798043	-2.70547605	2.91980314	-1.53927134
O-O	-0.668540	-3.28685575	3.89452775	-2.02554036

N-N	-0.789350	-1.59730628	3.05477692	-0.75091548
H-H	0.412194	-1.12308594	2.21114378	-0.45448301
C-CT	0.259399	-1.59908380	3.82362692	0.48322304
H-H1	0.062041	-2.49636298	4.44955873	0.47073007
C-CT	-0.058171	-0.35072268	4.71793752	0.60130758
H-H1	0.042927	-0.33466805	5.17579430	1.58993069
H-H1	0.140433	0.56061401	4.12120124	0.49502615
S-SH	-0.359477	-0.21831885	6.02989092	-0.68112449
H-HS	0.220776	-1.24218437	6.79161369	-0.24549221
C-C	0.659071	-1.72550840	2.96726998	1.76378390
O-O	-0.670089	-1.80526488	3.51895177	2.86873760
N-N	-0.592156	-0.68637788	2.19153334	1.40261817
H-H	0.311719	-0.67601874	1.74575378	0.49282556
C-CT	0.173864	0.42016720	1.92583534	2.29626698
H-H1	0.069658	0.08048733	1.44162987	3.21964576
H-H1	0.060170	0.92376379	2.85301238	2.59350164
C-C	0.758408	1.42280626	1.02216700	1.61327728
O-O	-0.602581	1.30394437	0.58342266	0.48462659
O-OH	-0.671462	2.46726004	0.76094923	2.41239189
H-HO	0.485632	3.08717956	0.17620172	1.94091820

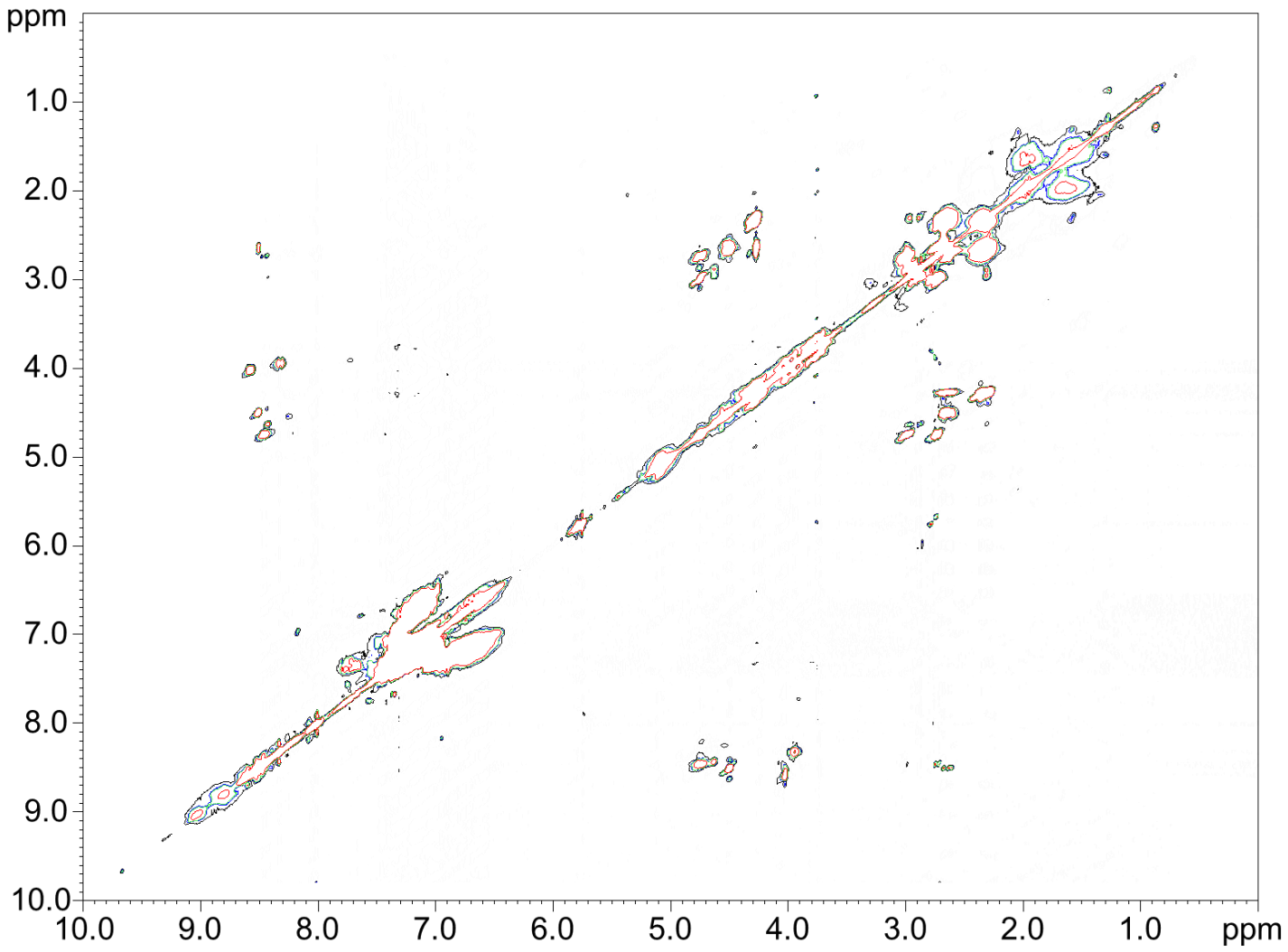


Figure A: Full TOCSY spectrum of the resin-bound GSH peptide. This experiment detects the interactions between ^{1}H within the same spin-system. Three spin systems and their respective R groups were assigned using this spectrum.

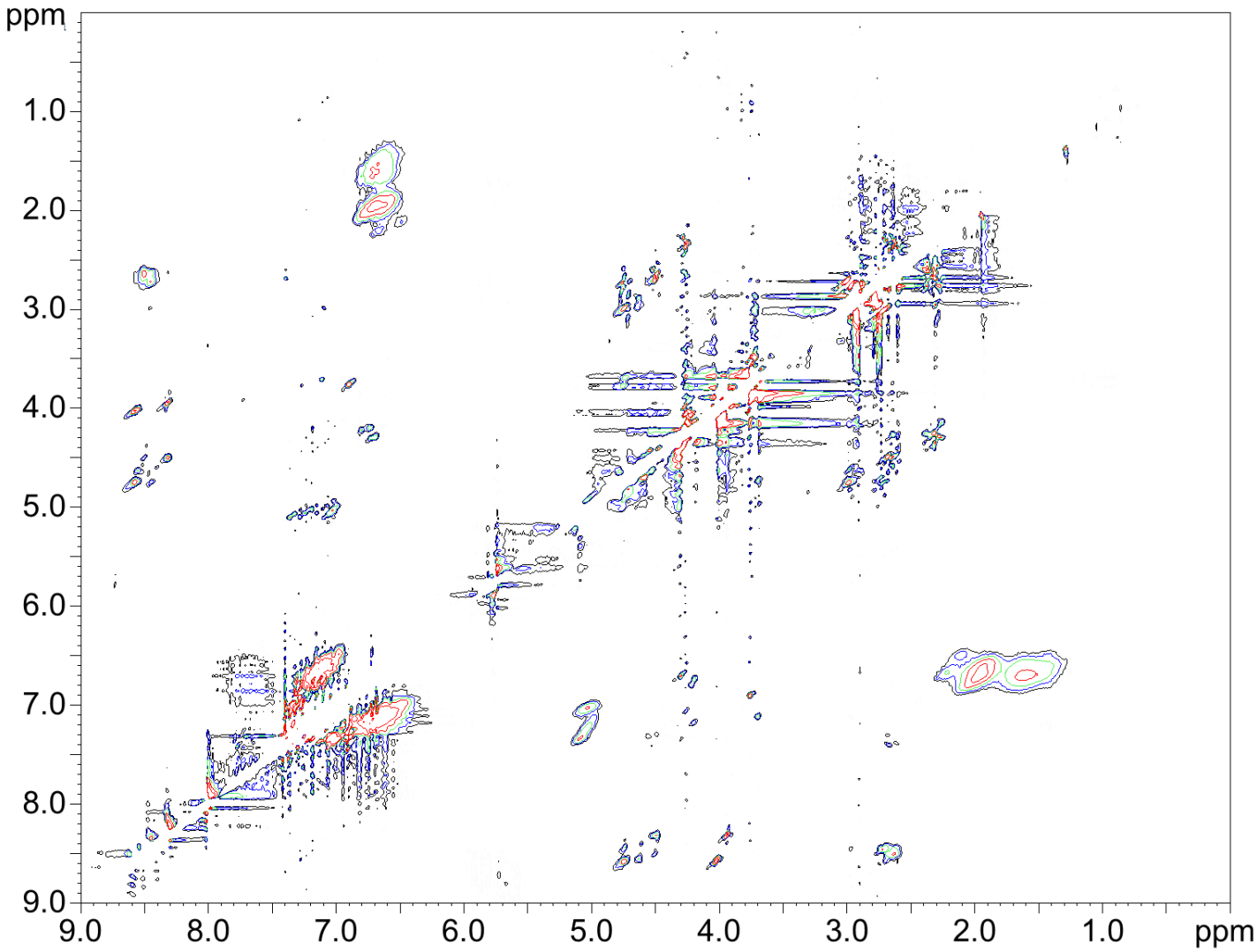


Figure B: Full ROESY spectrum of the resin-bound GSH peptide. This experiment detects the through-space interactions between ^{1}H atoms. The key interactions from this experiment will allow for the three-dimensional structure of the peptide to be determined

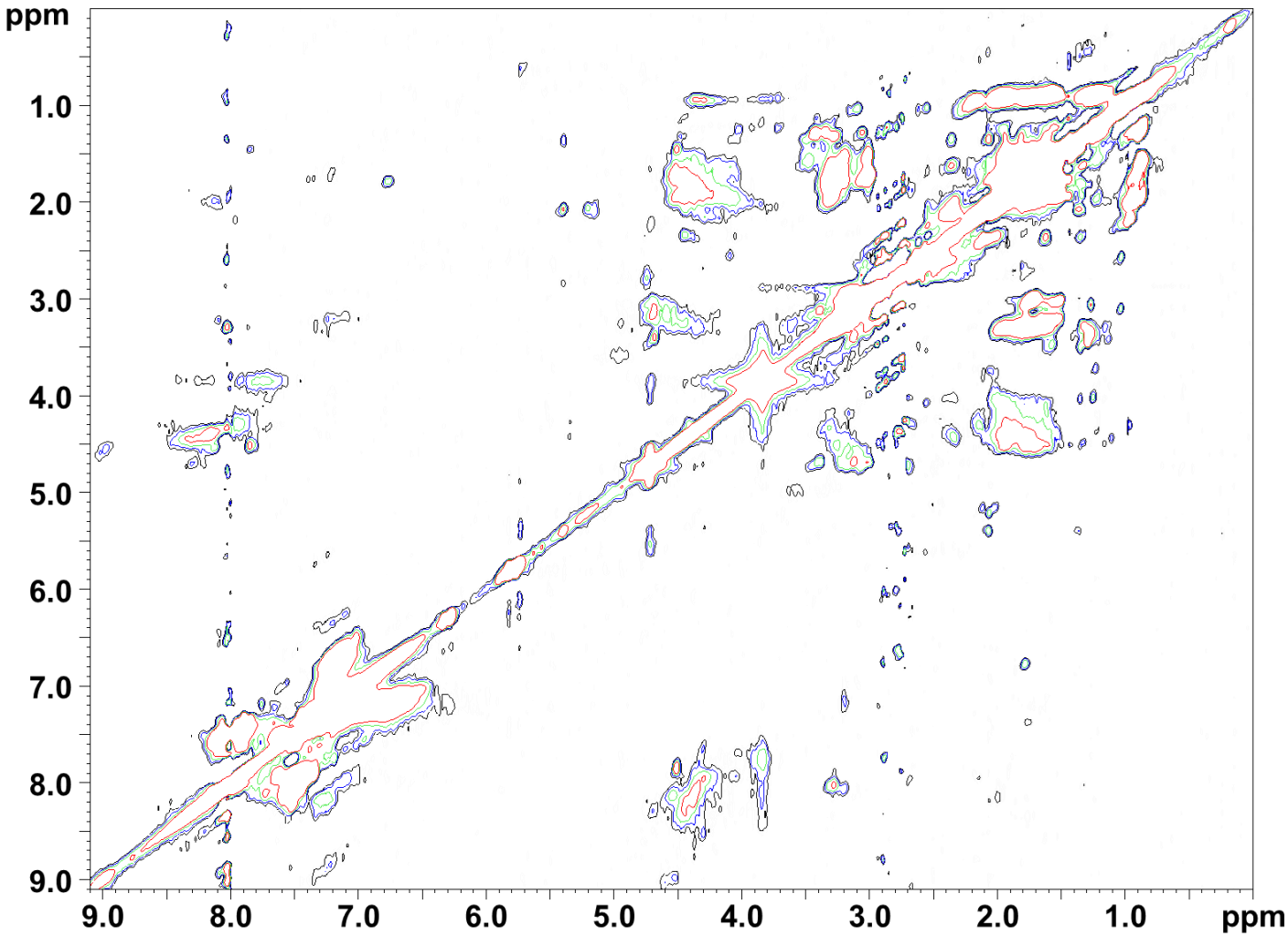


Figure C: Full TOCSY spectrum of the resin-bound FK-13 peptide. The thirteen different spin systems and their respective R groups were assigned ^{1}H shifts using this spectrum.

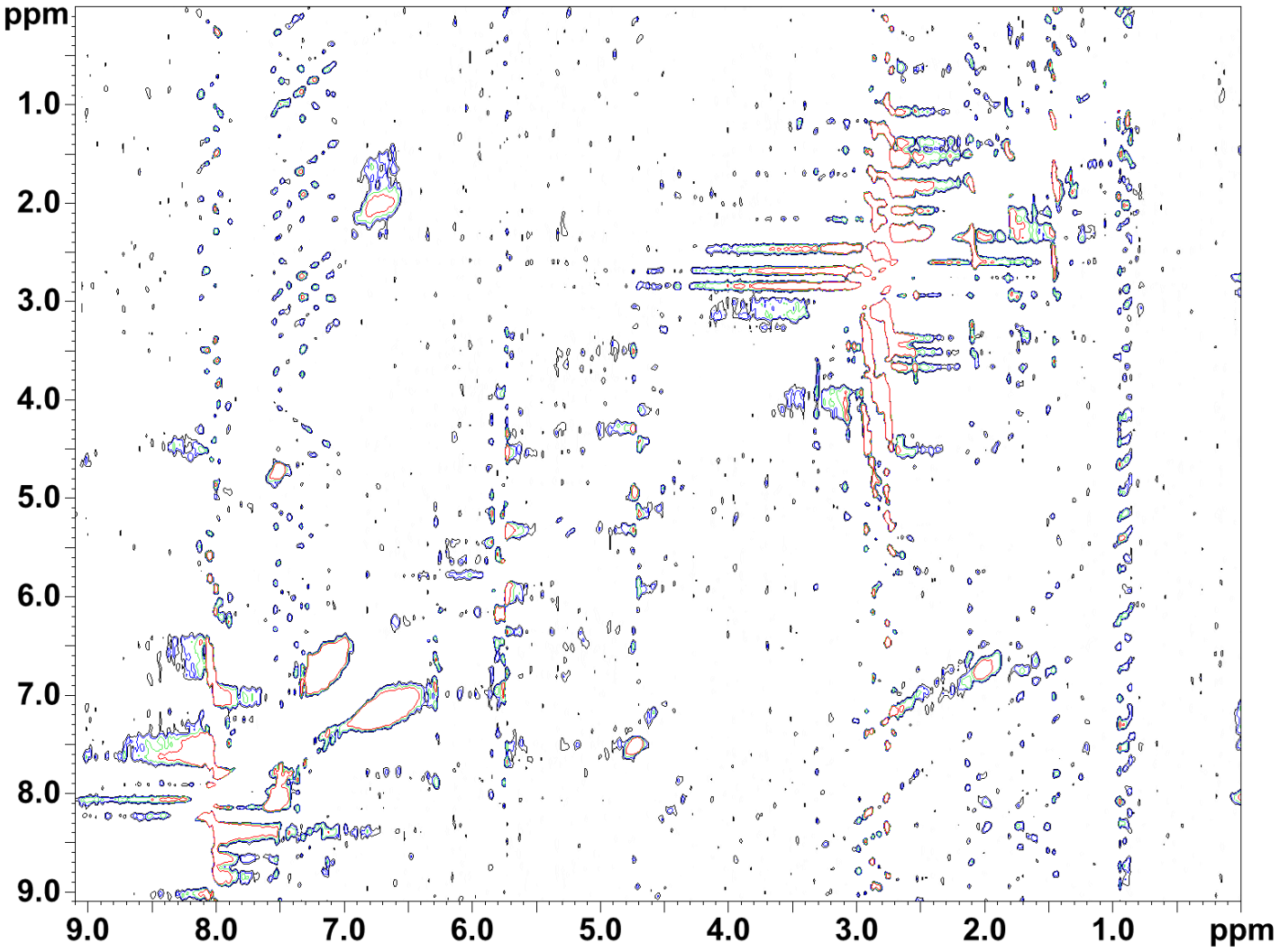


Figure D: Full ROESY spectrum of the resin-bound FK-13 peptide. The through-space interactions were detected between the ^{1}H atoms of the peptide. Using the key interactions from this spectrum the three-dimensional structure of the peptide can be determined.

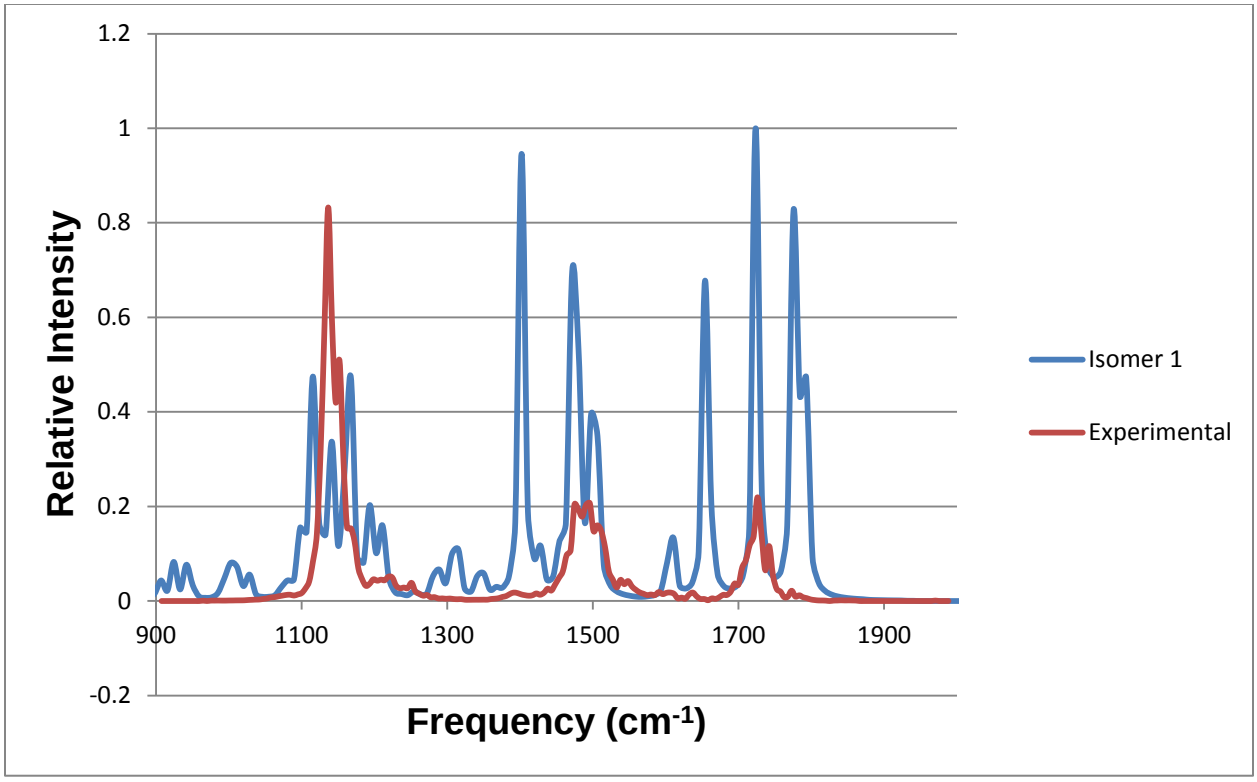


Figure E: Superimposed IRMPD spectra comparing the experimental GSH (red line) to the calculated spectrum for isomer 1. The spectra are both reported as frequency versus IRMPD efficiency (relative intensity). The three areas of interest in these spectra are the carbonyl region (~1900 cm^{-1}), amide region (~1600 cm^{-1}) and low IR region (~1100 cm^{-1}).

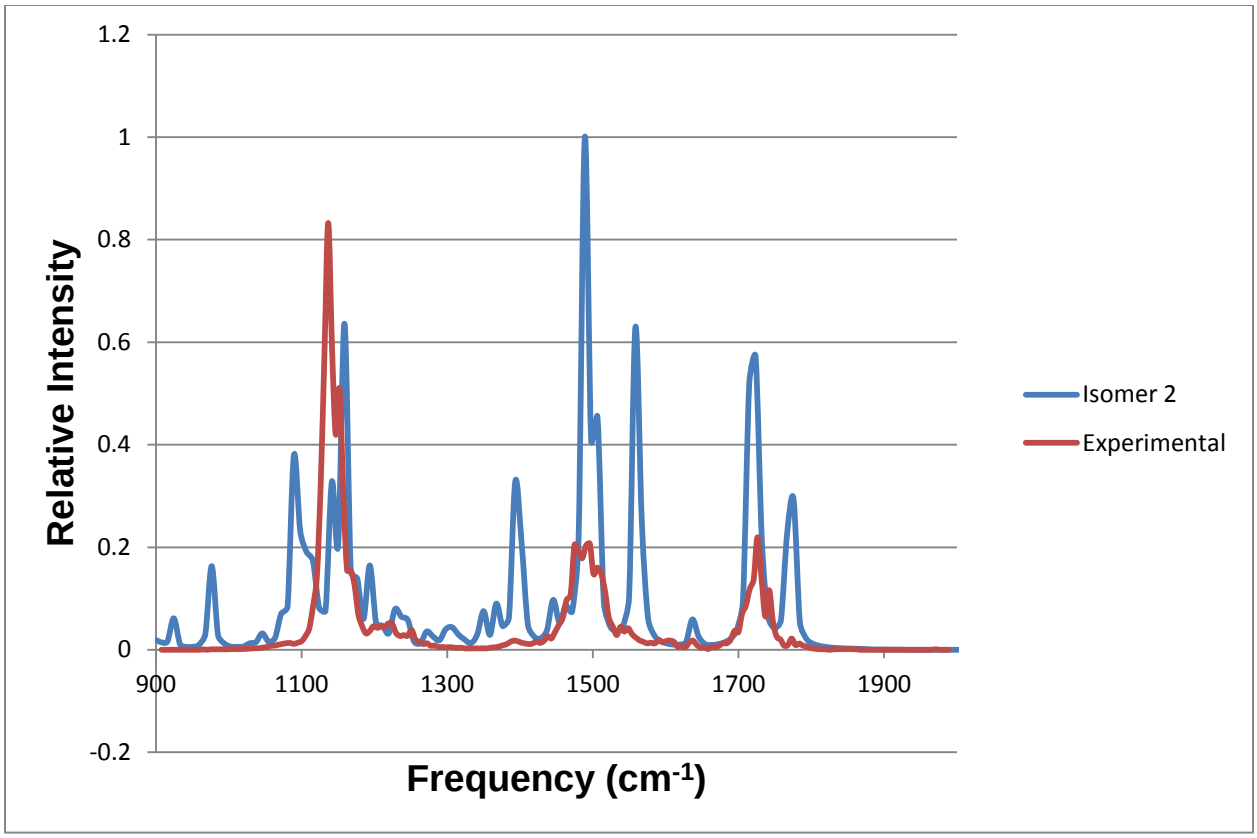


Figure F: Superimposed IRMPD spectra comparing the experimental GSH (red line) to the calculated spectrum for isomer 2. The spectra are both reported as frequency versus IRMPD efficiency (relative intensity). The three areas of interest in these spectra are the carbonyl region (~1900 cm^{-1}), amide region (~1600 cm^{-1}) and low IR region (~1100 cm^{-1}).

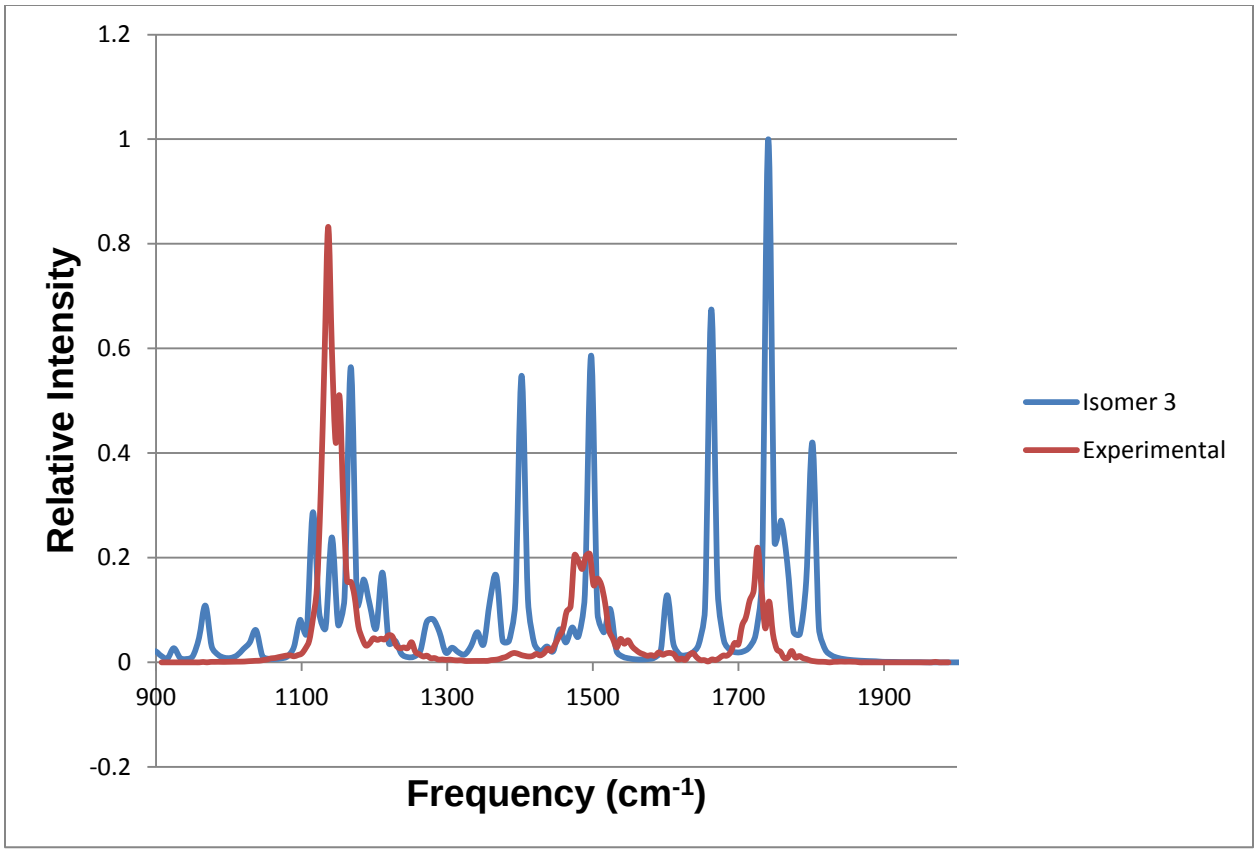


Figure G: Superimposed IRMPD spectra comparing the experimental GSH (red line) to the calculated spectrum for isomer 3. The spectra are both reported as frequency versus IRMPD efficiency (relative intensity). The three areas of interest in these spectra are the carbonyl region (~1900 cm^{-1}), amide region (~1600 cm^{-1}) and low IR region (~1100 cm^{-1}).

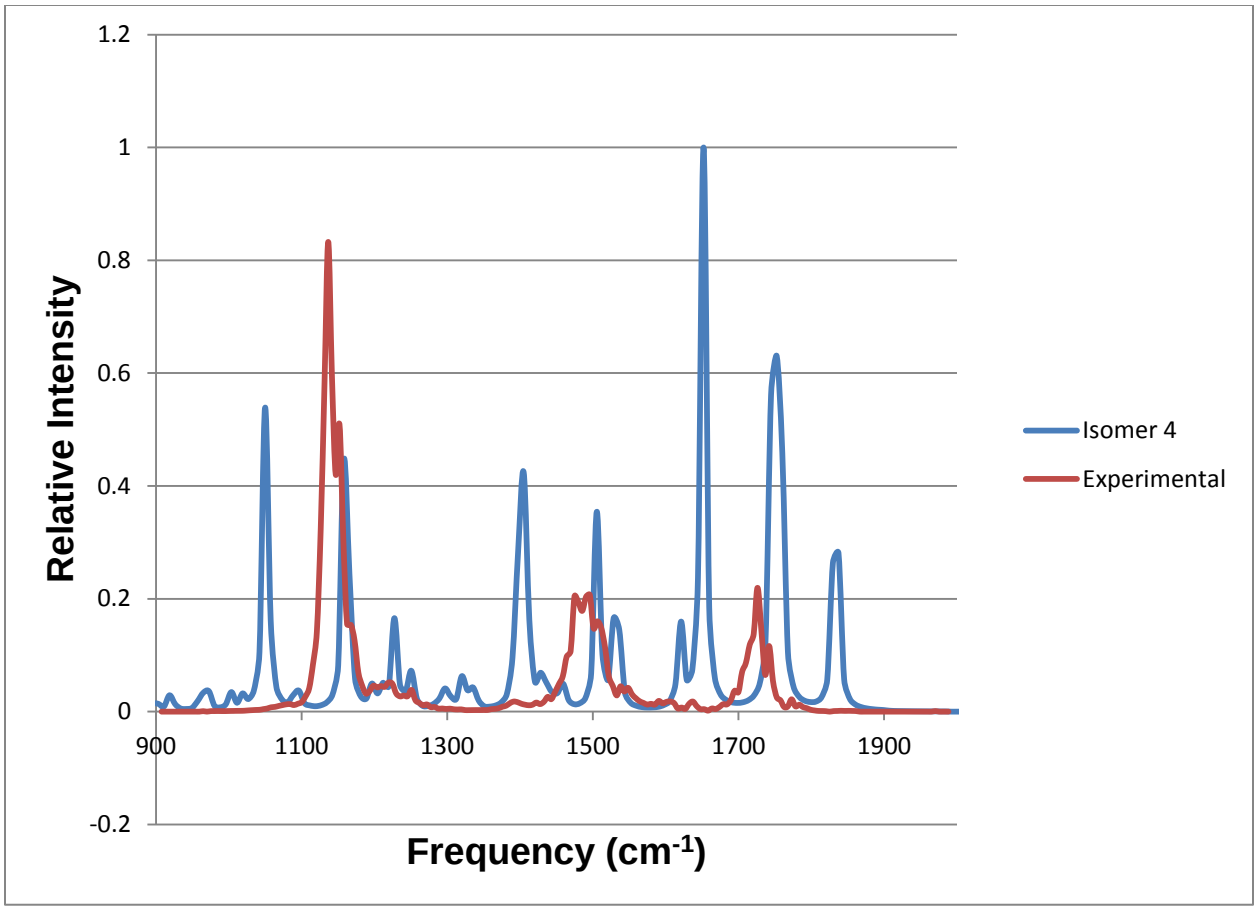


Figure H: Superimposed IRMPD spectra comparing the experimental GSH (red line) to the calculated spectrum for isomer 4. The spectra are both reported as frequency versus IRMPD efficiency (relative intensity). The three areas of interest in these spectra are the carbonyl region (~1900 cm^{-1}), amide region (~1600 cm^{-1}) and low IR region (~1100 cm^{-1}).

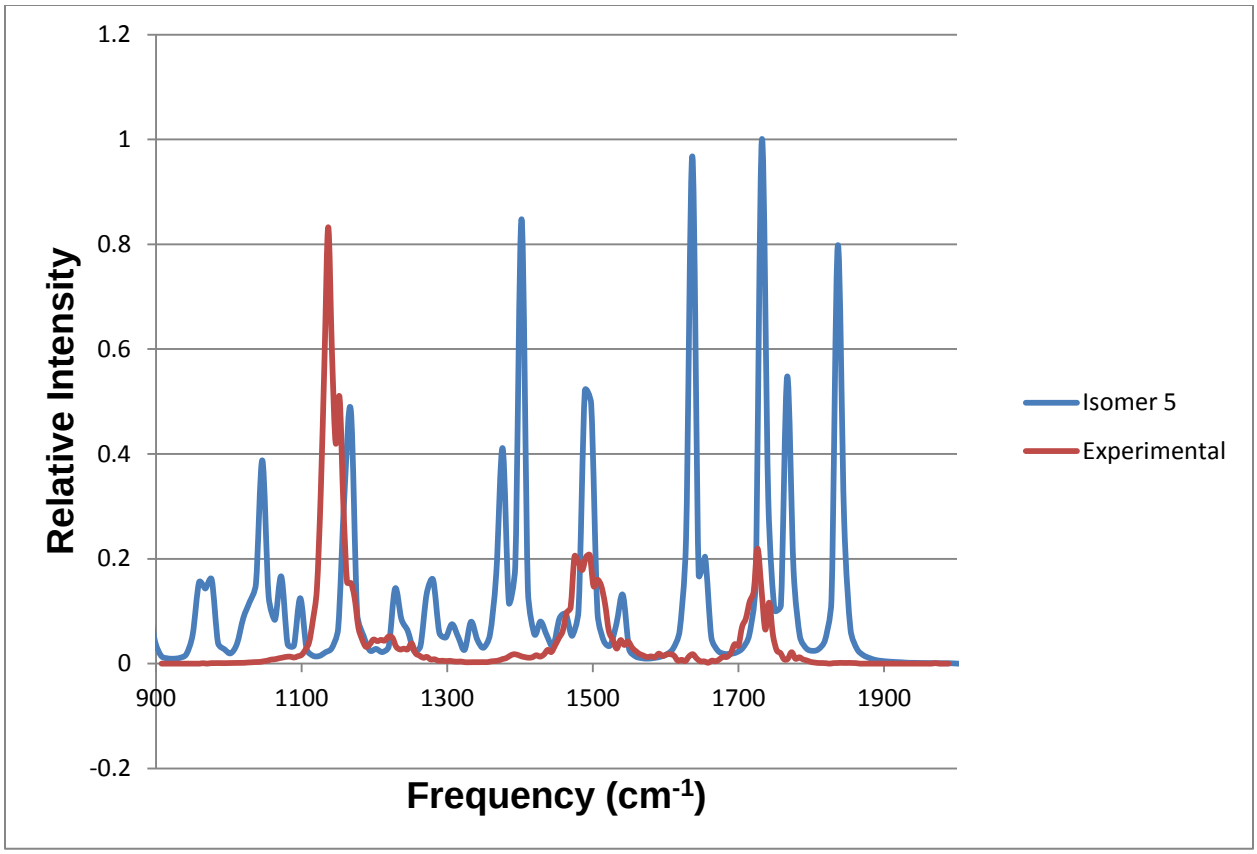


Figure I: Superimposed IRMPD spectra comparing the experimental GSH (red line) to the calculated spectrum for isomer 5. The spectra are both reported as frequency versus IRMPD efficiency (relative intensity). The three areas of interest in these spectra are the carbonyl region (~1900 cm^{-1}), amide region (~1600 cm^{-1}) and low IR region (~1100 cm^{-1}).