6-1997

Design, synthesis, and characterization of short salt and non-salt bridged peptides

Clark K.N. Choi
Union College - Schenectady, NY

DESIGN, SYNTHESIS, AND CHARACTERIZATION OF SHORT SALT AND NON-SALT BRIDGED PEPTIDES

By

Clark K.N. Choi

****

Submitted in partial fulfillment

of the requirements for

Honors in the Department of Chemistry

UNION COLLEGE

June, 1997

ABSTRACT

CHOI, CLARK Design, Synthesis, and Characterization of Short Salt and Non-Salt Bridged Peptides. Department of Chemistry, June 1997.

Peptides with secondary characteristics such as α-helices are among the most important structures in nature. Previous studies from Baldwin and coworkers have shown that short alanine-based peptides with specific salt spacing adopt helical conformations. In the study of helix formation, we have synthesized our versions of alanine- and glutamine-based peptides with $(i, i + 4)$ and $(i, i + 5)$ salt spacing. The sequence of our QEKR 4.0[b] and QEKR 5.0 peptides is as follows:

QEKR 4.0[b] = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-E_{13}-A_{12}-Q_{11}-A_{10}-K_9-A_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

QEKR 5.0 = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-E_{13}-A_{12}-Q_{11}-A_{10}-A_9-K_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

The design of the salt bridge is most suitable when it contains $(i, i + 4)$ salt spacing. The rational is that the $(i, i + 4)$ salt spacing yields a value close to 3.6, the number of amino acids it takes to stabilize one turn of an α-helix, than the $(i, i + 5)$ salt spacing. Using molecular modeling, we have shown that peptides with $(i, i + 4)$ salt spacing indeed formed stabilized salt bridge as opposed to no salt bridge for the $(i, i + 5)$ peptide.

The two main strategies for synthesis of the peptides utilize Fmoc as N-α main chain protection and Boc as side chain protection. Both groups display chemical orthogonality in acidic and basic conditions. The yields of the isolated di-TFA salts of both peptides were about 40%. MALDI-MS confirmed that both peptides were synthesized with 65% purity. Using reverse phase HPLC, we purified both $(i, i + 4)$ and $(i, i + 5)$ salt spaced peptides. NMR spectra of both peptides were similar.

In future studies, we would like to take CD measurements of our peptides to determine their degree of helicity. In addition, we would like to use cyanogen, a known reagent that traps salt bridges in proteins, and other coupling reagents to react with the salt spacing of our synthesized peptides to form amide links.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my sincere gratitude for all the countless help, effort, advice, and time my mentor, Professor Leslie A. Hull, has invested with me in the completion of my senior thesis. Professor Hull has made me more appreciative in learning different aspects of chemistry as a science which can apply to different areas of research. Second, I would like to thank my advisor and friend, Professor John R. Sowa, for fostering my maturity and furthering my goals throughout my four years at Union. Third, I would like to acknowledge the support of other chemistry faculties at Union who I have the pleasure in meeting and learning in and outside of classes. In addition, I am very grateful of the MALDI spectra provided by New York University for making my senior project possible. Finally, I would like to express my eternal love and affection for my parents who have taught me a lifetime of learning and have given me an opportunity to life, education, and success.

Clark K.N. Choi
June 6, 1997

TABLE OF CONTENTS

Chapter		Page
1	**Introduction**	
	Background	1
	Problem	12
	Peptide Design of QEKR 4.0^b and QEKR 5.0	14
	Instrumental Analysis	17
2	**Experimental**	
	Synthesis of QEKR 4.0^b and QEKR 5.0	20
	Deprotection of Fmoc Main Chain Protecting Groups	22
	Amino Acid Activation	24
	Coupling of Amino Acid to the Peptide Chain	26
	Kaiser Test	26
	N-terminus Capping	28
	Cleavage of Resins and Side Chain Protecting Groups	29
	Molecular Weight Determination	37
	Analysis of Peptides by HPLC	38
	Purification of Peptides by HPLC	39
	Characterization of Peptides by NMR	40
	Molecular Modeling	41
	Titration of QEKR 4.0^b by NMR	42
3	**Results and Discussion**	
	Percent Yields of QEKR 4.0^b and QEKR 5.0	44
	Confirmation of QEKR 4.0^b and QEKR 5.0 by MALDI-MS	44

Determination of Retention Times and Purification of QEKR 4.0^b and QEKR 5.0 by HPLC 49

^{1}H-NMR Spectra of Individual Amino Acids 57

1-D ^{1}H-NMR and 2-D COSY Spectra of QEKR 4.0^b and QEKR 5.0 66

Minimization of QEKR 4.0^b and QEKR 5.0 72

Determination of the pKa of Glutamic Acid in QEKR 4.0^b 78

Future Work 83

References

TABLE OF FIGURES

Figure		Page
1	A typical amino acid at pH 7 is represented as a tetrahedral structure. It consists of an α-carbon, amino and carboxyl groups, and a variable side chain.	1
2	Salt bridge among two opposite charged amino acids can be trapped using cyanogen to form amide links.	13
3	The structure of the Rink Amide resin.	20
4	The diagram of a glass sintered vessel.	23
5	Base sensitive Fmoc protecting group on main chain of amino acid 1 is removed by piperidine.	25
6	Amino acid A (9) is coupled to the peptide chain (6) at its amino end via the BOP, HOBt, and DIEA reagents.	28
7	The amino end of the peptide chain (13) is capped with an acetyl group.	30
8	Apparatus set-up for peptide suction filtration.	31
9	Cleavage of the peptide chain (14) off the Rink Amide linker (15) using TFA.	33
10	TFA deprotection of a *t*-butyl ester group on glutamic acid.	34
11	TFA deprotection of a Boc group on lysine.	35
12	TFA deprotection of a *t*-butyl ether group on tyrosine.	36
13	TFA deprotection of a Pmc group on arginine.	37
14	MALDI-MS of QEKR 4.0[b]: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.	47
15	MALDI-MS of QEKR 5.0: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.	48
16	Isocratic runs (70% H_2O:30% acetonitrile) of QEKR 4.0[b]: A) 254 nm; B) 275 nm; C) 288 nm.	51
17	Isocratic runs (70% H_2O:30% acetonitrile) of QEKR 5.0: A) 254 nm; B) 275 nm; C) 288 nm.	52
18	Solvent programming of QEKR 4.0[b] at 275 nm: A) 1st purification;	53

B) 2nd purification.

19 Solvent programming of QEKR 5.0 at 275 nm: A) 1st purification; B) 2nd purification. 54

20 MALDI-MS of purified QEKR 4.0^b: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix. 55

21 MALDI-MS of purified QEKR 5.0: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix. 56

22 1-D ^{1}H-NMR spectrum of glycine. 59

23 1-D ^{1}H-NMR spectrum of alanine. 60

24 1-D ^{1}H-NMR spectrum of glutamic acid. 61

25 1-D ^{1}H-NMR spectrum of glutamine. 62

26 1-D ^{1}H-NMR spectrum of arginine. 63

27 1-D ^{1}H-NMR spectrum of lysine. 64

28 1-D ^{1}H-NMR spectrum of tyrosine. 65

29 1-D ^{1}H-NMR spectrum of QEKR 4.0^b. 68

30 1-D ^{1}H-NMR spectrum of QEKR 5.0. 69

31 2-D COSY NMR spectrum of QEKR 4.0^b. 70

32 2-D COSY NMR spectrum of QEKR 5.0. 71

33 Side views of the minimized structures of QEKR 4.0^b: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen. 74

34 Barrel views of the minimized structures of QEKR 4.0^b: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen. 75

35 Side views of the minimized structures of QEKR 5.0: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen. 76

36 Barrel views of the minimized structures of QEKR 5.0: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen. 77

37 Titration of QEKR 4.0^b by NMR at pH: A) 1.20; B) 9.37. 80

38 Determining the pKa of glutamic acid in QEKR 4.0^b. Peak labeled as H7 was followed through the course of the titration. 81

39 Determining the pKa of glutamic acid in QEKR 4.0[b]. Peak labeled as H8 was followed through the course of the titration. 82

TABLE OF TABLES

Table		Page
1	The 20 amino acids are basic building blocks of proteins. They are classified as either nonpolar (hydrophobic), polar (uncharged), acidic, and basic. Properties, R groups at pH 7, and the 1- and 3-lettered codes of the amino acids are shown.	3-4
2	Amounts of amino acids and reagents used in the peptide synthesis.	21
3	Amounts of 12 M DCl (1:10 dilution) added.	43
4	Percent yields of the synthesized peptides.	44
5	Mass to charge (*m/z*) values of QEKR 4.0[b] and QEKR 5.0.	46
6	Proton chemical shifts of individual unblocked amino acids. All spectra contain a water peak at 4.7 ppm.	58
7	Proton chemical shifts of QEKR 4.0[b]. The spectrum contains a water peak at 4.7 ppm. (Note: Only 1 set of data is presented because QEKR 5.0 also have nearly identical proton chemical shifts.)	67
8	The minimized energies and electrostatic values for both QEKR 4.0[b] and QEKR 5.0 peptides were obtained in either water solvent or no solvent.	73
9	Proton chemical shifts of various labeled peaks at different pH.	79

CHAPTER 1: INTRODUCTION

Background

Amino acids are the fundamental building blocks of proteins. They are molecules that have structural as well as functional importance in a biological system. There are 20 amino acids that are known to exist naturally in this world. All natural occurring amino acids have the ability to rotate the plane of polarized light counterclockwise. Therefore, they are said to display levorotatory behavior.

The general structure of a single typical amino acid is shown in Figure 1. Central to this structure is the tetrahedral alpha (α) carbon (C_α) which is covalently linked to both the amino ($-NH_3^+$) and carboxyl ($-COO^-$) groups. The charges on the amino and carboxyl

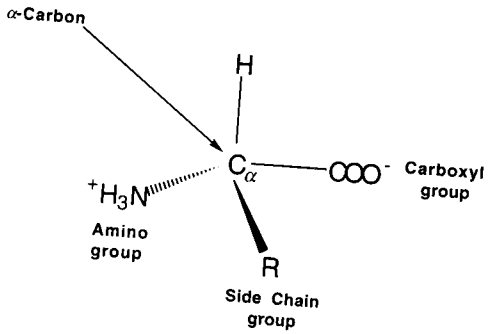


Figure 1. A typical amino acid at pH 7 is presented as a tetrahedral structure. It consists of an α-carbon, amino and carboxyl groups, and a variable side chain.

groups are formed as a result of hydrogen protonation and deprotonation in neutral (pH 7) solution, respectively; this species is known as zwitterion, a molecule that contains both a positive and negative charge. In addition, a hydrogen atom and a variable side chain group

are also attached to the α-carbon. The side chain (R group) gives each amino acid its different and unique properties. These properties can be grouped into four categories---- nonpolar (hydrophobic), polar (uncharged), acidic, and basic.[1] Table 1 shows a list of the 20 amino acids with their structures and their appropriate designations of 1- and 3-lettered codes.

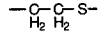
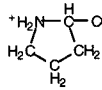
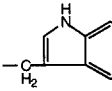
The nonpolar or hydrophobic amino acids include all those with an alkyl side chain or R group (alanine, valine, leucine, and isoleucine), as well as proline (with its unusual cyclic structure), methionine (one of the two sulfur-containing amino acids), and two aromatic amino acids (phenylalanine and tryptophan). Peptides that made of these amino acids tend to be insoluble in aqueous solution, owing to their alkyl side chains. In large protein molecules, these amino acids can aggregate into clusters inside the protein surface, forming hydrophobic pockets.

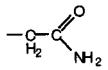
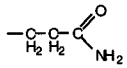
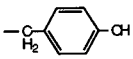
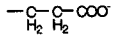
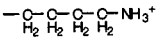
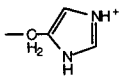
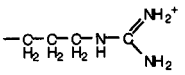
The polar and uncharged amino acids include serine, asparagine, glutamine, and glycine, threonine, cysteine, and tyrosine. These amino acids tend to form hydrogen bonds via electronegative atoms such as oxygen and nitrogen on the side chain group except glycine, which has a hydrogen atom. Hence, glycine is the only amino acid that displays achirality. Interestingly, glycine can be classified as nonpolar because its side chain contributes little or no polarity to the overall molecule. However, for organizational purposes, we have placed glycine under the polar category because like many small molecules, it tends to be relatively soluble in aqueous solution. Unlike bulky nonpolar amino acids, polar and uncharged amino acids are always soluble.

The acidic amino acids include only aspartic (aspartate) and glutamic (glutamate) acids. Their side chain groups contain a carboxyl group, which ionizes at neutral pH. In addition to hydrogen bonding, the two amino acids can participate in electrostatic interaction via their negative charged carboxyl containing side chains. They play a crucial role in stabilizing large protein molecules.

The basic amino acids include histidine, lysine, and arginine. These three amino

Table 1. The 20 amino acids are basic building blocks of proteins. They are classified as either nonpolar (hydrophobic), polar (uncharged), acidic, and basic. Properties, R groups at pH 7, and the 1- and 3-lettered codes of the amino acids are shown.

Property	Name of Amino Acid	3-Lettered Code	1-Lettered Code	R group
Nonpolar	Alanine	Ala	A	
Nonpolar	Valine	Val	V	
Nonpolar	Leucine	Leu	L	
Nonpolar	Isoleucine	Ile	I	
Nonpolar	Methionine	Met	M	
Nonpolar	Proline	Pro	P	
Nonpolar	Phenylalanine	Phe	F	
Nonpolar	Tryptophan	Trp	W	
Polar	Glycine	Gly	G	

Polar	Serine	Ser	S	
Polar	Asparagine	Asn	N	
Polar	Glutamine	Gln	Q	
Polar	Threonine	Thr	T	
Polar	Cysteine	Cys	C	
Polar	Tyrosine	Tyr	Y	
Acidic	Aspartic acid	Asp	D	
Acidic	Glutamic acid	Glu	E	
Basic	Lysine	Lys	K	
Basic	Histidine	His	H	
Basic	Arginine	Arg	R	

acids have side chains with a net positive charge at neutral pH. Like the acidic amino acids, they can also participate in electrostatic interaction via their positive charged side chains with other negative charged amino acid residues. They play an important role as proton donors and acceptors in many enzymatic reactions, and also as buffers in solution.

A common feature of all the amino acids is their ability to polymerize into peptide chains via their amino and carboxyl groups. A typical reaction joining two amino acids requires the removal of a water molecule to form a peptide (amide) bond. Careful examination of the amide bond (N-C_o) reveals partial double bond characteristics, due to delocalization of electrons on the nitrogen atom about 40% of the time. Therefore, the degree of rotation about the N-C_α (phi; ϕ) and C_α-C_o (psi; ψ) bonds of the planar molecule is restricted. The dipeptide can be further extended to larger molecules such as an oligopeptide (12-20 amino acids) and polypeptide (> 20 amino acids). Amino acids adopt certain hierarchical structures, which can be broken down into four distinct categories (small to large)----the primary (1°), secondary (2°), tertiary (3°), and quaternary (4°) structures.

The primary structure is among the simplest models found in nature. It consists of linear sequence of amino acids. The higher level of structures such as secondary, tertiary, and quaternary are involved in hydrogen bonding, hydrophobic interactions, electrostatic interactions, and Van der Waals interactions.

The two most common secondary structures found in proteins are α-helices and beta (β)-pleated sheets. The structure of the α-helix was first proposed by Taylor[2] as early as 1941 and was later discussed in detail by Bragg *et al*[3]. It was not until 1951 when Pauling *et al.*[4] elucidated and characterized much of the structure. The 3.6_{13}-helix (α-helix), where a single turn of the α-helix involves 13 atoms or 3.6 residues in the peptide chain is most common and is more stable than the 3_{10}-helix (10 atoms or 3 residues per turn).[5] The α-helix and 3_{10}-helices have backbone dihedral angles (ϕ, ψ) of -63°, -42°

and -57°, -30°, respectively. The α- and 3_{10}-helix are marked by their ability to participate in intramolecular hydrogen bonding from their amide protons to the carbonyl oxygens. For example, the α-helix has hydrogen bonds that point parallel to the helix axis where all the carbonyl oxygens are bonded to the amide protons four residues up the chain. Both C=O and N-H groups point in opposite directions.

Similarly, the β-pleated sheets can also participate in hydrogen bonding, but between two adjacent peptide chains. They are classified either parallel or antiparallel. The parallel β-pleated sheet contains both adjacent chains that run in the same direction (N-terminus ----> C-terminus or C-terminus ----> N-terminus) as oppose to opposite direction of the antiparallel β-pleated sheet.

The third type of secondary structure is "random" coil.[6] They are less prevalent in nature, but are of our interest in the study of short helical peptides. The random coil serves as a good model to study fluctuation between random coils and α-helices. We can apply this model to large protein in order to understand better of the folding and unfolding mechanisms of native structures.

The tertiary structure of proteins involves interactions of folding of a single polypeptide chain in three-dimensional space. Interactions of the side chain group through hydrogen bonding and electrostatic interactions are very common. Unlike folding of a single polypeptide chain into its tertiary structure, the quaternary structure of proteins is formed from interactions between two or more polypeptide chains.

One of many interactions that influences the stability of α-helices is hydrogen bonding. The stability of hydrogen bonding depends not only on the interaction between the N-H and C=O groups, but also on the environment such as solvent, concentration, and temperature. In 1955, Schellman[7] investigated on the stability of hydrogen-bonded peptide structures in aqueous solution. He proposed that free energy stabilization was more important than the energy; he estimated the free energy of the two chains to be greater than

600 cal. Furthermore, Schellman suggested that the effect of temperature on hydrogen-bonded peptides was enthalpically and entropically based; he showed that the temperature of the structure's unfolding was approximately equal to $\Delta H_{res} / \Delta S_{res}$, where "res" means amino acid residues.

During the late 1950s and early 1960s, statistical mechanical models were developed for the transition between α-helix and random-coil states in polypeptides. The principal models at the time were Zimm-Bragg[8] (ZB) and Lifson-Roig[9] (LR) theories. The procedure for obtaining parameters for the ZB theory is to measure thermal unfolding curves using circular dichroism (CD) for a set of polypeptides with uniform sequences and varying chain length. Zimm and Bragg employed polymers of γ-benzyl-L-glutamate in their calorimetric studies to find enthalpies. Ackermann and Neumann[10], who also used calorimetric determination, later confirmed the ZB results. The nucleation constant and propagation parameter are included in the ZB and LR theories. The helix nucleation constant (σ for ZB and υ^2 for LR) enters once into the sequence peptide whereas the propagation parameter or helix growth (s for ZB and w for LR) enters repeatedly, once for every stabilized helical residue.[11] The s and w terms in ZB and LR, respectively measure how well each amino acid stabilizes or destabilizes helical formation. The LR theory is preferred over the ZB theory in analysis of helix-coil transition because the usage is simpler; the w term is associated with the amino acid residue whereas s is associated with amide link of two amino acids. To make more use of the two helix-coil transitional theories, Qian and Schellman[12] interconvert the ZB and LR parameters into one general set of equations. They are given as follows:

$$s = w / (1 + \upsilon) \qquad \text{and} \qquad \sigma = \upsilon^2 / (1 + \upsilon)^4$$

Modification of the LR theory to include other parameters such as N- and C-cappings has arisen as a result of studies which suggest that certain amino acids occur more frequently than others at specific locations in α-helices, β-pleated sheets, and reverse turns

of globular proteins.[13] Frequency differences occur at N-terminal end (N-cap) and C-terminal end (C-cap) of secondary structures. The significance of the N- and C-caps was first described by Richardson and Richardson[14], and later by Presta and Rose[15]. Presta and Rose suggested that side chain-main chain hydrogen bonding is an important determinant in the stability of the α-helix such that if hydrogen bonding among the four N-H and C=O groups at both ends does not occur, then helix fraying might result.

In the classical ZB and LR theories, N- and C-cap residues are not included in the statistical calculation because they are believed to have no effect on helical stability. For example, the LR theory defines the N- and C-caps as coil residues with no helical (ϕ, ψ) dihedral angles, and the caps contribute no significance to thermodynamic calculations. However, recent studies suggest that they do add considerable thermodynamic weight to the free energy of the 20 amino acid residues[16,17,18]----observations which were described by earlier researchers. Mutation experiments from barnase enzyme indicate that glycine stabilizes both the N- and C-caps of the α-helix more than alanine by 0.4 to 2 kcal/mol.[19] Other reports indicate that asparagine is better in stabilizing the N-cap than glycine whereas changing amino acids stabilizing the C-cap has minimal effect.[31]

Helix propensities, or the intrinsic tendencies of amino acids to form helices are not enough to satisfy physical chemists who are interested in obtaining both values of change in enthalpy (ΔH) and entropy (ΔS) to determine the effect of helical formation. Knowing thermodynamic values are valuable because they offer a quantitative answer to test whether helix formation is spontaneous or not. From Schellman's work, he estimated the enthalpy and entropy per residue to be roughly -1.5 kcal/mol[20] and -1.4 kcal/mol[7], respectively. From this information, he concluded that formation of stable helical peptides could be possible in aqueous solvent.

The 1960s and early 1970s was marked by extensive syntheses of polypeptides with repeating amino acid residues. Poly-α-amino acids derived from L-lysine and its lower homologs, L-ornithine and L-α-γ-diaminobutyric acid were studied for helical

stability.[21,22] Researchers found that helix propensities of these basic poly-α-amino acids appear to increase as the length of their side chains increases with the addition of methylene groups. More recently, alanine-based short peptides containing L-lysine and its lower homologs, spaced five amino acid residues apart also display high helix propensities when their side chains are lengthened.[23]

Other poly-α-amino acids such as poly-N^5-(3-hydroxypropyl)-L-glutamine (PHPG) were used to study on the influence of thermodynamics[24], hydrophobicity[25], and solvent[26] have on helical stability. Similar to the result of the poly-α-amino acids of L-lysine and its homologs, generic prototypes of poly-N^5-(w-hydroxyalkyl)-L-glutamine with long side chain length also display an increase in helicity in the following order from high to low: poly-N^5-(4-hydroxybutyl)-L-glutamine (PHBG) > PHPG > poly-N^5-(2-hydroxyethyl)-L-glutamine (PHEG). In water, the PHPG is partly helical because its non-ionized side chain groups are unable to form stable interactions. The degree of helicity of PHPG increases with decreasing temperature and increasing molecular weight. Solutes such as sodium hydroxide was found to be more helical-inducing than sodium chloride or water. The ΔH and ΔS values for the random coil to the helical transition of a single amino acid of PHPG in water at room temperature are -130 cal/mol and -0.45 e.u., respectively; in methanol-water (3:7 v/v), ΔH = -170 cal/mol and ΔS = -0.45 e.u. The helix growth parameter, w in both solvents is 0.017. These thermodynamic data were analyzed through the LR theory by Lotan and coworkers.

Homopolymers were not the only subject of extensive investigation, but also copolymers consisting of mixtures of two different amino acid components were studied as well. Fasman and coworkers[27,28] used random copolymers of PHPG with L-leucine to study factors that influence helix stability. The helicity of the copolymers was found to increase with: (1) decreasing temperature; (2) increasing methanol concentration; and (3) increasing mole ratios of leucine in the copolymer. Contrary to Lotan and coworkers' data, the thermodynamic values analyzed via the ZB theory at 25°C of a single amino acid from

random coil to a helical state of PHPG in water were determined to be: ΔH = -155 cal/mol, ΔS = -0.57 e.u., and s = 0.98; likewise, for a single amino acid in the leucine polymer in water: ΔH = -109 cal/mol, ΔS = +0.88 e.u., and s = 1.2.

Host-guest studies were also performed on random copolymers (host) using natural occurring amino acids (guest) to determine statistical weight of the propagation parameter (s) and nucleation constant (σ) in the ZĿ theory.[29,30,31] From helix-coil transitional calculations, Scheraga and coworkers observed large range of helix-making and helix-breaking tendency of the amino acids. They determined that if the s term is greater than 1 ($s > 1$), then this indicates helical tendency of an amino acid; if $s < 1$, then the amino acid has helix-breaking tendency. However, if s is equal to 1 ($s = 1$), then this signifies helix indifference of an amino acid to helix formation. For all the twenty amino acid residues at 20°C, the s values were close to 1 (± 20%) except for proline and glycine, which are strong helix breakers. However, the σ values fluctuate among each amino acid.

By the 1970s, scientists began to look at helix formation of defined peptide fragments on proteins. In 1971, Brown and Klee[32] reported that C-peptide (residues 1-13 of ribonuclease A) shows partial helix formation in water at 1°C and is strongly pH dependent. But early studies on other protein models such as myoglobin[33] and staphylococcal nuclease[34] showed no stable helix formation at 25°C. Brown and Klee's experiment was later confirmed by Baldwin and coworkers who showed that there were two important amino acid residues important in the formation of helical segments in the protein. They were glutamate and histidine located on the opposite ends at position 2 and 12 of the C-peptide fragment, respectively.[35] The formation was the result of electrostatic interaction. In addition to the Glu 2-His 12 stabilization, they found that Arg 10 and Glu 2 also participate in salt bridge interaction. Furthermore, hydrogen bond interaction of Phe 8 and His 12 was also reported. Other work was investigated in the S-peptide (residues 1-20 of ribonuclease A) using nuclear magnetic resonance (NMR) to confirm helical formation

as well.

By discovering that certain regions of the ribonuclease A can form a stable structure, the next logical experiment is to test if altering or substituting these amino acids could also produce a similar result. Baldwin *et al.* found that by interchanging Ala 11 with His 12, the hydrogen bond interaction disintegrated, probably result from no (*i*, *i* + 4) interaction.[36] On the other hand, replacing either Phe 8 or His 12 by Ala produces an increase in helicity when compared to Phe 8-His 12. But replacing both amino acid residues at position 8 and 12 with Ala, produces an even higher degree in helical content. This finding has prompted an ongoing investigation to study how short alanine-based peptides could produce such high helical content.

Experiments by Marqusee *et al.*[37] in 1989 have led to the findings that alanine-based short peptides are indeed very stable and show helical formation in water. The result has been surprising because in past studies, α-helix segments from proteins were found to be unstable in aqueous solution. The alanine-based peptides are solubilized by insertion of three or more residues of negative charged glutamate and positive charged lysine in defined salt spacing.[38] There are no clear explanations on why alanine residues have high helix-inducing potential, or any other amino acid residues that show this kind of property. Researchers are now looking at the origin of this phenomenon.

Most of the work on helix stabilization of salt bridges in short peptides were investigated by Baldwin and coworkers. They were interested as to how glutamate and lysine could form electrostatic interactions in defined salt spacing. They synthesized two subclasses of (*i*, *i* + 3) and (*i*, *i* + 4) peptides[37,38,39]:

1. **(*i*, *i* + 3) 3.0 Peptides**

a. EK 3.0 = Ac-A-E-A-A-K-A-E-A-A-K-A-E-A-A-K-A-NH_2

b. KE 3.0 = Ac-A-K-A-A-E-A-K-A-A-E-A-K-A-A-E-A-NH_2

2. ***(i, i + 4) 4.0 Peptides***

a. EK 4.0 = Ac-A-E-A-A-A-K-E-A-A-A-K-E-A-A-A-K-A-NH_2

b. KE 4.0 = Ac-A-K-A-A-A-E-K-A-A-A-E-K-A-A-A-E-A-NH_2

Using CD, Baldwin found the order by which these four peptides displayed helicity at pH 7 (high to low): EK 4.0 > KE 4.0 > EK 3.0 > KE > 3.0.

From this novel discovery, others have attempted to size a variation of these peptides. Berger *et. al.*[40] created two hybrid peptides, combining both (*i, i* + 3) and (*i, i* + 4) salt spacing into a mixed peptide:

EK 3.3 = Ac-Y-E-A-A-K-A-A-E-A-A-A-K-A-A-E-A-A-A-K-NH_2

EK 3.7 = Ac-Y-E-A-A-A-K-A-E-A-A-K-A-E-A-A-A-K-A-NH_2

The purpose of modeling such hybrid peptides is to maximize and minimize the distance of both (*i, i* + 3) and (*i, i* + 4) salt spacing, respectively to one turn of an α-helix. The closer the value of the spacing gets to 3.6 amino acids per turn, the greater the stability of the helix will become when compared to the EK 3.0 and EK 4.0 peptides.

Problem

Large and complex structures of proteins can be complicated to study. Multiple side chain interactions and other intermolecular and intramolecular forces can influence surface folding through stabilization of protein structures. A simpler approach to study

these interactions is examine their simple and short peptide structures such as α-helices on proteins. The focus of this project is to synthesize similar version of Baldwin's short peptides to determine their helical conformations and also their salt bridge interactions.

The investigation of determining how Baldwin's alanine-based peptides induce helicity still remains a mystery. It is clear from previous experiments that when opposite charged amino acids are placed in certain defined sequence on a peptide, it displays a high degree of helicity via forming stable electrostatic interactions. The goal of the project is to observe if our version of the peptides (QEKR 4.0[b] and QEKR 5.0) meet this condition, tested through a series of instrumental techniques which are described in the later section. If so, can we detect this type of interaction using a trapping reagent probe?

One such trapping reagent probe is cyanogen (ethanedinitrile). Cyanogen is known to trap salt bridges directly through condensation of paired groups.[41,42] Figure 2 shows a covalent (amide) bond is being formed as a result of this reaction. As a conformational

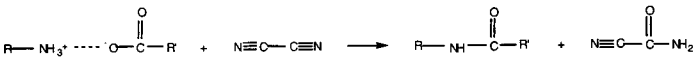


Figure 2. Salt bridge among two opposite charged amino acids can be trapped using cyanogen to form amide links.

probe, cyanogen can specifically trap salt bridges of certain catalytic enzymes such as hens egg white lysozyme and ribonuclease S protein. In addition, it is able to trap only intramolecular paired species as oppose to intermolecular species. Hence, this probe reagent gives both selectivity and specificity.

We would like to use cyanogen to trap salt bridges, specifically lysine and glutamate in peptides with $(i, i + 4)$ salt spacing that we have chosen in our design. Theoretically, $(i, i + 4)$ peptide should form an intramolecular salt bridge. The concept of

selectivity and specificity serves the purpose in our study of peptide helicity. As a control, we will use (i, i + 5) salt spacing. This should not form a salt bridge and should not react with cyanogen, if it is indeed both selective and specific for salt bridges. If amide links are confirmed through trapping, then we know definitively that the peptides must show some sort of helicity. However, the degree of helicity depends on which peptide that we analyze. Ultimately, we would like to determine what kind of factors influence and stabilize our synthesized peptides?

Peptide Design of QEKR 4.0[b] and QEKR 5.0

Much of this frontier work was made possible by the advent of the solid-phase peptide synthesis (SPPS) by Merrifield[43] in which he first synthesized a tetrapeptide using this approach. As a result of his novel idea, SPPS has become more diverse in many modern day applications. Researchers can now synthesize very long peptides up to nearly 200 amino acid residues long. The two main strategies for synthesis of peptides utilize 9-fluorenylmethoxycarbonyl (Fmoc) as N-α protecting group and *tert*-butoxycarbonyl (Boc) as side chain protecting group.[44] The Fmoc and Boc groups both exploit either the main chain or side chain of an amino acid for coupling or deprotection reactions. However, we are using Fmoc as main chain and Boc as side chain protecting groups in our strategy. The two protecting groups are chosen because they work orthogonal to one another. For example, Fmoc protecting group can be removed in slightly basic conditions whereas the Boc protecting group is removed in slightly acidic conditions. Therefore, the protecting groups would not interfere with each other's chemistry since one group is insensitive to the other's removal conditions. The synthesis and mechanism of our short peptides will be discussed in more details in the Experimental section.

The design of our peptides has to meet certain criteria. The length of the peptide is 18 amino acid residues long (18 mer). The synthesis of the peptides is divided into 2 parts---experimental and control. The experimental involves synthesizing a helical peptide that

contains residues which can form salt bridge with (i, $i + 4$) salt spacing. Also, peptide with (i, $i + 5$) salt spacing will be synthesized and used as the control.

Of the 18 amino acids in the experimental peptide, we will use one each of tyrosine (Y), glycine (G), glutamate (E), lysine (K), and arginine (R); three of glutamines (Q); and ten of alanines (A). The majority of the peptide consists of alanines because they are known to induce helicity in short peptides. Not only alanines provide a good foundation for this peptide design, their small nonpolar side chains ($-CH_3$) would not interfere with side chains of other amino acid residues. Glutamines are used in our peptide to ensure solubility because they are polar and uncharged. In addition, they are (i, $i + 5$) spaced to avoid any hydrogen bonding among the amide group with other electronegative or hydrogen atoms. More importantly, the amide groups of glutamines do not ionize readily in solution, so they can not interfere with the electrostatic interaction between lysine and glutamate. Like glutamines, an arginine is incorporated near the C-terminus for solubility. In our cyclic peptides, there are no charged residues (lysine and glutamate) present to provide solubility since they will be formed as lactam bridge. Unless an additional acidic or basic amino acid is used, the cyclic peptide will not be suited for some instrumental analyses. Therefore, both glutamines and arginines are needed to overcome this weakness.

Negative charged glutamate and positive charged lysine are placed toward the N- and C-termini of the peptide, respectively. The charges of glutamate and lysine can better stabilize their respective terminus ends due to opposing dipole moment, and they can also provide extra solubility to the helix like glutamine and arginine. In addition, the (i, $i + 4$) salt spacing can further stabilize the helicity of the peptide. Glutamate and lysine are arranged in such a way that they are centered to give both symmetry and less fraying at the ends of the helix. The N- and C-termini of the peptide are acetylated (-Ac) and amidated ($-NH_2$), respectively to avoid dipole destabilization by having charged amino acid residues at the ends of the peptide. Tyrosine is placed at the end of the C-terminus as a ultraviolet (UV) marker to monitor the peptide's concentration whereas glycine, a helix breaker, is

placed next to tyrosine to enhance the precision and reproducibility of the CD measurements. The glycine ensures that the helix always terminate at, or before this position in order to prevent tyrosine from disrupting CD measurements when placed as part of the C-terminal cap[18] as well as the N-terminal cap[45].

Two peptides with both (i, $i + 4$) salt spacing are being considered in the experimental:

QEKR 4.0^a = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-A_{13}-E_{12}-Q_{11}-A_{10}-A_9-K_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

QEKR 4.0^b = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-E_{13}-A_{12}-Q_{11}-A_{10}-K_9-A_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

The QEKR 4.0^a peptide presents some problems: (1) Q_{11}----K_8 produces weak (i, $i + 3$) interaction; (2) Q_{16}----E_{12} produces strong (i, $i + 4$) interaction; and (3) E_{12} might interact with Q_{16}, therefore, disrupting its salt bridge with K_8. Although the QEKR 4.0^b peptide contains both weak (i, $i + 3$) Q_{16}----E_{13} and (i, $i + 3$) K_9----Q_6 interactions, none of these interactions are detrimental to the formation of the salt bridge because of no strong (i, $i + 4$) interactions. Thus, QEKR 4.0^b is preferred over QEKR 4.0^a and will be used in our synthesis.

Similar in amino acids composition to that of the experimental, the peptide of the control is designed in such a way to have no or few interactions among the glutamate and lysine as much as possible. One feasible solution is to synthesize peptides with (i, $i + 3$) or (i, $i + 5$) salt spacing. Opposite charged amino acids in (i, $i + 3$) or (i, $i + 5$) spacing are less likely to form stabilized salt bridge than in (i, $i + 4$) salt spacing because of their positioning on the helix. The most stable helical spacing found in nature is 3.6 residues per turn. Since neither (i, $i + 3$) salt spacing nor (i, $i + 5$) salt spacing is close to 3.6 residues per turn, the rationale is that they will form very unstable interaction. We can use trapping reagents to detect if amide links had formed as a result of salt bridge interaction. If no

amide links were detected, then the reagent is specific in trapping only peptide with (i, i + 4) salt spacing since (i, i + 3) and (i, i + 5) salt spacing are weak and too unstable to form. However, if there were amide links, then the trapping reagent is nonspecific and reacts randomly with charged amino acid residues on the helix.

Two peptides with (i, i + 3) and (i, i + 5) salt spacing are being considered in the control:

QEKR 3.0 = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-E_{13}-A_{12}-Q_{11}-K_{10}-A_9-A_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

QEKR 5.0 = Ac-A_{18}-A_{17}-Q_{16}-A_{15}-A_{14}-E_{13}-A_{12}-Q_{11}-A_{10}-A_9-K_8-A_7-Q_6-A_5-A_4-R_3-G_2-Y_1-NH_2

K_{10}----Q_6 forms strong (i, i + 4) interaction in the QEKR 3.0 peptide. On the other hand, both Q_{16}----E_{13} and Q_{11}----K_8 in the QEKR 5.0 form weak (i, i + 3) interaction. QEKR 5.0 is preferred over the QEKR 3.0 peptide because it lacks strong (i, i + 4) interaction. Therefore, QEKR 5.0 will be used as the control.

Finally, in order to compare the peptides that we have selected in the experimental and control, cyclic peptides with appropriate (i, i + 4) and (i, i + 5) lactam bridges, respectively will be synthesized in our peptide design. Ösapay and Taylor[46] recently showed that cyclic peptides with lactam bridges adopt helical formation.

Instrumental Analysis

Once these short peptides are synthesized, they will need to be purified and concentrated before accurate CD measurements and NMR spectral data are taken. The purification process will be done using a gradient high performance liquid chromatography (HPLC) and the purity will be confirmed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF). CD spectra will be taken between 180 nm and 260 nm because α-helices are known to show two minima at 208 nm and 222 nm and a maximum

close to 1(nm.[47]

A ugh CD spectra tell us quantitative result on the degree of helicity of the peptide, it does not provide qualitative data to analyze the dynamics of the structure. We would like to use NMR spectroscopy to determine its helical state[48] as well as follow the course of reaction over time. In secondary structures of large polypeptides or proteins, strong nuclear Overhauser enhancement (NOE) signals in spaced proton-proton distances can be detected. The distance between two r otons can be denoted as $d_{AB}(i, j)$, where A and B are amide protons and i and j are the sequence of amino acids on a peptide. In addition, other NOE signals between the N-H proton (A) and C_α hydrogen (B) can also be detected as well.[49,50] There are recent experimental reports that show hydrogen-deuterium exchange rate and equilibrium thermodynamic stability can be analyzed on amide and C_α hydrogens of short peptide (Ac-Y-K-E-D-R-D-R-A-A-A-A-NH_2) using ^{1}H-NMR.[51] An alternative way is to use ^{13}C-NMR to analyze thermal helix-coil transition of isotopic labeled short peptide. Shalongo *et. al.*[52] analyzed an enriched ^{13}C alanine-labeled Ac-W-(E-A-A-A-R)$_3$-A-NH_2 peptide.

Another method to analyze secondary structure of polypeptides, but not used in the Experimental, is infrared (IR) spectroscopy. Previous IR studies have shown that α-helices and β-pleated sheets absorb strongly at band regions between 1650-1660 cm^{-1} and 1630-1640 cm^{-1}, respectively.[53] However, the bands resolved by conventional IR spectroscopic techniques offer poor resolutions. Susi and Byler[54] used Fourier transform IR (FT-IR) spectroscopy to interpret structures of proteins. FT-IR spectroscopy provides several advantages: (1) high resolution; (2) increase sensitivity; (3) high signal-to-noise (S/N) ratio; and (4) stronger and more accurate frequency.

Finally, Protein Data Bank (PDB) searches will be employed to look for homology among our short peptides with other defined protein sequences. PDB was established in 1971 as a computer-based archival file for macromolecular structures.[55] ts purpose is to collect, standardize, and distribute atomic coordinates and other data from crystallographic

studies. In addition, molecular modeling studies of our synthesized peptides will be used extensively to visualize the structures for any salt bridge formation and other intramolecular interactions. Minimization on the geometry of the α-helices will be run to calculate the energy.

CHAPTER 2: EXPERIMENTAL

Synthesis of QEKR 4.0[b] and QEKR 5.0

The protocol for the synthesis of short peptides uses the standard Fmoc for N-α protection and Boc (or *t*-butyl) for side chain protection types of strategy. These N-α and side chain protecting groups react orthogonally under base and acid conditions, respectively. Beside Boc group, 2,2,5,7,8-pentamethylchroman-6-sulfonyl[56] (Pmc) was used as N^G side chain protection for arginine in our synthesis. Like Boc, Pmc is cleaved under acidic conditions. Coupling of the residues was performed with benzotriazole-1-yl-oxy-tris(dimethylamino)-phosphonium-hexafluorophosphate (BOP) (Castro's Reagent) and *N*-hydroxybenzotriazole (HOBt) reagents. All the procedures listed in this method section are modified and adapted from *Amino Acid and Peptide Synthesis*[44].

The peptides were synthesized using Rink Amide resins[57] (4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin) on a polystyrene polymer. In Figure 3, the amino group on the Rink Amide resin is blocked with a Fmoc group to protect the nitrogen from undergoing nucleophilic attack. All resins were purchased from

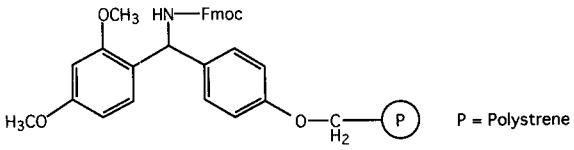


Figure 3. The structure of the Rink Amide resin.

Novabiochem®. The (i, i + 4) and (i, i + 5) peptides were synthesized starting with approximately 0.50 g of the resins. The resins employ a loading factor of 0.46 mmol/g,

which theoretically can produce 0.23 mmol of peptide in a given synthesis, provided that each of the amino acid coupling reactions goes to completion. However, to ensure that the reaction goes to 100% completion, a five-fold excess (1.15 mmol) amount of amino acid was added during each of the coupling reaction. Table 2 lists the amounts of each amino acid and reagent used in our synthesis.

Table 2. Amounts of amino acids and reagents used in the peptide synthesis.

Amino Acid/Reagent	Residue MW (g/mol)	Overall MW (g/mol)	Mmoles (5-fold excess)	Amount Added (g)
Lys (Boc)	129.18	468.5	1.15	0.5388
Tyr (*t*-butyl)	163.18	459.6	1.15	0.5285
Glu (*t*-butyl)	129.12	425.5	1.15	0.4893
Arg (Pmc)	157.20	662.8	1.15	0.7622
Ala	71.08	311.3	1.15	0.3580
Gly	57.05	297.3	1.15	0.3419
Gln	128.13	368.4	1.15	0.4237
BOP	-----	442.5	1.15	0.5089
HOBt	-----	135.1	1.15	0.1554
DIEA/DMF	-----	129.3	1.15 of DIEA	5.0-mL
Acetic anhydride	-----	102.1	1.15	0.1174 (127.5-μL)

Seven different N-α-Fmoc protected amino acids were used to make the (i, i + 4) and (i, i + 5) peptides. They were N-α-Fmoc-L-alanine, N-α-Fmoc-N-ε-Boc-L-lysine, N-α-Fmoc-L-glutamic acid-γ-t-butyl ester, N-α-Fmoc-N^{G}-Pmc-L-arginine, N-α-Fmoc-L-glutamine, N-α-Fmoc-O-t-butyl-L-tyrosine, and N-α-Fmoc-glycine. The amino terminus of all the N-α amino acids are protected with a Fmoc group. However, for the side chain, lysine is protected by Boc; glutamic acid by a t-butyl ester group; arginine by Pmc; and tyrosine by a t-butyl ether group. The side chains of alanine, glutamine, and glycine are free and unblocked because their R groups are insensitive to acidic and basic conditions.

Alanine and BOP were purchased from both Bachem California and Novabiochem®; arginine, lysine, glutamic acid, and HOBt from Bachem California; and glycine, glutamine, and tyrosine from Novabiochem®. All other reagents such as piperidine (Pip), anhydrous *N*,*N*-dimethylformamide (DMF), *N*,*N*-diisopropylethylamine (DIEA), trifluoroacetic acid (TFA), anisole, thioanisole, acetic anhydride, and anhydrous ether were supplied by Aldrich.

Amino acid coupling reactions were carried out in a reaction flask that is cemented with a sintered and fritted disc filter on the bottom. The excess reagents trapped inside the reaction flask can be removed via suction filtration. Figure 4 is a diagram of the glass sintered vessel. The reaction flask was secured in a Burrel Wrist Action Shaker® to agitate the coupling mixture and to increase the rate of the reaction.

Deprotection of Fmoc Main Chain Protecting Groups

The first step in the synthesis was removal of the Fmoc group from the amino group on the linker part of the resins. First, the resins (0.50 g) were soaked in 15-mL of DMF for 15 minutes to wash off any impurities that might have been picked up from manufacturer's processing. The soaking process is important in order to expand the size of the resin beads and to increase the surface area, thereby, allowing effective coupling

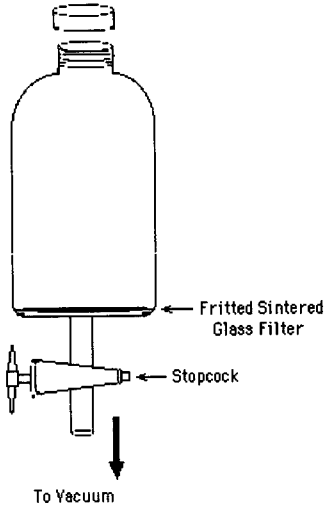


Figure 4. The diagram of a glass sintered vessel.

reaction to take place. Fmoc blocking groups on the coupled amino acids of the peptide chain were also removed in a similar fashion. The procedures for this part are listed as follows:

Step 1: 15-mL of 30% by volume Pip/DMF solution is added into the reaction flask and the mixture is agitated in a Burrel Wrist Action Shaker® for one minute

***Note*:** To ensure freshness and to avoid contamination, the 30% by volume Pip/DMF solution is to be replaced every four to five days

Step 2: The reaction mixture is removed via suction filtration

Step 3: Small amount of DMF is squirted onto the inner wall of the reaction flask to wash

down any remain of the resins and other contaminants

Step 4: Another 15-mL Pip/DMF solution is added into the reaction flask and is shaken for ten minutes to ensure all Fmoc groups are removed

Step 5: The reaction mixture is removed via suction filtration after the ten minutes

Step 6: The now deprotected peptide is washed with five consecutive times of 15-mL of DMF each for one minute to remove any excess reaction mixture and contaminants

Figure 5 illustrates the mechanism by which a Fmoc protecting group is cleaved. The removal of the Fmoc is easily achieved by addition of piperidine. Piperidine, a mild base and a secondary amine, is able to remove one of the acidic hydrogens on the Fmoc group. When cleaved, the piperidine is in its salt form.

A typical peptide chain (**1**), where n is the number of different side chain groups, consists of a sequence of amino acid residues with the N-terminus blocked by a Fmoc group. Removal of a hydrogen on the Fmoc group by piperidine forms the carbanion species (**2**) and the piperidium ion (**3**). Electron delocalization of intermediate **2** produces carbon dioxide (**4**), compound **5**, and a deprotected amino acid peptide chain (**6**). Further attack by piperidine on compound **5** forms product **7**.

Amino Acid Activation

In this part of the protocol, activation of an amino acid is necessary to elongate the length of the peptide chain through coupling. As described previously, five-fold excess of the amino acid and activation reagents (BOP and HOBt) were added to the reaction to ensure 100% completion. The activation solution (1.15 mmol of DIEA in DMF) was prepared by mixing 4.0-mL of DIEA with 96.0-mL of DMF, producing a solution concentration of DIEA that is 0.23 M. The procedures for this part are listed as follows:

Step 1: Mix 5-mL of 0.23M DIEA/DMF solution with 5-mL DMF

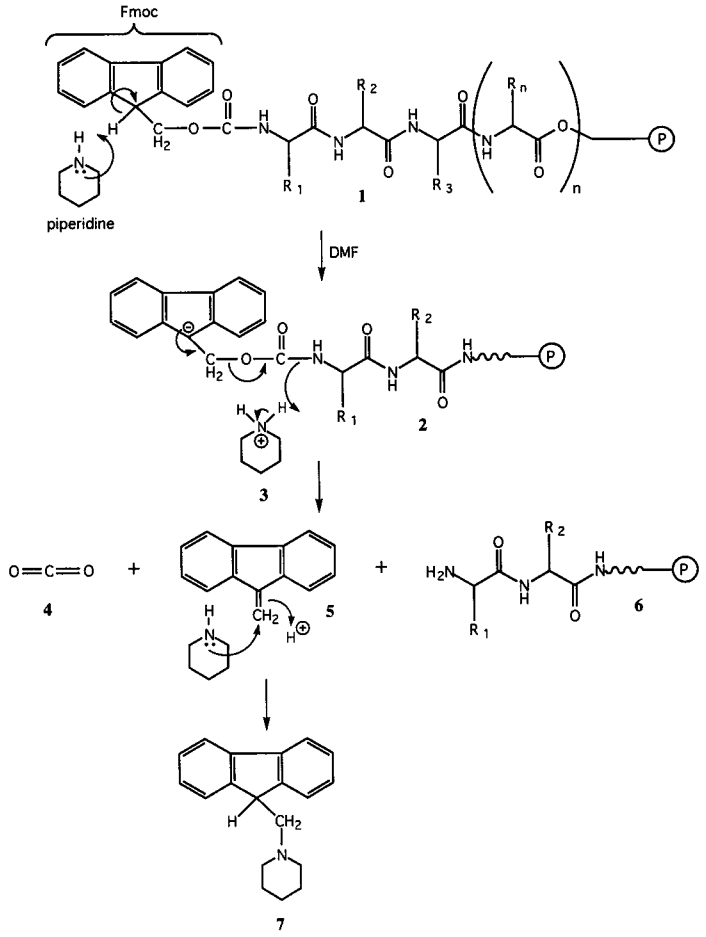


Figure 5. Base sensitive Fmoc protecting group on main chain of amino acid 1 is removed by piperidine.[6]

Step 2: Weigh out five-fold excess of BOP (0.5089 g), HOBt (0.1554 g), and the amino acid to be activated (see Table 2)

Step 3: Combine the above reagents into a 50-mL Erlenmeyer flask and shake well for three minutes until the BOP, HOBt, and amino acid are totally dissolved

Coupling of Amino Acid to the Peptide Chain

After the amino acid has been activated, it can be added to the deprotected peptide chain or resin (if it was the first coupling). The reaction was shaken for approximately one hour. A Kaiser test[58] was performed periodically during the middle and the end of the coupling reaction to detect if free amines were present. The procedures for this part are listed as follows:

Step 1: Add the activated amino acid solution into the reaction flask

Step 2: Allow the activated amino acid to react for approximately one hour by shaking

***Note*:** Amino acids that are coupled to the further end along the peptide chain usually take longer time than the beginning, therefore, it is advisable to monitor the activity of the coupling at a 15 minute interval to check for completion by the Kaiser test

Kaiser Test

The Kaiser test was performed by placing a few grains of the peptide resins into a small test tube and a drop of each of the three Kaiser test reagents. One bottle of the Kaiser test reagent contained 0.281 M ninhydrin in *n*-butanol; the second with 42.5 M phenol in *n*-butanol; and the third with 2.00×10^{-4} M KCN in pyridine. All the Kaiser reagents were supplied by Aldrich. The test tube with the peptide resins and Kaiser reagents, and a control with only the Kaiser reagents were placed into a beaker of boiling water for five minutes. The color of the solution indicates whether or not free amines are present. A dark

purple blue indicates a failure of the coupling; a light blue suggests that not all of the peptides have been coupled with the activated amino acids; and a colorless or a dimmed light blue indicates essentially complete coupling. The intensity of the color might vary depending on how many grains are added into the test tube. The procedures for this part are listed as follows:

Step 1: After one hour, coupling reagents, side products, and other waste are removed via suction filtration

Step 2: The now lengthened and protected peptide chain is washed with five consecutive times of 15-mL of DMF each for one minute

Step 3: A Kaiser test is performed for free amine detection

Step 4: If the Kaiser test is negative, then the protected peptide chain can be further extended by repeating the above deprotection, amino acid activation, and coupling procedures again or it can be safely stored overnight

***Note*:** To avoid a false reading from the Kaiser test, make sure that the test tubes are immaculately clean

Figure 6 illustrates both mechanisms by which an amino acid (**8**) is activated and then coupled to the growing peptide chain (**6**). The removal of a hydrogen from the carboxyl group by DIEA activates the amino acid (**9**). DIEA, a basic tertiary amine, is capable on prolonged reaction of deprotonating the C_α hydrogen and racemizing the amino acid. However, short reaction times and steric factors limit this reaction. The deprotonated amino acid (**9**) attacks BOP to form a phospho-ester intermediate (**10**) where the tris(dimethylamino)-phosphonium group gets displaced as a result of nucleophilic attack by the nitrogen on the amino group of peptide chain **6**. The final product is an elongated peptide chain (**11**).

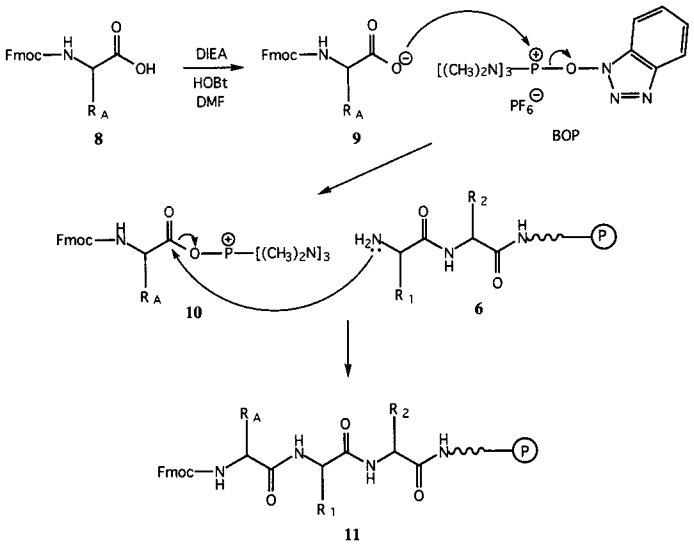


Figure 6. Amino acid A (**9**) is coupled to the peptide chain (**6**) at its amino end via the BOP, HOBt, and DIEA reagents.[6]

N-terminus Capping

After the elongation process has been accomplished and the last Fmoc group is removed by piperidine, the deprotected N-terminal end of the peptide chain is capped with an acetyl group. Acetic anhydride serves the purpose to cap the N-terminal end. A five-fold excess of 0.1174 g acetic anhydride (127.5-μL) was used. The procedures for this part are listed as follows:

Step 1: Fmoc protecting group is removed from the amino terminus of the last amino acid on the peptide chain, just like the deprotection procedure

Step 2: Five-fold excess of acetic anhydride (0.1174g; 127.5-μL) and 10-mL of DMF are combined and mixed well in a 50-mL Erlenmeyer flask

Step 3: The capping mixture is added into the reaction flask and allowed to react for one hour

Step 4: The capping mixture and other contaminants are removed via suction filtration after one hour

Step 5: Excess capping reagents are removed with five consecutive washes of 15-mL of DMF each for one minute

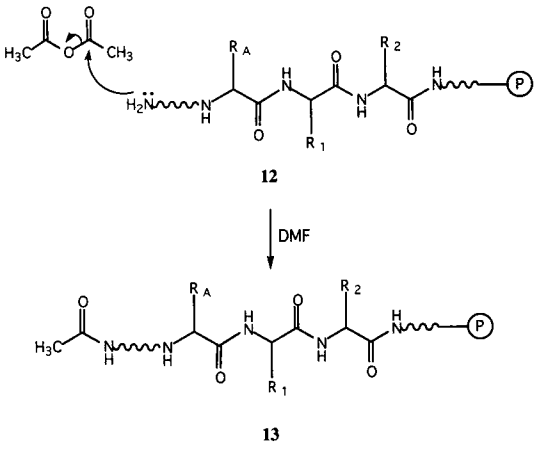
Step 6: The now capped peptide chain is tested for the presence of free amines using the Kaiser test

Figure 7 illustrates the mechanism by which the amino end of the peptide chain (**12**) is capped with an acetyl group. The nucleophilic nitrogen at the amino end attacks the carbonyl group of acetic anhydride by displacing one of its acyl groups, which behaves as a good leaving group. Product **13** is the result of nucleophilic acyl substitution.

Cleavage of Resins and Side Chain Protecting Groups

After the N-terminus of the peptide chain has been capped, the peptide needs to be cleaved from the polymer resin linker and also the side chain protected groups on the blocked amino acid residues. All of these groups are sensitive to acidic conditions. The concentration of the cleavage solution was composed of 90% TFA, 5% anisole, and 5% thioanisole by weight. A 35-mL portion of cleavage solution was prepared by adding 30.41-mL of TFA, 2.51-mL of anisole, and 2.36-mL of thioanisole. Anisole and thioanisole are scavengers that react with *t*-butyl carbocations and other positive charged species to prevent them from alkylating electron-rich side chain groups of amino acid residues.

One must take extreme caution when handling this part. The protocol for the



12

13

Figure 7. The amino end of the peptide chain (**13**) is capped with an acetyl group.[6]

experiment is to be performed under the hood to avoid toxic fumes given off by TFA. But more importantly in step 7 of the procedure, it is imperative that the filtrate is never to be discarded because the peptides are dissolved in it. The procedures for this part are listed as follows:

Step 1: The capped peptide chain is washed with five consecutive times of 15-mL of DMF each for one minute

Step 2: The resins are physically removed from the reaction flask and placed into a 50-mL Erlenmeyer flask

Step 3: The Erlenmeyer flask is sealed with a piece of parafilm, punctured with small holes to allow air and vapor to be removed under vacuum

Step 4: The resins inside the Erlenmeyer flask are frozen with dry ice and placed into a

Labconco Freeze Dry System/Freezone® 4.5 apparatus for lyophilization at -50°C overnight

Step 5: A cleavage solution is prepared from 30.41-mL of TFA, 2.51-mL of anisole, and 2.36-mL of thioanisole

Step 6: After lyophilization, 15-mL of the cleavage solution is added to the capped peptide chain into a 50-mL Erlenmeyer flask, sealed with a cork and is stirred with a magnetic stir bar inside the hood at room temperature for two hours

Step 7: The resins are removed via suction filtration through a glass sintered filter using a 50-mL side-arm Erlenmeyer flask (See Figure 8)

***Note*:** The now unbound and free-up peptide chain is dissolved in solution, therefore, the filtrate is *not* to be discarded but only the crude resins

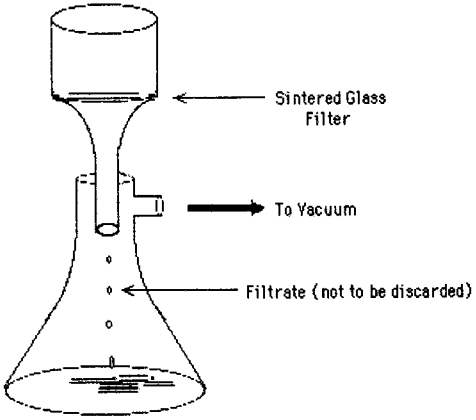


Figure 8. Apparatus set-up for peptide suction filtration.

Step 8: The resins are washed with four consecutive times of 5-mL of the leftover cleavage solution to ensure that all of the unbound peptide is in the filtrate

Step 9: The filtrate is transferred into a 200-mL round-bottom flask and is evaporated down to approximately 10-mL using a Büchi Rotavapor with a dry ice-acetone trap

Step 10: The peptide is precipitated out of solution by adding 125-mL of cold anhydrous ether in a drop-wise fashion through a 150-mL funnel while the solution is kept stirring

Step 11: The peptide-ether solution is kept in the refrigerator overnight to ensure that all the peptide precipitates out of solution

Step 12: The white peptide precipitate is removed via suction filtration through a glass sintered filter and is washed with three consecutive times of 5-mL of cold anhydrous ether

Step 13: The precipitate is removed from the sintered glass and is placed inside a vial for lyophilization (see Step 3 and 4)

Step 14: The white crystallized solid can be stored in the freezer for future analysis

Figure 9 illustrates the mechanism by which the peptide chain is cleaved from the polymer resin using TFA. Protonation occurs at the nitrogen on the linker. As a result of bond breakage, two compounds are formed, one of which is the cleaved peptide chain (**14**) and the other is an intermediate species (**15**). The intermediate species can be represented as several resonance structures such as **16** and **17**.

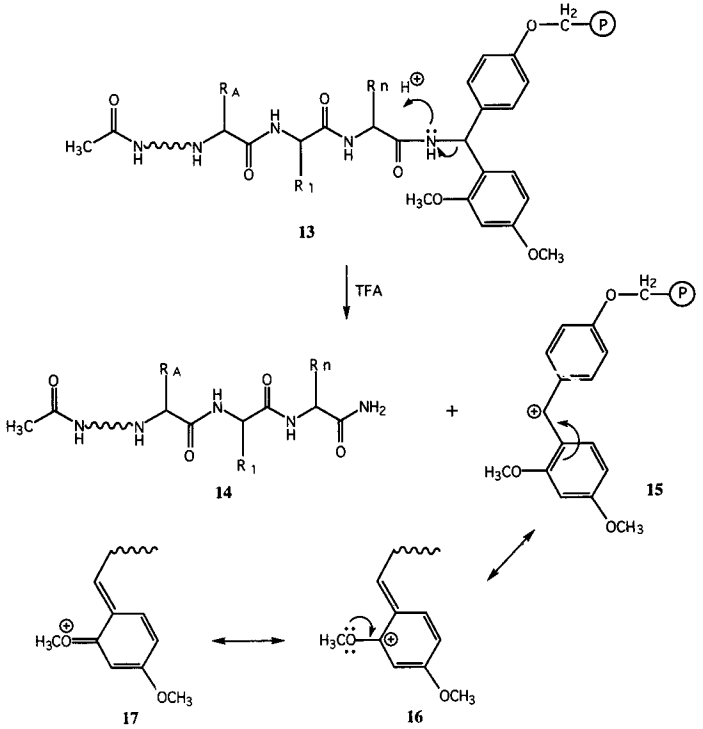


Figure 9. Cleavage of the peptide chain (**14**) off the Rink Amide linker (**15**) using TFA.[6]

Figure 10 illustrates the mechanism by which the *t*-butyl ester group on glutamic acid is cleaved using TFA. Protonation on the carbonyl oxygen produces intermediate **19**. Electron delocalization among intermediate **19** displaces a *t*-butyl group to produce a protonated form of glutamic acid (**20**). The *t*-butyl carbocation (**21**) can either form 2-methylpropene (**22**) as result of deprotonation or alkylate the strongly activated anisole and

thioanisole scavengers.

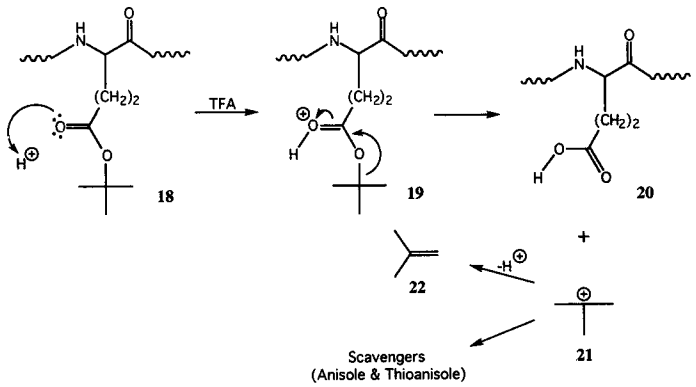


Figure 10. TFA deprotection of a *t*-butyl ester group on glutamic acid.[6]

Figure 11 illustrates the mechanism by which the Boc group on lysine is cleaved using TFA. The first step proceeds the same fashion as cleavage of a *t*-butyl ester group on glutamic acid. Protonation occurs at the carbonyl oxygen to form intermediate **24**. Electron delocalization among intermediate **24** generates compound **25** and a *t*-butyl carbocation (**26**) that further loses a hydrogen to form 2-methylpropene (**27**) or to be consumed by scavengers. In addition, decarboxylation of reactive compound **25** produces the protonated form of lysine (**28**) and a carbon dioxide molecule (**29**).

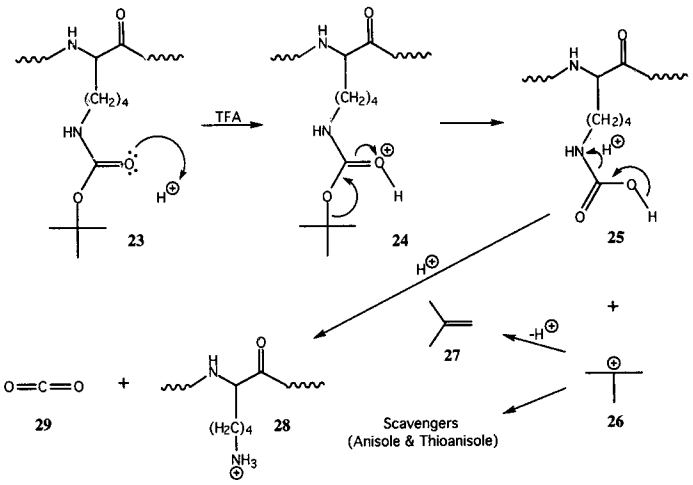


Figure 11. TFA deprotection of a Boc group on lysine.[6]

Figure 12 illustrates the mechanism by which the *t*-butyl ether group on tyrosine is cleaved using TFA. Protonation occurs at the ether oxygen to form a positive charged intermediate **31**. Electron delocalization among intermediate **31** kicks out a free-up tyrosine (**32**) and a *t*-butyl carbocation (**33**), which further loses a hydrogen to form 2-methylpropene (**34**) or react with the scavengers.

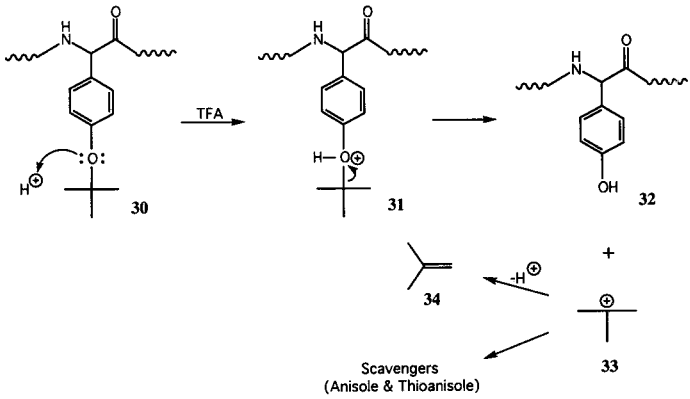


Figure 12. TFA deprotection of a *t*-butyl ether group on tyrosine.[6]

Figure 13 illustrates the mechanism by which the Pmc group on arginine is cleaved using TFA. The first step in the cleavage starts with protonation at one of the nitrogens to form intermediate **36**. Electron delocalization among intermediate **36** frees-up a protonated arginine (**37**) and a Pmc carbocation (**38**). The sulfonyl carbocation reacts either with anisole and thioanisole, or with trifluoroacetic anion to form species **39**. Electron delocalization among the ring of species **39** introduces a positive charge on the oxygen to form intermediate **40**. The addition of water on the carbonyl releases a 2,2,5,7,8-pentamethylchroman group (**41**), sulfur trioxide (**42**), and TFA (**43**).

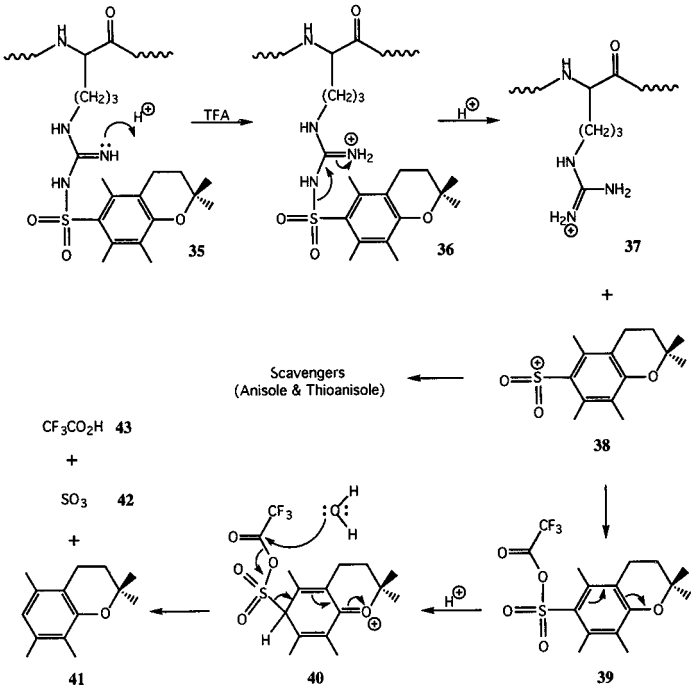


Figure 13. TFA deprotection of a Pmc group on arginine.[56]

Molecular Weight Determination

The purity of QEKR 4.0[b] and QEKR 5.0 peptides was analyzed and confirmed through MALDI-TOF at New York University. Approximately 3 mg of each peptide were sent for analysis.

Analysis of Peptides by HPLC

The QEKR 4.0[b] and QEKR 5.0 peptides were analyzed by reverse phase HPLC. The model of the HPLC instrument was a Hewlett Packard 1100 Series, equipped with a Windows PC workstation. Deionized water and acetonitrile were the two main solvents used in the purification. The filtered deionized water was poisoned with 2% acetonitrile to kill bacteria that might live in the reservoir environment. On the other hand, the acetonitrile contained 0.1% TFA to maintain slight acidic condition so that the peptides will remain in their protonated form during the purification process. The concentration of our peptide samples was 5 mM, dissolved in 2-mL of deionized water. Samples were eluted through a Zorbax® R_X-C8 (4.6 x 250 mm) reverse phase column with 5 µm packing using isocratic conditions at a flow rate of 1-mL/min over 10-25 minutes. However, samples eluted from the C8 column failed to show adequate resolution and separation among the base peaks. Therefore, a Zorbax® SB-C18 (4.6 x 250 mm) reverse phase column with 5 µm packing was used instead throughout the experiment for determining the retention time and for purification of the peptide of interest.

The variable UV detector was set at three different wavelengths (254 nm, 275 nm, and 288 nm) in order to detect the presence of aromatic compounds or other sensitive species that absorb strongly at the UV region of the electromagnetic spectrum. The reason for exploiting three different wavelengths is to discriminate the products of our interest to that of other interfering or undesirable materials that might have been eluted from the column. By employing isocratic runs, the retention time of the peaks will remain the same while only the intensity changes. Therefore, peaks that show a greater intensity than other peaks at a specific retention time can be distinguished based on a set wavelength.

In order to be able to analyze where the desired peaks lie, adequate separation between peaks must be met. The most favorable isocratic conditions for our samples were determined to be 70% H_2O and 30% acetonitrile eluted over a period of 10 minutes. Both of the peptide samples were run using these conditions. We flushed the column with 100%

acetonitrile to observe if there were other materials that had not been eluted from the column.

Tyrosine contains an aromatic benzene ring which absorbs maximum at 275 nm.[40] Since both QEKR 4.0[b] and QEKR 5.0 contain a tyrosine residue, the variable UV detector was set at that wavelength to monitor its concentration. Beside tyrosine, the UV detector was set at another different wavelength to detect other strongly absorbing species as well.

One such species that absorbs at the UV region is the aromatic ring of Pmc from blocked arginine. From mass spectral data, it was determined that the Pmc side chain group failed to be removed on the arginine residue for both peptides. Therefore, the presence of Pmc contained in the synthesized peptides might mask the absorbance of the peaks at 275 nm and might generate distortion in the peak intensity values. To be sure that Pmc can be detected under the variable UV detector, a cuvette containing approximately 30 mg of the blocked arginine in DMF was analyzed using a Hewlett Packard 8452A Diode Array Spectrophotometer. Blocked arginine was determined to absorb maximum at 288 nm. The peptide samples were run using the HPLC at that wavelength without changing the previous parameters. In addition, the wavelength of 254 nm was also used arbitrarily to compare differences between spectra taken at 275 nm and 288 nm.

Purification of Peptides by HPLC

Although the peak of interest for both QEKR 4.0[b] and QEKR 5.0 peptides showed resolution, the spectra failed to demonstrate enough separation of the peak of interest for peptide purification. Instead of isocratic runs, we switched to solvent programming to improve peak separations. The parameters for the solvent programming required a solvent range from 100% to 5% acetonitrile (0% to 95% H_2O) at time period from 0 to 10 minutes, respectively. The detector was set at wavelength of 275 nm to collect the filtrate at a flow rate of 1-mL/min using a C18 reverse phase column.

Since the retention time of peak of interest for both QEKR 4.0[b] and QEKR 5.0

peptides were approximately at 2.4 minutes, we started collecting the filtrate from 2.5 to 2.8 minutes. Approximately five collections of the filtrate were gathered and were lyophilized at -50°C to expel the solvent. The lyophilized samples were sent to New York University for mass spectral analysis to determine their purity.

To determine if the lyophilized samples were pure by HPLC, we again collected 10 samples of the filtrate and lyophilized them at -50°C. The lyophilized samples were dissolved in 500-μL of deionized water and injected into the HPLC using solvent programming.

Characterization of Peptides by NMR

NMR spectroscopy was used to study the ^1H-^1H coupling patterns and chemical shifts of each amino acid residue on the peptide structure. The model of the NMR instrument was a Varian Gemini-200 MHz, equipped with a Sun SPARC workstation. Before the samples were submitted for NMR analysis, the N-H and -OH groups of the peptides had to be exchanged with D_2O. Deuterium does not give signals under the conditions of proton NMR spectroscopy. Therefore, the extent to which the peaks and coupling patterns formed would eliminate some difficulties in interpreting the result between hydrogen and deuterium exchange. The procedures for this part are listed as follows:

Step 1: 30 mg of each QEKR 4.0[b] and QEKR 5.0 peptides are dissolved in 1-mL of D_2O

Step 2: Each sample is frozen with dry ice and lyophilized at -50°C overnight

Step 3: Repeat Step 1 and 2 again

Step 4: Each lyophilized sample is dissolved with 1-mL of D_2O for NMR analysis

Correlated spectroscopy (COSY), a procedure that gives a 2-D spectrum of ^1H-^1H

connectivities, was also performed in both QEKR 4.0[b] and QEKR 5.0 peptides. Proton spectrum of COSY appears along the diagonal of the contours that correlate to peaks to the off-diagonal contours called crosspeaks. A horizontal line drawn from the crosspeak will intercept a contour on the diagonal which is correlated to another contour on the diagonal drawn vertically from the same crosspeak. Correlation or coupling among the ^1H-^1H connectivities can also give some useful information in solving the proton chemical shifts of our peptides.

In addition to the above procedures, we analyzed spectra of individual unblocked amino acids (Ala, Arg, Gln, Glu, Gly, Lys, and Tyr) so that the result could help us further piece together information in characterizing and assigning peaks to both our peptide spectra. For each preparation, 50 mg of amino acid residues were dissolved in 1.5-mL of D_2O for NMR analysis.

Molecular Modeling

The α-helical geometry of our synthesized peptides, QEKR 4.0[b] and QEKR 5.0, were visualized using molecular modeling. The computational studies of our minimized 3-D constructed peptides were built using MacroModel Version 5.0 program on the Silicon Graphics workstation. The program used modified AMBER force field parameters. The constructed peptides were minimized in environment either with or without H_2O as solvent. The minimized structures were manipulated and printed using CAChe translated from a .pdb file to inspect salt bridge interactions and to verify the α-helical conformation of both peptides. The minimized energy, electrostatic, and other quantum mechanical parameters were calculated. The procedures for this part are listed as follows:

Step 1: Enter MacroModel on the SGI

Step 2: Once inside MacroModel, select the <PEPTID> and <GROW> options

Step 3: To grow the peptide in an α-helical conformation, select <Conf> and then click <Alpha Helix>

Step 4: Grow the peptide to the desired length by selecting the appropriate amino acid options

Step 5: After finish growing the peptides, the N- and C-termini are capped with an acetyl and -NH_2 group, respectively by selecting the appropriate carbon, oxygen, and nitrogen options

Step 6: To minimize the peptide structures, select <ENRGY> and then click <AMBER>, <MINIM>, and <PRCG>

Step 7: To select the solvent, click <SLVNT>

Step 8: To run, select <Start>

Step 9: To save the file as .pdb, select <WRITE> and then click <PDB>

Step 10: To transfer the .pdb file from the SGI to CAChe, open the Fetch 3.0.1 program found on the hard drive of the Power Macintosh

Step 11: Type chandler as the host, your user ID, and password

Step 12: Once inside chandler, save the .pdb file into the hard drive by clicking <get file>

Step 13: To translate the .pdb file into .mol file, open CAChe Translators

Step 14: Once inside CAChe Translators, select the .pdb file and then click <Brookhaven> as input and <CAChe> as output for translation

Step 15: Open CAChe Editor to manipulate and view the peptide structure

Titration of QEKR 4.0^b by NMR

Spectra of both titrated peptides were taken by NMR at various measured pH units. We used the same samples (30 mg of samples dissolved in 1-mL of D_2O) from the previous NMR experiment to perform the titration. Measurements were monitored at each pH unit from pH 10 to pH 1 using a pH microelectrode with an internal calomel reference. The starting samples were added with enough of the lyophilized deuterated DPO_4^{2-}

(approximately 50-75 mg) to adjust a basic pH of about 10. The pH of the NMR samples was lowered to each unit by adding varying microliter amounts of 12 M DCl in a 1 to 10 dilution. The total amounts of 12 M DCl (1:10 dilution) added are listed below in Table 3.

Table 3. Amounts of 12 M DCl (1:10 dilution) added.

pH	9.37	9.09	8.07	7.98	7.06	5.84	3.98	3.08	2.09	1.20
DCl (μL)	0	9	140	150	390	510	540	590	770	920

CHAPTER 3: RESULTS AND DISCUSSION

Percent Yields of QEKR 4.0^b and QEKR 5.0

The syntheses of QEKR 4.0^b and QEKR 5.0 afford a yield of approximately 174 mg and 177 mg, respectively. Both peptides have similar yields of 38% (Table 4).

Table 4. Percent yields of the synthesized peptides.

Peptide	Mass Collected (mg)	Theoretical Weight (mg)	Percent Yield
QEKR 4.0^b	174	464	38
QΞKR 5.0	177	464	38

Confirmation of QEKR 4.0^b and QEKR 5.0 by MALDI-MS

Mass to charge (*m/z*) values for both peptides were obtained from MALDI spectra (Figure 14 and 15). All peaks above the *m/z* of 1500 for both QEKR 4.0^b and QEKR 5.0 had similar values. Table 5 shows 65% of each sample contained the peptide of interest at the observed *m/z* of 1791.2, similar to the calculated *m/z* of 1791.0. The protonated molecular parent ion (1791.2) was probably the result of proton transfer to the analyte (M) from the matrix during electron impact. All *m/z* values that are shown in Table 5 indicate protonation of the analyte.

The other 35% were impurities generated either from incomplete coupling of amino acids onto the growing peptide chain or from incomplete cleavage of side chain protecting groups on the residues. The failure of Gln and Ala to couple efficiently onto both peptides was indicated by an observed *m/z* of 1591.8, the same as the calculated *m/z* of 1591.8.

Lys (calculated *m/z* of 1590.7) or Glu (calculated *m/z* of 1590.8) could be residues which failed to couple onto the peptide instead of Gln because of their similar calculated *m/z* values when compared to the observed *m/z* values. However, the chance for Lys or Glu to be residues that were lost was slim because there were more possibilities for 3 Gln to be lost for a single Lys or Glu residue.

An observed *m/z* of 1663.8 and 1662.6 was due to Gln failing to couple onto the growing peptide chain during elongation. The observed *m/z* of 1662.6 from QEKR 5.0 was similar to the calculated *m/z* of 1662.9, but it was about 1 mass unit different compared to the other observed *m/z* of 1663.8 from QEKR 4.0[b]. The difference might be contributed by the relative amount of isotope ^{13}C on the carbon-based peptide. Again, the possibility for Lys (calculated *m/z* of 1661.8) or Glu (calculated *m/z* of 1661.9) to be residues that were lost was highly unlikely.

An observed *m/z* of 1721.3 and 1720.1 suggests that Ala failed to couple onto the peptide. The discrepancy between the observed *m/z* of 1721.3 and the calculated *m/z* of 1719.9 was probably due to the difference in the amount of isotope ^{13}C. The lose of Ala was prevalent because the peptide consisted the majority of Ala residues.

Finally, the observed *m/z* of 2060.0 and 2058.7 indicates Pmc failed to be cleaved off on the Arg residue. Both observed *m/z* values were similar to the calculated *m/z* of 2058.4. Slight variation in mass units might be caused by isotopic abundance of ^{13}C atoms. Some suggestions to improve upon the conditions for cleavage of Pmc protecting group on Arg are either increase the concentration of TFA from 90% to 95% or lengthen the time from 2 to 3 hours for the cleavage procedure.

Table 5. Mass to charge (*m/z*) values of QEKR 4.0[b] and QEKR 5.0.

QEKR 4.0[b]			QEKR 5.0		
m/z	**Percent Purity**	**Species Lost or Added**	***m/z***	**Percent Purity**	**Species Lost or Added**
1591.8	7	-Gln, -Ala and $+H^+$	1591.8	6	-Gln, -Ala, and $+H^+$
1663.8	14	-Gln and $+H^+$	1662.6	11	-Gln and $+H^+$
1721.3	8	-Ala and $+H^+$	1720.1	8	-Ala and $+H^+$
1791.2	64	$+H^+$	1791.2	65	$+H^+$
2060.0	7	+Pmc and $+H^+$	2058.7	10	+Pmc and $+H^+$

Note: The lose of Ala and Gln and the failure of Pmc to be cleaved on the peptide were due to the reaction and not caused by fragmentation from electron impact. Only proton transfer to the analyte (M) accounts for the electron impact.

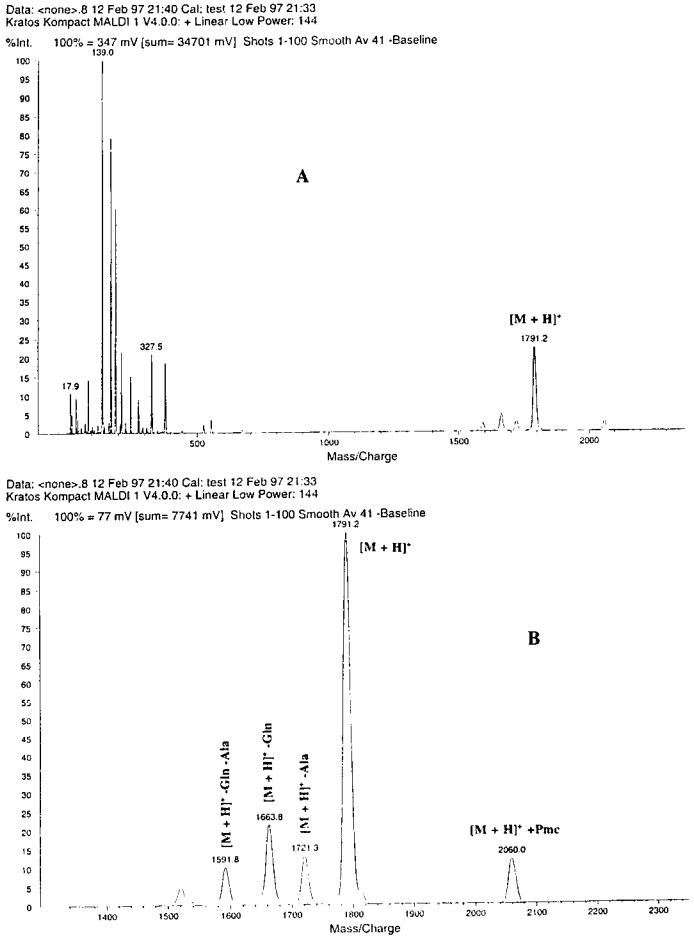


Figure 14. MALDI-MS of QEKR 4.0[b]: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.

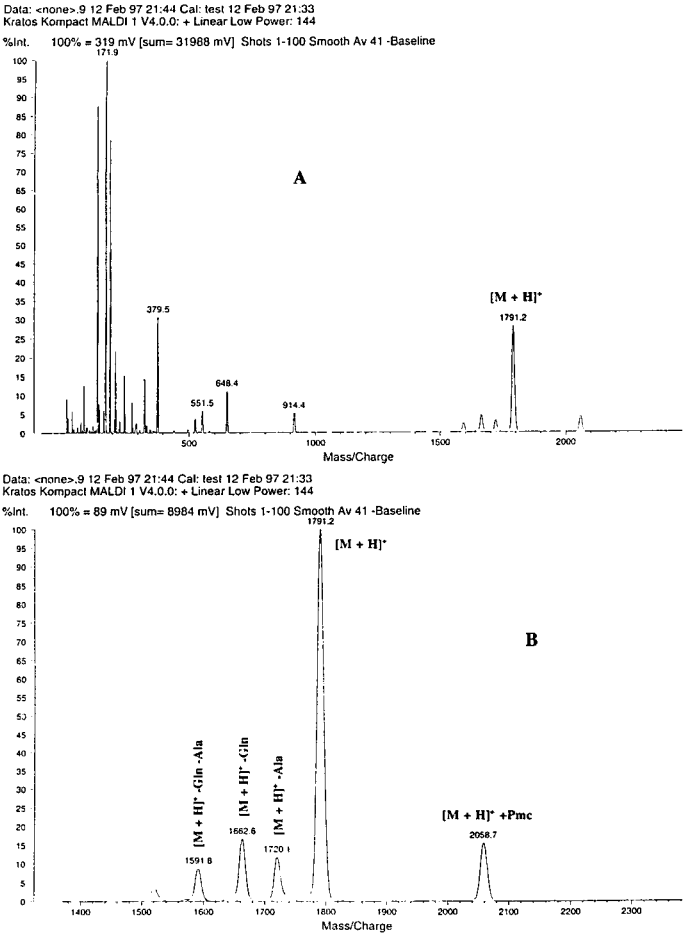


Figure 15. MALDI-MS of QEKR 5.0: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.

Determination of Retention Times and Purification of QEKR 4.0^b and QEKR 5.0 by HPLC

Isocratic conditions were used to determine the retention time for both QEKR 4.0^b and QEKR 5.0. The peak of interest was marked at the retention time of approximately 2.5 minutes for all three wavelengths that were used to follow the runs (Figure 16 and 17). Spectrum B in Figure 16 and 17 indicates that the product peak absorbs more than 3 times as intensely as in spectrum C and twice as intensely as in spectrum A. The strong absorption at 275 nm suggests that tyrosine must be present in the product peak. Although we were able to determine the product peak, we could not collect the fractions until impurities beyond the retention time of 2.5 minutes were well separated from the product.

The purification of the isolated product was performed under solvent programming. The retention time of the purified peptide for both QEKR 4.0^b and QEKR 5.0 was marked at approximately 2.4 minutes (Figure 18A and 19A). A second purification procedure was acquired to confirm if the collected fractions contained pure products of QEKR 4.0^b and QEKR 5.0. The retention time for both pure peptides was again marked at 2.4 minutes, indicating the products were present but in low concentration, about a 6-fold decrease (Figure 18B and 19B). Although most impurities beyond the retention time of 2.4 minutes were removed on spectrum B when compared to spectrum A (Figure 18 and 19), other residual materials could be seen at the retention time between 2.5 and 2.6 minutes.

MALDI-MS was used to confirm if the purified fractions contained the QEKR 4.0^b and QEKR 5.0 products. The MALDI spectrum of QEKR 4.0^b on Figure 20B shows three major peaks at *m/z* of 1788.0, 1814.5, and 1836.1. The observed *m/z* of 1788.0 indicates the peptide of interest, but the *m/z* value was 3 mass units away from the calculated *m/z* of 1791.0. The difference might be due to some mechanical problems of the MALDI mass spectrometer.

An observed *m/z* of 1814.5 indicates a sodium ion was transferred to the analyte (M) during electron impact. The observed *m/z* was similar to the calculated *m/z* of 1813.0.

Any variations in mass units were contributed either by problems with the mass spectrometer or by isotope differences. A major discrepancy was noticed when we assigned the peak of *m/z* of 1836.1. Although the mass of the observed peak was similar to the calculated value of 1836.0, the transfer of two positive charged sodium ions to the analyte (M) would make the *m/z* to be half as much for the observed (918.05). The observed *m/z* value violated the assumption of the assignment.

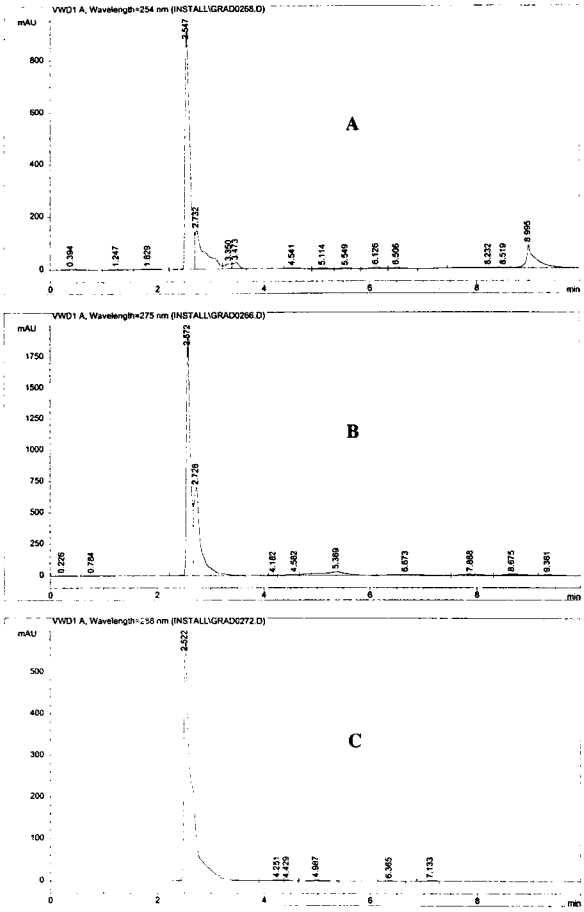


Figure 15. Isocratic runs (70% H_2O:30% acetonitrile) of QEKR 4.0[b]: A) 254 nm; B) 275 nm; C) 288 nm.

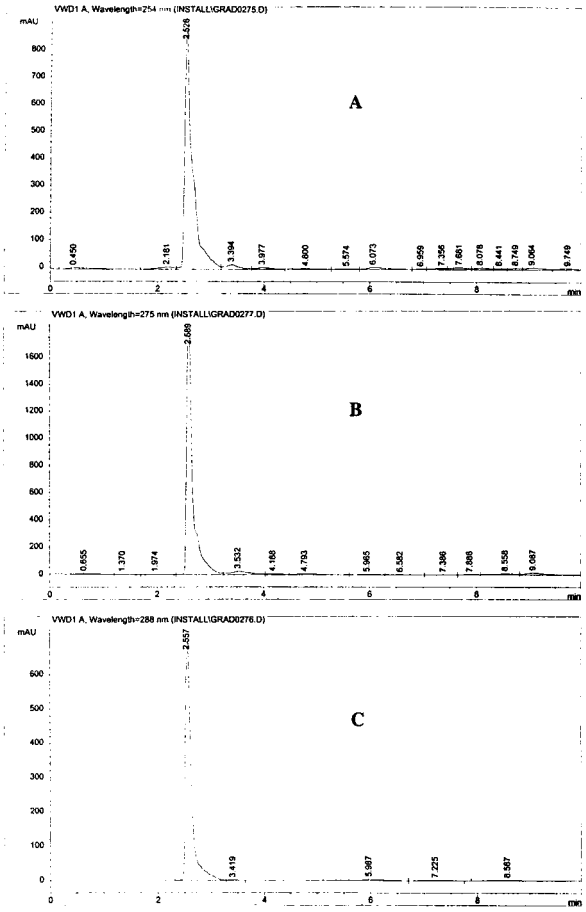


Figure 17. Isocratic runs (70% H_2O:30% acetonitrile) of QEKR 5.0: A) 254 nm; B) 275 nm; C) 288 nm.

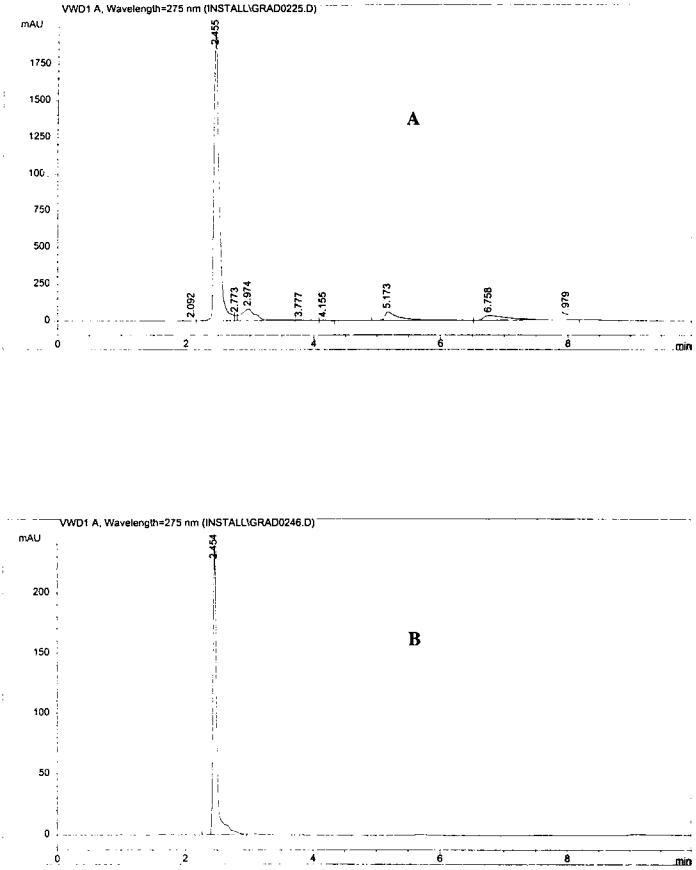


Figure 18. Solvent programming of QEKR 4.0[b] at 275 nm: A) 1st purification; B) 2nd purification.

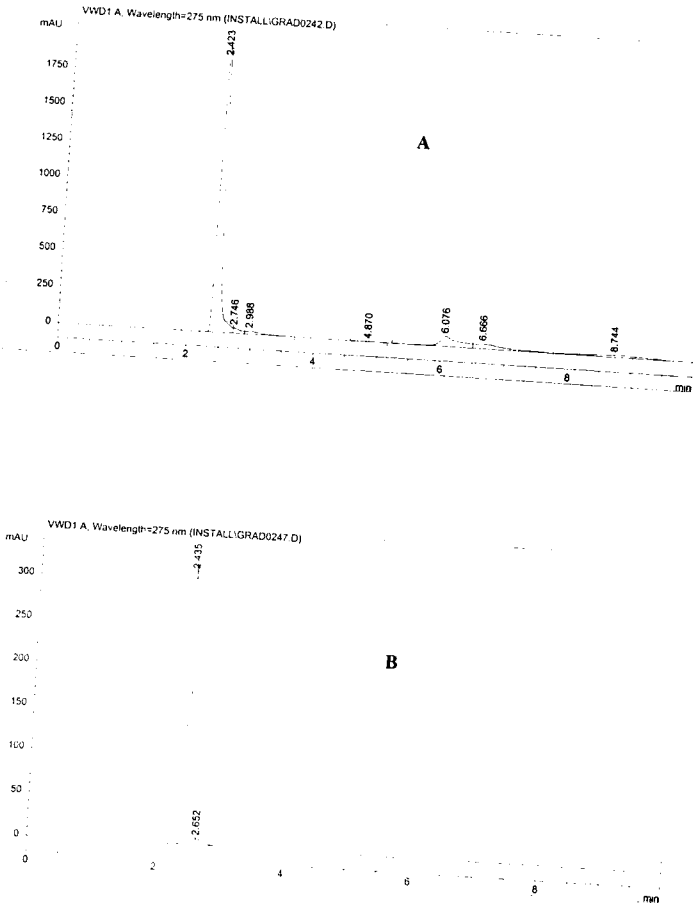


Figure 19. Solvent programming of QEKR 5.0 at 275 nm: A) 1st purification; B) 2nd purification.

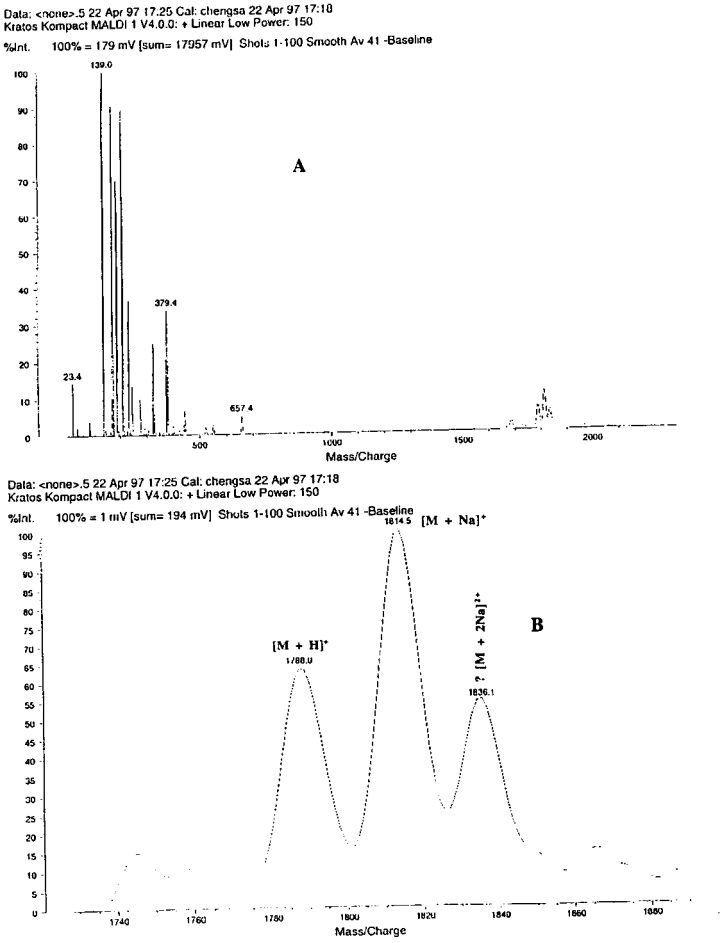


Figure 20. MALDI-MS of purified QEKR 4.0[b]: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.

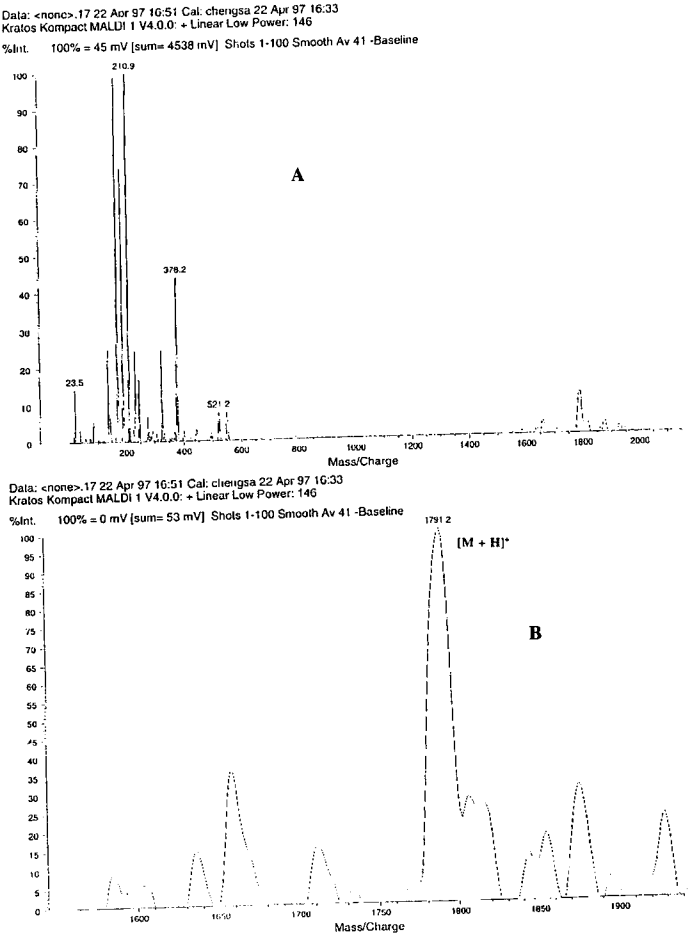


Figure 21. MALDI-MS of purified QEKR 5.0: A) overall view; B) blow-up view. Peaks below *m/z* of 600 correspond to the matrix.

^{1}H-NMR Spectra of Individual Amino Acids

Table 6 shows the relative assignments for proton chemical shifts of individual unblocked amino acids. Side band enhancement were prevalent in all amino acid spectra (Figure 22-28). Chemical shifts and splitting patterns of amino acids on the 200 MHz ^{1}H-NMR (D_2O) are as follows: Gly δ 3.3 (singlet); Ala δ 3.6 (quartet), 1.2 (doublet); Glu δ 3.6 (triplet), 1.9 (quartet), 2.3 (triplet); Gln δ 3.6 (triplet), 1.9 (quartet), 2.3 (triplet); Lys δ 3.6 (triplet), 1.7 (quartet), 1.2 (quintet), 1.5 (quintet), 2.8 (triplet); and Tyr δ 3.7 (singlet), 2.8 (octet), 6.7 (doublet), 7.0 (doublet). In general, the splitting patterns on the majority of peaks in Gly, Ala, Glu, Gln, Lys, and Tyr were well-behaved except Arg. The Arg spectrum (Figure 26) rendered poor resolution and provided no splitting of the α (3.0), β (1.4), γ (1.4), and δ (3.0) hydrogens.

Table 6. Proton chemical shifts of individual unblocked amino acids. All spectra contain a water peak at 4.7 ppm.

Amino Acid Residue	α-H (ppm)	β-H (ppm)	γ-H (ppm)	δ-H (ppm)	ε-H (ppm)	Aromatic (ppm)
Gly	singlet 3.3	-----	-----	-----	-----	-----
Ala	quartet 3.6	doublet 1.2	-----	-----	-----	-----
Glu	triplet 3.6	quartet 1.9	triplet 2.3	-----	-----	-----
Gln	triplet 3.6	quartet 1.9	triplet 2.3	-----	-----	-----
Arg	3.0	1.4	1.4	3.0	-----	-----
Lys	triplet 3.6	quartet 1.7	quintet 1.2	quintet 1.5	triplet 2.8	-----
Tyr	quartet 3.7	octet 2.8	-----	-----	-----	doublet/ doublet 6.7/7.0

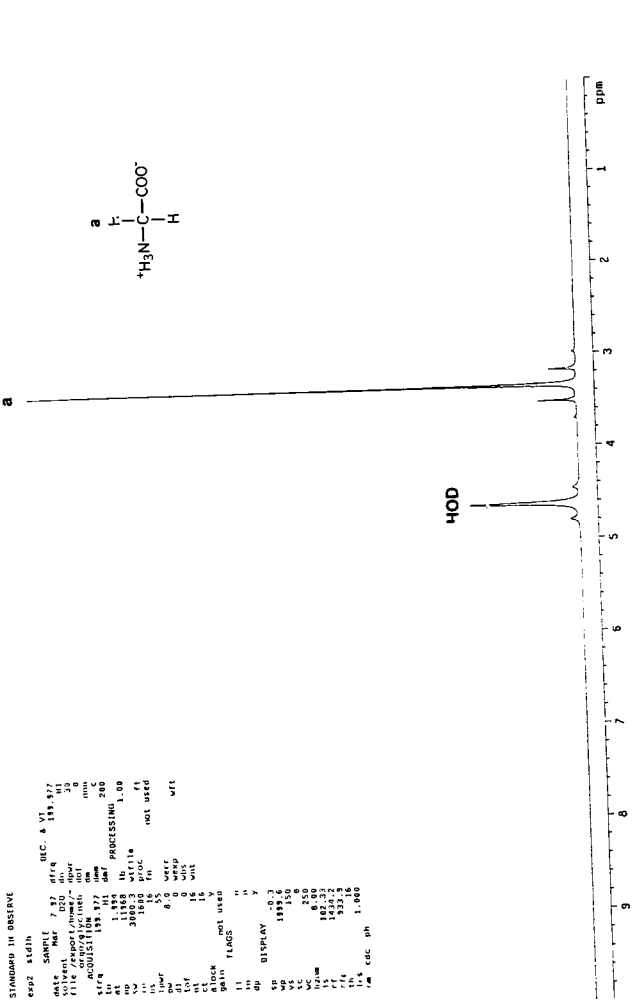


Figure 22. 1-D ^{1}H-NMR spectrum of glycine.

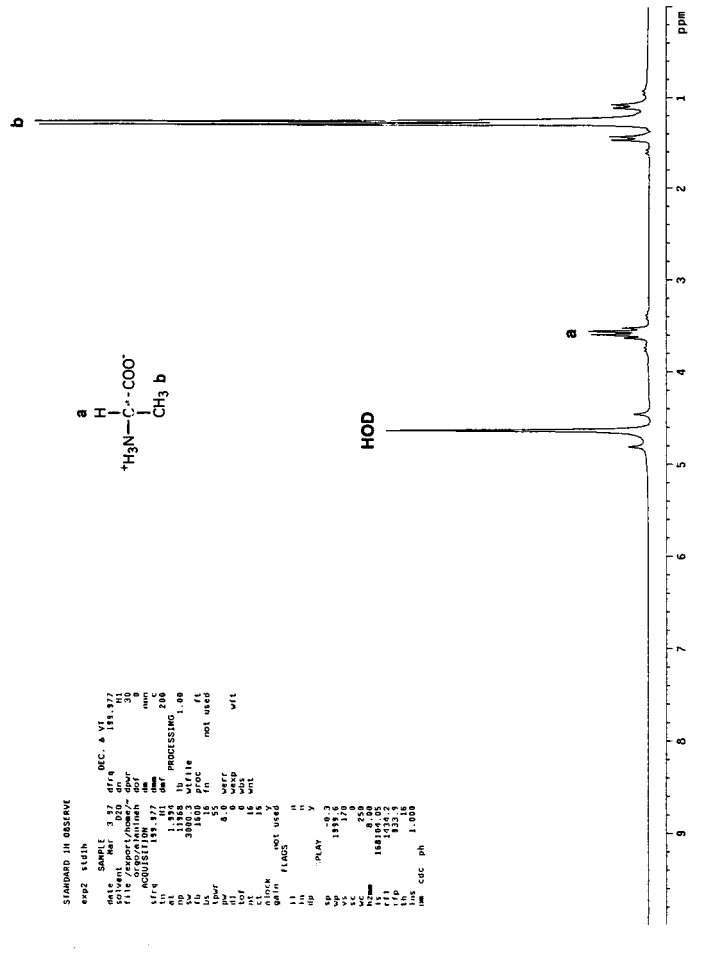


Figure 23. 1-D ^{1}H-NMR spectrum of alanine.

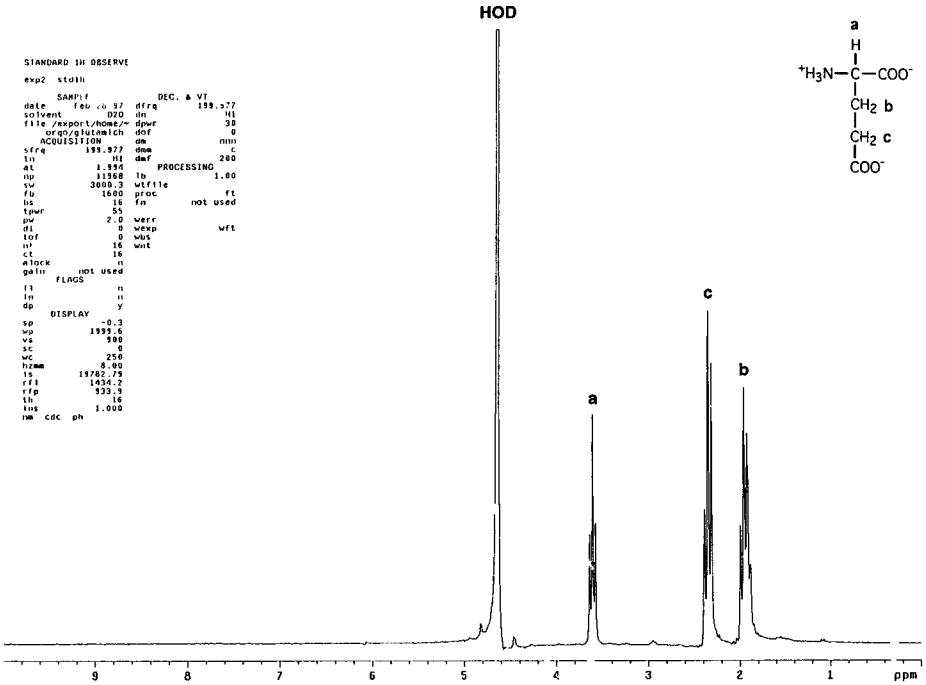


Figure 24. 1-D ^{1}H-NMR spectrum of glutaric acid.

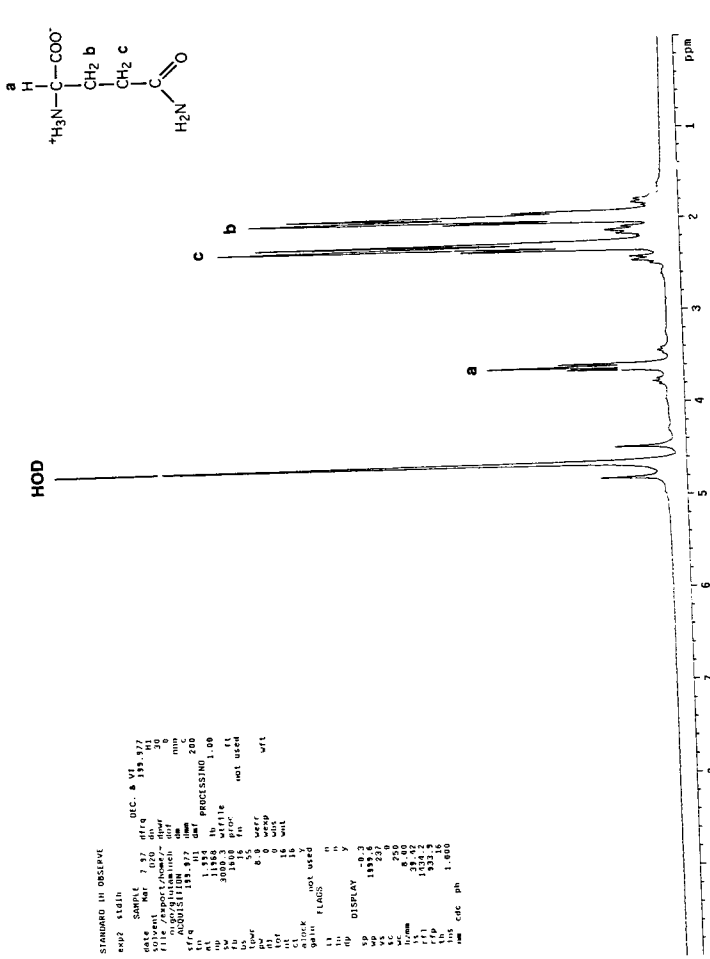


Figure 25. 1-D ^{1}H-NMR spectrum of glutamine.

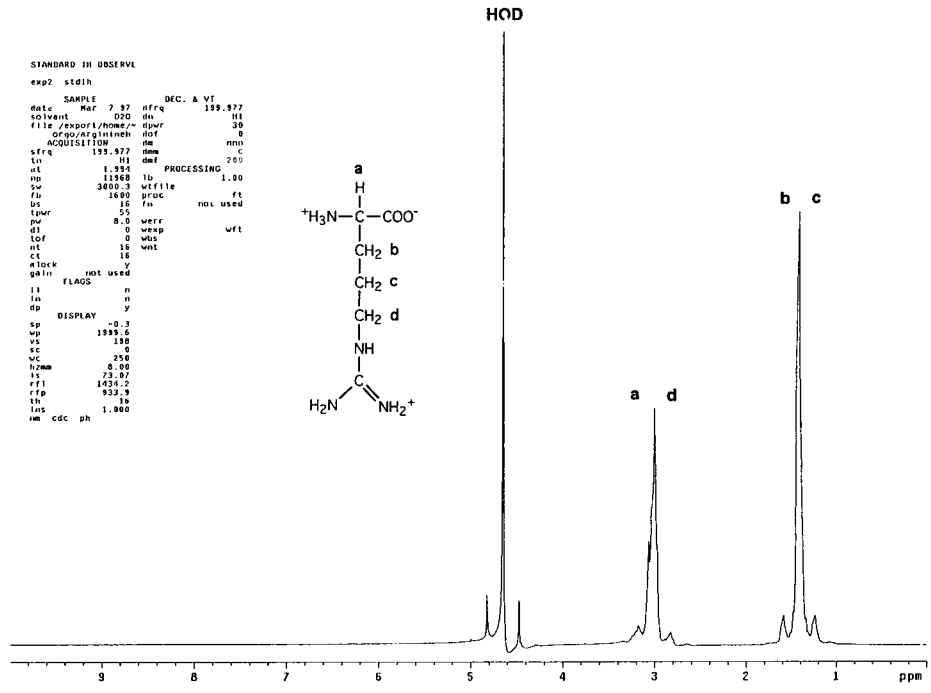


Figure 26. 1-D ^{1}H-NMR spectrum of arginine.

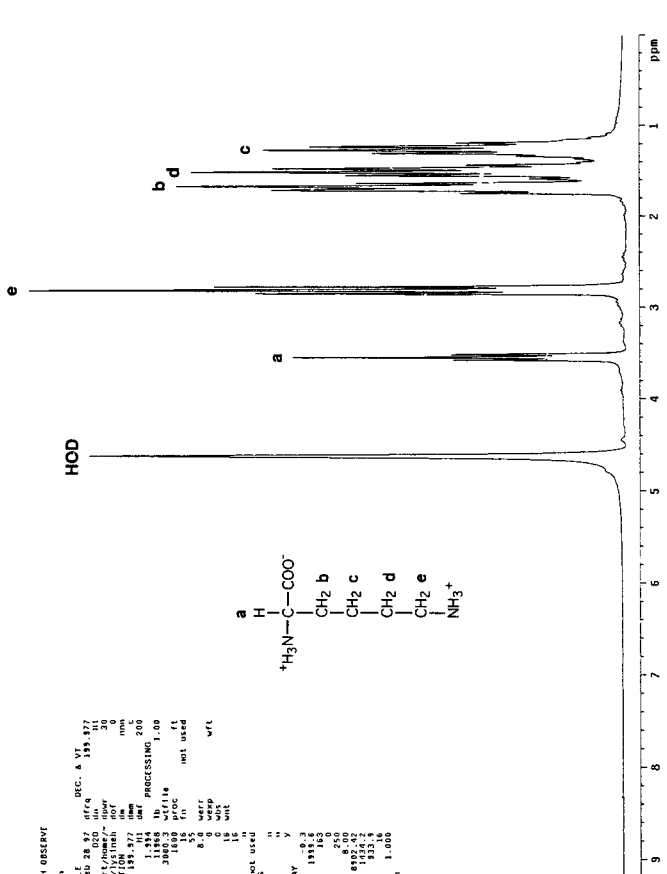


Figure 27. 1-D ^{1}H-NMR spectrum of lysine.

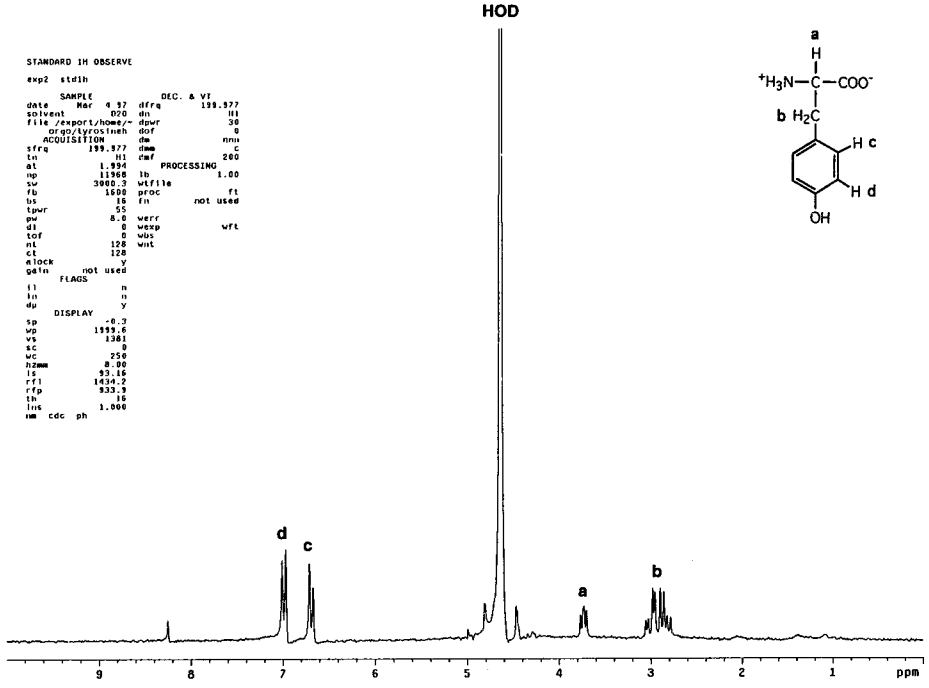


Figure 28. 1-D ^{1}H-NMR spectrum of tyrosine.

1-D ^{1}H-NMR and 2-D COSY Spectra of QEKR 4.0^b and QEKR 5.0

1-D ^{1}H-NMR and 2-D COSY spectra of both QEKR 4.0^b and QEKR 5.0 resembled each other (Figure 29-32). Although some residue assignments were ambiguous, important chemical shifts of visible and obvious couplings were noted. Two doublets from coupling on the aromatic hydrogens of tyrosine were marked at 6.7 ppm and 7.0 ppm. However, the β-hydrogen on tyrosine was unable to be assigned. A broad α-hydrogen peak was noted at 4.0 ppm, which was probably the result of the accumulation of all the α-hydrogens contributed by each residue on the peptide. Interestingly, chemical shifts of α-hydrogens on the peptide were different from expected chemical shifts of individual amino acids. An explanation might be that there were amide linkages along the peptide chain which influenced α-hydrogens on the residues to experience different chemical shifts as opposed to zwitterions in the individual amino acids. Similar to α-hydrogen, a broad and intense β-hydrogen (Ala) peak was marked at 1.2. COSY spectra (Figure 31 and 32) indicate well correlation in ^{1}H-^{1}H connectivity between the α- and β-hydrogens of Ala.

The assignments of Glu was partly determined by titration studies. The crucial γ-hydrogen was shown to move from 2.1 ppm to 2.3 ppm (Figure 37), going from high to low pH, respectively. The change in chemical shift implied that the γ-hydrogen was probably experiencing a different environment due to deprotona ion and protonation of the carboxyl group on the chain group of Glu when DPO_4^{2-} and DCl were added, respectively. COSY spectra showed well correlation in ^{1}H-^{1}H connectivity between the α- and β-hydrogens of Glu, and also between the β- and γ-hydrogens. Gln had similar proton chemical shifts as Glu except for the γ-hydrogen marked at 2.2 ppm.

The hydrogens on the alkyl side chain of Lys were partially assigned to some peaks on the peptide spectra via looking for comparison at the couplings in the Lys spectrum (Figure 27). The β-hydrogen was assigned at 1.6 ppm which correlated with the α-hydrogen in the COSY. However, the γ-hydrogen of Lys was unable to be assigned.

COSY did not indicate well correlation between the β- and γ-hydrogens or between the γ- and δ-hydrogens of Lys. There was correlation in ^1H-^1H connectivity between the δ- (1.4 ppm) and ε-hydrogens (2.8 ppm) of Lys in the COSY spectra. The assignments for Arg were tentatively labeled based on the peaks noted on the Arg spectrum (Figure 26) when compared to relevant peaks on the peptide spectra. COSY spectra revealed correlation in ^1H-^1H connectivity between the β- (1.4 ppm) and δ-hydrogens (3.0 ppm) and between the γ- (1.4 ppm) and δ-hydrogens (3.0 ppm) of Arg.

Table 7. Proton chemical shifts of QEKR 4.0[b]. The spectrum contains a water peak at 4.7 ppm. (Note: Only 1 set of data is presented because QEKR 5.0 also have nearly identical proton chemical shifts.)

Peptide Residue	α-H (ppm)	β-H (ppm)	γ-H (ppm)	δ-H (ppm)	ε-H (ppm)	Aromatic (ppm)
Gly	4.0	-----	-----	-----	-----	-----
Ala	4.0	1.2	-----	-----	-----	-----
Glu	4.0	1.9	2.3	-----	-----	-----
Gln	4.0	1.9	2.2	-----	-----	-----
Arg	4.0	1.4	1.4	3.0	-----	-----
Lys	4.0	1.6	?	1.4	2.8	-----
Tyr	4.0	?	-----	-----	-----	doublet/ doublet 6.7/7.0

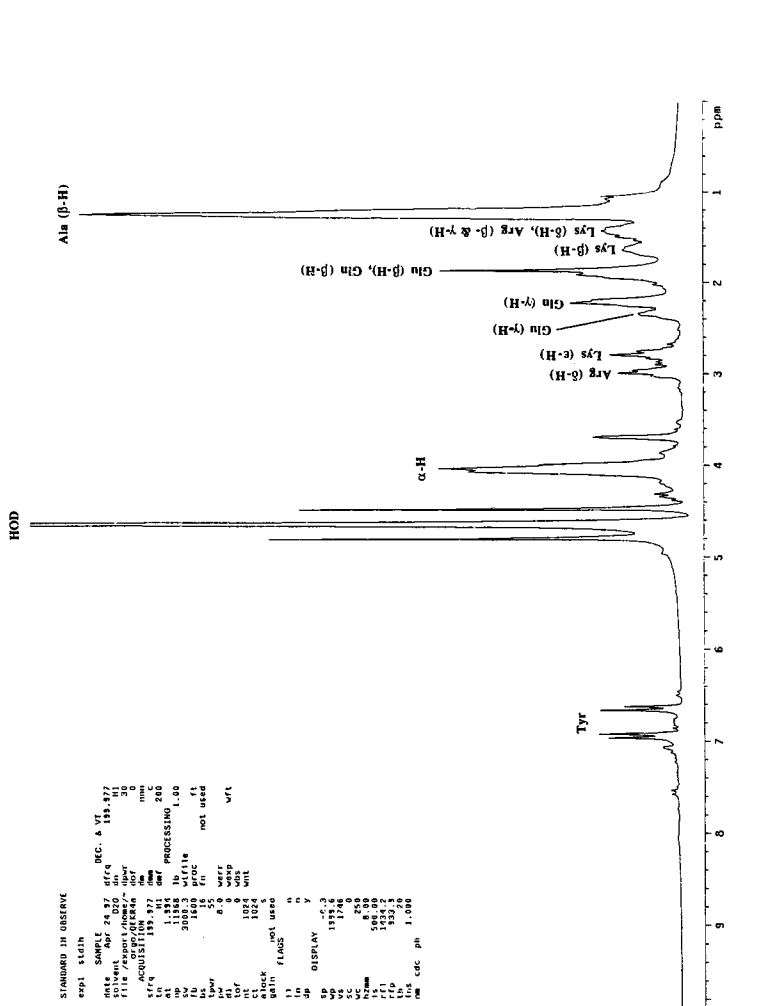


Figure 29. 1-D ^{1}H-NMR spectrum of QEKR 4.0[b].

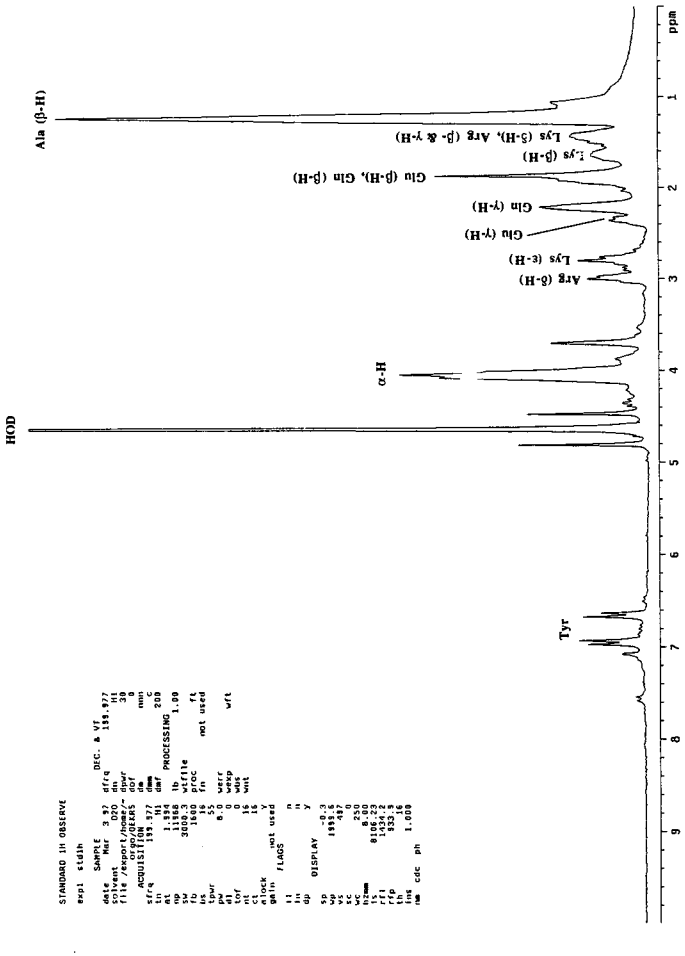


Figure 30. 1-D ^{1}H-NMR spectrum of QEKR 5.0.

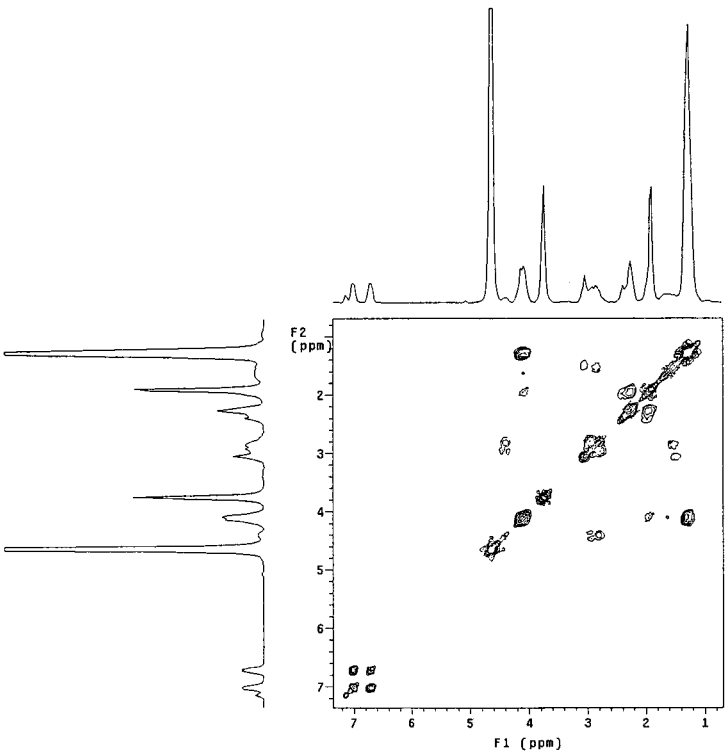


Figure 31. 2-D COSY NMR spectrum of QEKR 4.0[b].

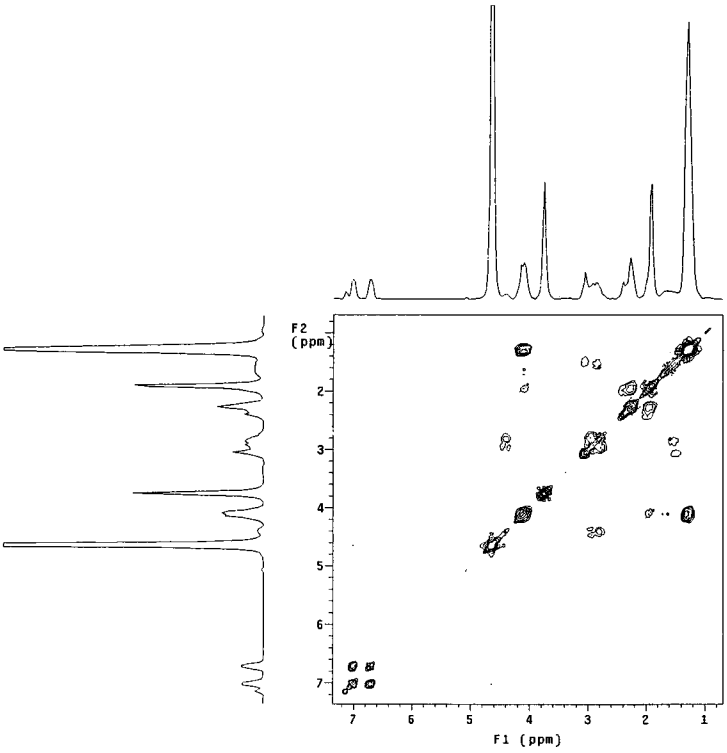


Figure 32. 2-D COSY NMR spectrum of QEKR 5.0.

Minimization of QEKR 4.0^b and QEKR 5.0

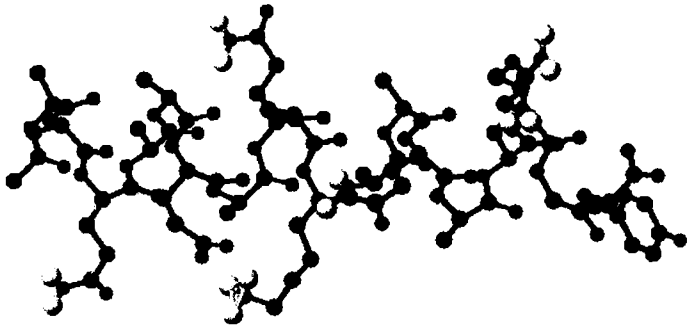
The minimized structures of both QEKR 4.0^b and QEKR 5.0 were viewed and generated from CAChe. Side and barrel views of both peptide structures were relatively similar when minimized either under no solvent or H_2O solvent conditions (Figure 33-36). Salt bridge formation could be seen between Glu and Lys in QEKR 4.0^b and not in QEKR 5.0. The distance of the salt bridge under no solvent and water solvent conditions was the same (2.55 Å). The distance of the non-salt bridge under no solvent (16.20 Å) and water solvent (15.65 Å) conditions was similar. However, the distance of non-salt bridge was about 6.5 times farther than the distance of the salt bridge.

The minimized energies of QEKR 4.0^b and QEKR 5.0 under water solvent conditions were -3243.07 kJ/mol and -3232.80 kJ/mol, respectively (Table 8). However, under no solvent conditions, the minimized energies were more than 4 times greater in magnitude for both QEKR 4.0^b (-848.10 kJ/mol) and QEKR 5.0 (-708.83 kJ/mol) when compared to helices under water solvent conditions. This large change in energy suggests that water might be participating in hydrogen bonding which stabilized the peptide structures.

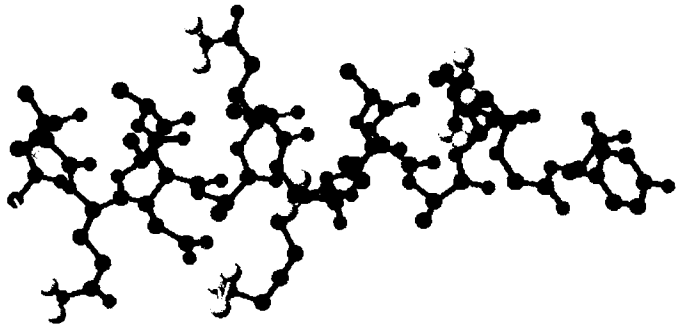
The electrostatic energy of QEKR 4.0 under both water solvent (-3165.77) and no solvent (-1364.33) conditions were lower than the electrostatic energy of QEKR 5.0 under water solvent (-2843.65) and no solvent (-1193.02) conditions, respectively. One reason might be that there were fewer repulsions in QEKR 4.0^b because the negative charged Glu and the positive charged Lys already had formed stable electrostatic interaction as opposed to no salt bridge interaction in QEKR 5.0.

Table 8. The minimized energies and electrostatic values for both QEKR 4.0[b] and QEKR 5.0 peptides were obtained in either water solvent or no solvent.

	QEKR 4.0[b]		QEKR 5.0	
	Water Solvent	**No Solvent**	**Water Solvent**	**No Solvent**
Energy (kJ/mol)	-3243.07	-848.10	-3232.80	-708.83
Electrostatic	-3165.77	-1364.33	-2843.65	-1193.02

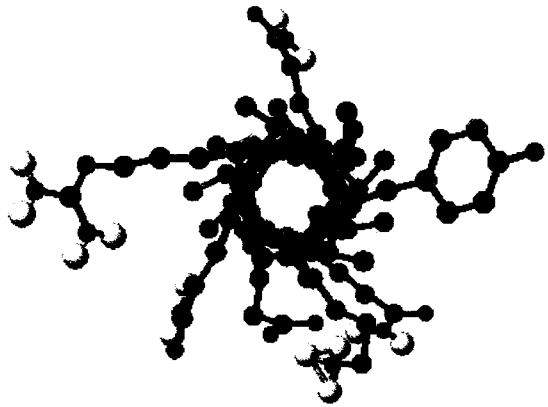


A

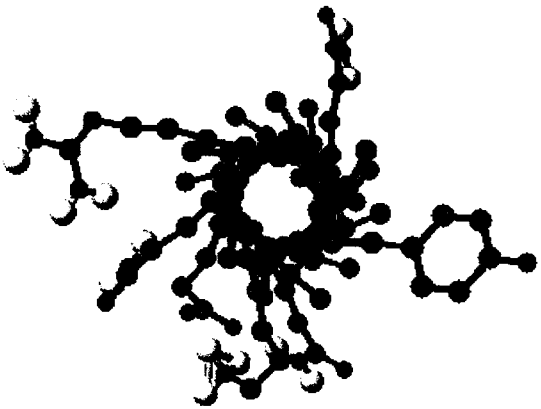


B

Figure 33. Side views of the minimized structures of QEKR 4.0°: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen.



A



B

Figure 34. Barrel views of the minimized structures of QEKR 4.0[b]: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen.

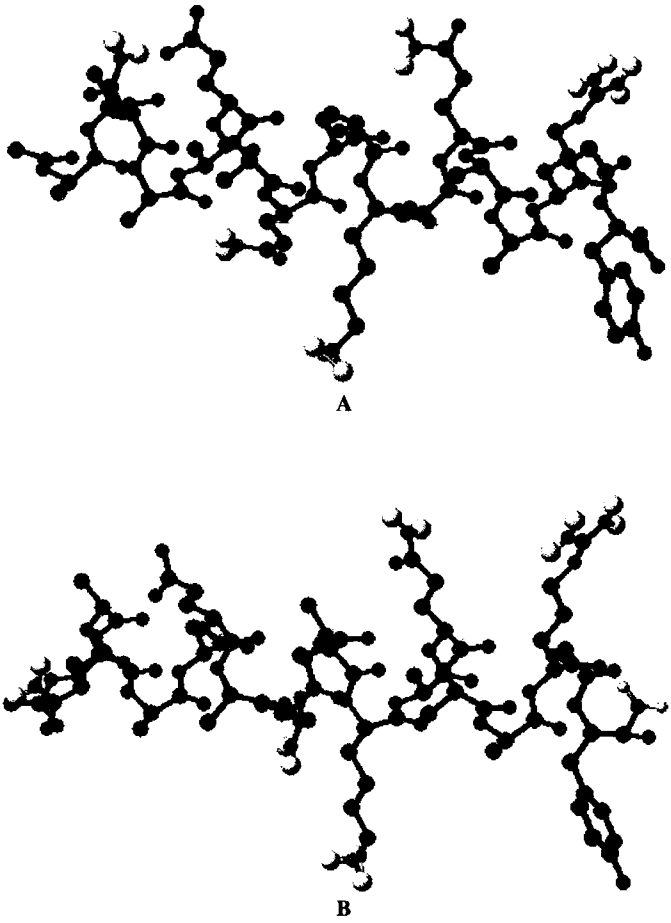


Figure 35. Side views of the minimized structures of QEKR 5.0: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen.

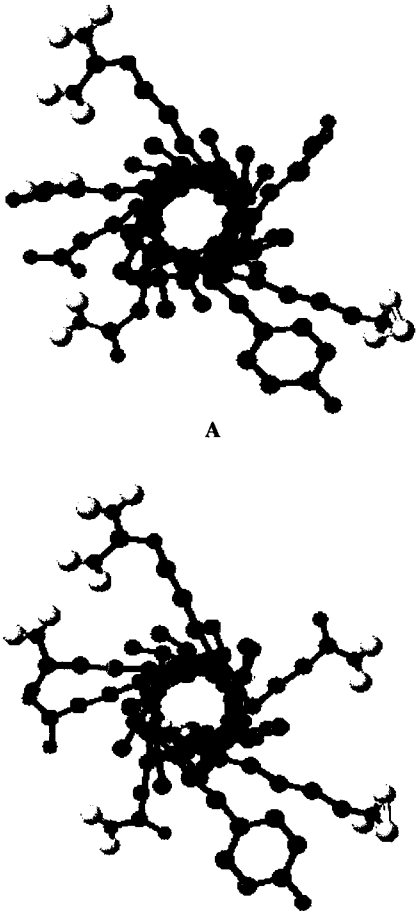


Figure 36. Barrel views of the minimized structures of QEKR 5.0: A) water solvent; B) no solvent. Blue: nitrogen; red: oxygen; black: carbon; and white: hydrogen.

Determination of the pKa of Glutamic Acid in QEKR 4.0^b

Titration of QEKR 4.0^b was performed to determine the pKa of glutamic acid. We followed the proton shifts of various peaks between 1.2 ppm and 4.4 ppm because the β- and γ-hydrogens of glutamic acid were sensitive to changes in that region. Nine peaks were annotated with chemical shifts at different pH (Table 9). Of the nine peaks, two (H7 and H8) were selected. ^{1}H-NMR revealed that at the two extremes of the pH values, chemical shifts of two peaks were detected with a change in a magnitude of 0.35 ppm (Figure 38 and 39). The change in proton shift of the other seven peaks was moderately constant (0.07 ppm).

A pH versus proton shift profile showed a sigmoidal shaped distribution curve of the plotted values (Figure 38 and 39). It was determined that both peaks (H7 and H8) probably came from the same ionized residue because they had relatively the same change in proton shifts. Furthermore, we determined the observed pKa of the residue to be about 4.75. We know from our peptide design that QEKR 4.0^b contained only 3 ionizable residues----arginine, lysine, and glutamic acid. A comparison between the literature pKa values of the three residues (pKa (Arg) = 12.5; pKa (Lys) = 10.5; and pKa (Glu) = 4.3)[1] and the observed pKa proved only glutamic acid was the most likely residue that conformed to the observed pKa value.

Table 9. Proton chemical shifts of various labeled peaks at different pH.

pH	H1 (ppm)	H2 (ppm)	H3 (ppm)	H4 (ppm)	H5 (ppm)	H6 (ppm)	H7 (ppm)	H8 (ppm)	H9 (ppm)
1.20	3.68	2.99	2.88	2.86	2.79	2.75	2.33	2.31	2.21
2.09	3.70	3.01	2.91	2.88	2.81	2.77	2.36	2.33	2.23
3.08	3.71	3.02	2.92	2.89	2.82	2.78	2.37	2.34	2.24
3.98	3.72	3.02	2.92	2.90	2.82	2.78	2.34	2.33	2.24
5.84	3.72	3.02	2.93	2.90	2.82	2.78	2.16	2.15	2.24
7.06	3.71	3.01	2.92	2.89	2.81	2.77	2.14	2.11	2.23
7.98	3.69	2.98	2.89	2.87	2.78	2.74	2.12	2.09	2.21
8.07	3.68	2.97	2.89	2.86	2.78	2.74	2.12	2.08	2.21
9.09	3.67	2.96	2.88	2.85	2.76	2.72	2.11	2.08	2.20
9.37	3.66	2.95	2.86	2.82	2.74	2.71	2.10	2.06	2.18

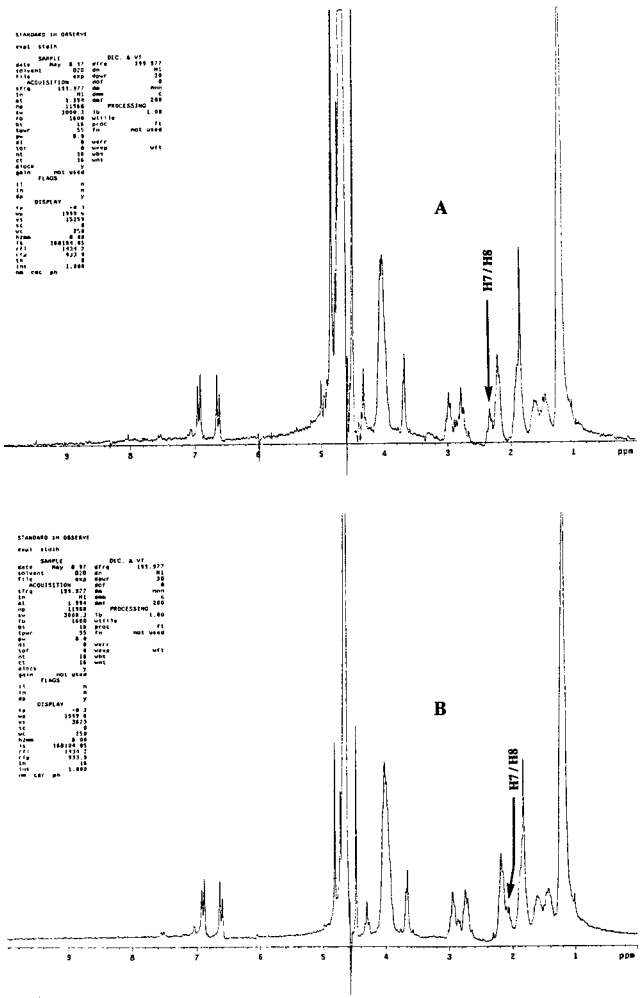


Figure 37. Titration of QEKR 4.0[b] by NMR at pH: A) 1.20; B) 9.37.

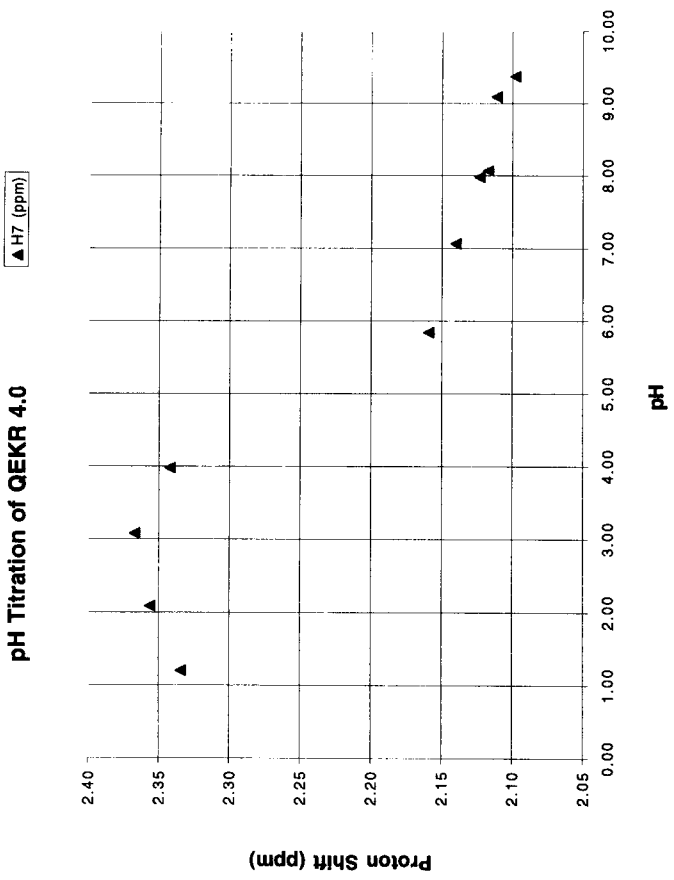


Figure 38. Determining the pKa of glutamic acid in QEKR 4.0[b]. Peak labeled as H7 was followed through the course of the titration.

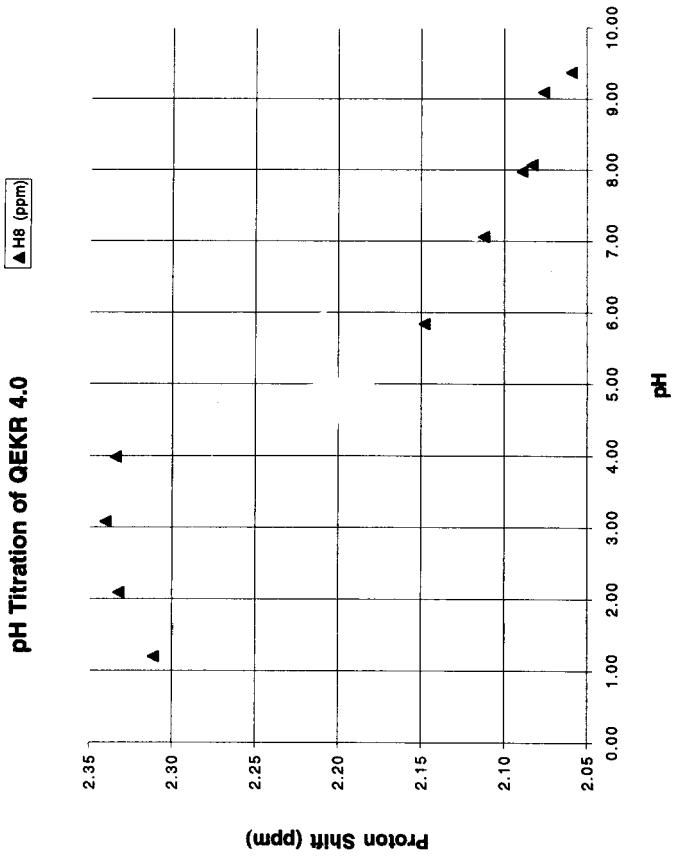


Figure 39. Determining the pKa of glutamic acid in QEKR 4.0[b]. Peak labeled as H8 was followed through the course of the titration.

Future Work

Although the design and synthesis of both peptides have been completed, we still have not taken CD measurements of our purified products to determine their degree of helicity. In addition, we have not acquired sequence searches to look for homology of our peptides to compare how they relate to other peptide sequences in proteins or large macromolecules. We would like to use MacroModel to generate thermodynamical values of our designed peptides as well as our proposed cyclic peptides for comparisons. Ultimately, our goal is to trap the designed peptides with cyanogen or other simpler coupling reagents to see if amide links were to form among the salt bridge in QEKR 4.0[b] and not in QEKR 5.0.

REFERENCES

1. Garrett, R.H.; Grisham, C.M. *Biochemistry*; Saunders College: New York, 1995; Chapter 3.

2. Taylor, H.S. *Proc. Am. Phil. Soc.* **1941**, *85*, 1-7.

3. Bragg, L.; Kendrew, J.C.; Perutz, M.F. *Proc. R. Soc. London* **1950**, *A203*, 321-357.

4. Pauling, L.; Corey, R.B.; Branson, H.R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205-211.

5. Toniolo, C.; Benedetti, E. *Trends Biochem. Sci.* **1991**, *16*, 350-353.

6. Nicoletta, A.C. BS Thesis, Union College, June 1995.

7. Schellman, J.A. *C. R. Trav. Lab. Carlsberg Ser. Chim.* **1955**, *29*, 230-259.

8. Zimm, B.H.; Bragg, J.K. *J. Chem. Phys.* **1959**, *31*, 526-535.

9. Lifson, S.; Roig, A. *J. Chem Phys.* **1961**, *34*, 1963-1974.

10. Ackermann, T.; Neumann, E. *Biopolymers* **1967**, *5*, 649-662.

11. Baldwin, R.L. *Biophys. Chem.* **1995**, *55*, 127-135.

12. Qian, H.; Schellman, J.A. *J. Phys. Chem.* **1992**, *96*, 3987-3994.

13. Chakrabartty, A.; Doig, A.J.; Baldwin, R.L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11332-11336.

14. Richardson, J.S.; Richardson, D.C. *Science* **1988**, *240*, 1648-1652.

15. Presta, L.G.; Rose, G.D. *Science* **1988**, *240*, 1632-1641.

16. Doig, A.J.; Chakrabartty, A.; Klinger, T.M.; Baldwin, R.L. *Biochemistry* **1994**, *33*, 3396-3403.

17. Chakrabartty, A.; Kortemme, T.; Baldwin, R.L. *Protein Sci.* **1994**, *3*, 843-852.

18. Doig, A.J.; Baldwin, R.L. *Protein Sci.* **1995**, *4*, 1325-1336.

19. Serrano, L.; Neira, J.-L.; Sancho, J.; Fersht, A.R. *Nature* **1992**, *356*, 453-455.

20. Schellman, J.A. *C. R. Trav. Lab. Carlsberg Ser. Chim.* **1955**, *29*, 223-229.

21. Hatano, M.; Yoneyama, M. *J. Am. Chem. Soc.* **1970**, *92*, 1392-1395.

22. Grourke, M.J.; Gibbs, J.H. *Biopolymers* **1971**, *10*, 795-808.

23. Padmanabhan, S.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *J. Mol. Biol.* **1996**, *257*, 726-734.

24. Lotan, N.; Bixon, M.; Berger, A. *Biopolymers* **1969**, *8*, 247-257.

25. Lotan, N.; Yaron, A.; Berger, A. *Biopolymers* **1966**, *4*, 365-368.

26. Lupu-Lotan, N.; Yaron, A.; Berger, A.; Sela, M. *Biopolymers* **1965**, *3*, 625-655.

27. Fasman, G.D.; Lindblow, C.; Bodenheimer, E. *Biochemistry* **1964**, *3*, 155-166.

28. Chou, P.Y.; Wells, M.; Fasman, G.D. *Biochemistry* **1972**, *11*, 3028-3043.

29. Scheraga, H.A. *Pure Appl. Chem.* **1978**, *50*, 315-324.

30. Sueki, M.; Lee, S.; Powers, S.P.; Denton, J.B.; Konishi, Y.; Scheraga, H.A. *Macromolecules* **1984**, *17*, 148-155.

31. Wójcik, J.; Altmann, K.-H.; Scheraga, H.A. *Biopolymers* **1990**, *30*, 121-134.

32. Brown, J.E.; Klee, W.A. *Biochemistry* **1971**, *10*, 470-476.

33. Epand, R.M.; Scheraga, H.A. *Biochemistry* **1968**, *7*, 2864-2872.

34. Taniuchi, H.; Anfinsen, C.B. *J. Biol. Chem.* **1969**, *244*, 3864-3875.

35. Bierzynski, A.; Kim, P.S.; Baldwin, R.L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470-2474.

36. Shoemaker, K.R.; Fairman, R.; Schultz, D.A.; Robertson, A.D.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *Biopolymers* **1990**, *29*, 1-11.

37. Marqusee, S.; Robbins, V.H.; Baldwin, R.L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286-5290.

38. Marqusee, S.; Baldwin, R.L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8898-8902.

39. Marqusee, S.; Baldwin, R.L. *In Protein Folding*; American Association for the Advancement of Science: Washington DC, 1990; pp 85-94.

40. Berger, J.S.; Ernst, J.A.; Nicoletta, A.C.; Hull, L.A.; Yang, J.; Qiu, R.; Morozov, V.N.; Kallenbach, N.R. *J. Biomol. Struct. Dynam.* **1996**, *14*, 285-291.

41. Karagözler, A.A.; Ghenbot, G.; Day, R.A. *Biopolymers* **1993**, *33*, 687-692.

42. Day, R.A.; Hignite, A.; Gooden, W.E. *Techniques in Protein Chemistry VI*; Academic: San Diego, 1995; pp 435-442.

43. Merrifield, R.B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

44. Jones, J. *Amino Acid and Peptide Synthesis*; Oxford University: New York, 1992.

45. Chakrabartty, A.; Kortemme, T.; Padmanabhan, S.; Baldwin, R.L. *Biochemistry* **1993**, *32*, 5560-5565.

46. Ösapay, G.; Taylor, J.W. *J. Am. Chem. Soc.* **1992**, *114*, 6966-6973.

47. Scholtz, J.M.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *J. Am. Chem. Soc.* **1991**, *113*, 5102-5104.

48. Liff, M.I.; Lyu, P.C.; Kallenbach, N.R. *J. Am. Chem. Soc.* **1991**, *113*, 1014-1019.

49. Billeter, M.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1982**, *155*, 321-346.

50. Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715-740.

51. Zhou, H.X.; Hull, L.A.; Kallenbach, N.R.; Mayne, L.; Bai, Y.; Englander, S.W. *J. Am. Chem. Soc.* **1994**, *116*, 6482-6483.

52. Shalongo, W.; Dugad, L.; Stellwagen, E. *J. Am. Chem. Soc.* **1994**, *116*, 2500-2507.

53. Elliott, A.; Ambrose, E.J. *Nature (London)* **1950**, *4206*, 921-922.

54 Susi, H.; Byler, D.M. *Methods Enzymol.* **1986**, *130*, 290-311.

55. Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer, Jr., E.F.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535-542.

56. Ramage, R.; Green, J.; Blake, A.J. *Tetrahedron Lett.* **1991**, *47*, 6353-6370.

57. Rink, H. *Tetrahedron Lett.* **1987**, *28*, 3787-3790.

58. Stewart, J.M.; Young, J.D. *Solid Phase Peptide Synthesis*; Pierce Chemical: Illinois, 1984; pp 105-106.