I. Charged Pair Hydrogen Bonding Interactions in Collagen Heterotrimers
II. Surface Enhanced Raman Spectroscopy of Aromatic Peptides

by

Fang Wei

A THESIS SUBMITTED IN PARTIAL Fulfillment of the REQUIREMENTS FOR THE DEGREE Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Jeffrey D. Hartgerink, Associate Professor of Chemistry and Bioengineering, Chair

Ronald Parry, Professor of Chemistry and Professor of Biochemistry and Cell Biology

K. Jane Grande-Allen, Associate Professor of Bioengineering

HOUSTON, TEXAS
August 2009
ABSTRACT

I. Charged Pair Hydrogen Bonding Interactions in Collagen Heterotrimers

II. Surface Enhanced Raman Spectroscopy of Aromatic Peptides

by

Fang Wei

Eight ABC heterotrimers whose self-assembly are directed through electrostatic interactions were studied here. Oppositely charged pairs of amino acids, with varying side chain length, were assessed for their ability to stabilize a triple helix. Aspartate-lysine was found to result in the most thermally stable helix followed by lysine-glutamate, ornithine-aspartate and finally ornithine-glutamate. When the sequence position of these charged amino acids was reversed from what is normally observed in nature, triple helix stability and compositional purity was significantly reduced. The effect of salt on triple helix stability was explored and it was observed that increased salt concentration reduces the thermal stability of heterotrimers by an average of 5°C, but does not disrupt helix assembly. It was also found that positively charged homotrimers can be stabilized in the presence of phosphate anions.

Raman and Surface-enhanced Raman spectroscopies (SERS) are potentially important tools in the characterization of biomolecules such as proteins and DNA. In this work, SERS spectra of three cysteine containing aromatic peptides: tryptophan-cysteine,
tyrosine-cysteine, and phenylalanine-cysteine, bound to Au nanoshell substrates, were obtained and compared to their respective normal Raman spectra. While the full widths at half maximum of the SERS peaks are significantly broadened (up to 70%), no significant spectral shifts (<6 cm\(^{-1}\)) of the major Stokes modes were observed between the two modalities. It is shown that the Raman and SERS spectra of penetratin, a cell-penetrating peptide, can be evaluated quite reliably from the spectra of its constituent aromatic amino acids except in the \(-\text{CH}_2-\) bending and amide I and III regions where the spectral intensities are critically dependent on the chain length and/or protein conformations. From this study we conclude that with aromatic amino acid residues provide the overwhelmingly dominant features in the Raman and SERS spectra of peptides and proteins, and that the Raman modes of these three small constructed peptides can apply directly to the assignment of Raman and SERS features in the spectra of larger peptides and proteins.
ACKNOWLEDGEMENTS

I gratefully acknowledge my mentor, professor Jeffrey Hartgerink, for his guidance, patience, encouragement and invaluable advice throughout my studies at Rice. I would like to thank Dr. Naomi Halas for her guidance on the SERS project. I truly appreciate her scientific excellence. I would also like to thank my other thesis committee members, Dr. Ronald Parry and Dr. Jane Grande-Allen for serving on my committee, very careful proofreading of my thesis and for their advice on many aspects.

I would like to acknowledge all my former and present lab members for their helpful discussions, support and help. I would also like to thank the past and current Halas group members for their support on the SERS project, especially to Dr. Hui Wang, Dr. Dongmao Zhang and Dr. Felicia Tam for their generous help and friendship.

Most importantly, I am grateful to my parents who not only love us but also put our education as the top priority no matter what. I am grateful to my sister and my two brothers for their encouragement and support. I am also grateful to my husband Yongqiang, my daughter Sandy, and my son Max, for their unconditional love and support. I feel blessed to have my son who was born in the first semester when I started my study at Rice and grows up with my research progresses. Without their love and support, I would not be able to overcome those rough times and come this far.
# Table of Contents

Abstract........................................................................................................................................... ii

Acknowledgements........................................................................................................................... iv

Table of Contents.............................................................................................................................. v

List of Tables..................................................................................................................................... ix

List of Figures...................................................................................................................................... xi

Abbreviations..................................................................................................................................... xv

Chapter 1  Introduction: Charged Pair Hydrogen Bonding Interactions in Synthetic Collagen Heterotrimers

1.1 Collagen Structure Characteristics.............................................................................................. 2

1.2 Benefits of Study Synthetic Collagen-like Peptides...................................................................... 4

1.3 Homotrimeric Study of Stability of Collagen-like Peptide Systems.............................................. 5

1.3.1 Homotrimeric Associated Triple Helical Collagen-like Peptides................................................ 6

1.3.2 Homotrimeric Branched Triple Helical Collagen-like Peptides................................................. 8

1.3.3 Brodsky's Host-guest Amino acid Propensity Study................................................................. 8

1.3.4 Heterotrimeric Collagen-like Peptides.................................................................................... 11

1.4 Peptide and System Design.......................................................................................................... 14

1.5 Peptide Synthesis and Purification............................................................................................... 16

1.6 Characterization Methods............................................................................................................ 17

1.6.1 Circular Dichroism Spectroscopy (CD).................................................................................... 17

1.6.2 Differential Scanning Calorimetry (DSC)................................................................................ 19

1.6.3 Amino Acid Analysis (AAA).................................................................................................. 22

1.6.4 Analytical Ultracentrifugation (AUC).................................................................................... 22

References.......................................................................................................................................... 26

Chapter 2  Homotrimeric Collagen-like Peptide Systems

2.1 Introduction................................................................................................................................. 30
2.2 Design of Peptides ................................................. 31
2.3 Characterization .................................................. 32
  2.3.1 Circular Dichroism (CD) ..................................... 32
  2.3.2 Differential Scanning Calorimetry (DSC) ................... 33
2.4 Results and Discussion ........................................... 33
  2.4.1 Host-guest Peptides ......................................... 34
  2.4.2 (XYG)$_{10}$ Homotrimer ..................................... 36
    2.4.2.1 Positively charged (XYG)$_{10}$ Systems ................. 36
    2.4.2.2 Negatively charged (XYG)$_{10}$ Systems ............... 41
2.5 Conclusion ....................................................... 43

References .................................................................. 46

Chapter 3 Study of Heterotrimeric Collagen-like Peptide Systems: Group 1
(A$_{X_{i-}}$B$_{Y_{+1}}$C$_{O}$)
3.1 Introduction ....................................................... 47
3.2 Characterization Methods ....................................... 50
  3.2.1 Circular Dichroism (CD) .................................... 50
  3.2.2 Differential Scanning Calorimetry (DSC) ................. 51
  3.2.3 Analytical Ultracentrifugation (AUC) ..................... 52
    3.2.3.1 (DOG)$_{10}$:(PKG)$_{10}$GY:(POG)$_{10}$ .................. 54
    3.2.3.2 (DOG)$_{10}$:(PKG)$_{10}$GY ............................. 54
3.3 Result and discussion ............................................. 55
  3.3.1 (DOG)$_{10}$•(PorG)$_{10}$•(POG)$_{10}$ ......................... 55
  3.3.2 (EOG)$_{10}$•(PorG)$_{10}$•(POG)$_{10}$ ........................ 57
  3.3.3 Effects of Different Ratio of (PorG)$_{10}$ on the ABC System ............ 59
  3.3.4 Effect of Salts on Helix Stability .......................... 60
  3.3.5 Thermodynamic Analysis ................................... 63
  3.3.6 Using Analytical Ultracentrifugation to Study the Composition
List of Tables

Table 1.1. Melting Temperature ($T_m$) of host-guest peptides with 20 amino acids' substitutions at guest region. ................................. 11

Table 1.2. Collagen-like peptide sequence. ..................................................... 15

Table 1.3. ABC heterotrimeric systems. ......................................................... 16

Table 2.1. Peptide sequence, abbreviation and melting temperature of homotrimers. ....................................................... 32

Table 2.2. Melting temperature of Ornithine containing peptides compared to lysine. .......................................................... 35

Table 2.3. Thermodynamic parameters obtained from DSC. ......................... 36

Table 3.1. Group 1 ABC heterotrimeric systems. ............................................. 49

Table 3.2. Comparison of $T_m$s from CD thermal unfolding studies. ............. 53

Table 3.3. Possible heterotrimers in (DOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture. .......................................................... 57

Table 3.4. Possible heterotrimers in (EOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture. .......................................................... 58

Table 3.5. Comparison of the area of $A_{(EOG)}B_{(POrG)}C_{(POG)}$ peak with different ratio of (POrG)$_{10}$ from DSC spectra. .................................. 60

Table 3.6. Summary of the $T_m$ of ABC heterotrimers in phosphate buffer and PBS. 63

Table 3.7. Thermodynamic parameters obtained from DSC (in phosphate buffer). 65

Table 3.8. Summary of sedimentation velocity data with Monte Carlo analysis. 66

Table 3.9. Expected molecular weight of different combinations. ..................... 68

Table 3.10. Summary of sedimentation velocity data with Monte Carlo analysis. 69

Table 4.1. Group 2 ABC heterotrimeric systems. ............................................ 74

Table 4.2. Heterotrimers in (KOG)$_{10}$, (PDG)$_{10}$ and (POG)$_{10}$ peptide mixture. 78
Table 4.3. Heterotrimers from (KOG)$_{10}$, (PEG)$_{10}$ and (POG)$_{10}$.

Table 4.4. Possible heterotrimers in (OrOG)$_{10}$, (PDG)$_{10}$ and (POG)$_{10}$ peptide mixture.

Table 4.5. Possible heterotrimers in (OrOG)$_{10}$, (PEG)$_{10}$ and (POG)$_{10}$ peptide mixture.

Table 4.6. Summary of the T$_m$ of ABC heterotrimers in phosphate buffer and PBS.

Table 4.7. Thermodynamic parameters obtained from DSC (in phosphate buffer).

Table 5.1. Peptide Sequences.

Table 5.2. Raman and SERS spectra peaking fitting parameters.

Table 5.3. Relative enhancements of the Stokes modes.
List of Figures

Figure 1.1. Biological superstructures formed by collagen triple-helix. 2
Figure 1.2. Ball and stick diagrams showing two projections of the currently accepted triple helical structure for collagen with one inter-chain hydrogen bond per tripeptide. 3
Figure 1.3. The staggered alignment of individual chains in a triple helix collagen model. 4
Figure 1.4. Conformational preferences of imino acids. 7
Figure 1.5. General schemes of branched, triple-helical peptides. 9
Figure 1.6. CD of collagen triple helix. 18
Figure 1.7. Sketch diagram of a typical DSC. 20
Figure 1.8. Sedimentation velocity and equilibrium cell design. 23
Figure 1.9. Sedimentation velocity data. 24
Figure 1.10. Schematic representation of sedimentation equilibrium. 26
Figure 2.1. Circular Dichroism spectra of Orn containing host-guest peptides 1-3. 34
Figure 2.2. DSC spectra of host guest homotrimers: (POG)_8G (in black), (POG)_3 (OrOG) (POG)_4G (in red), (POG)_3 (POrG) (POG)_4G (in green). 35
Figure 2.3. CD spectra of (OrOG)_{10} homotrimer in PBS and phosphate buffer. 37
Figure 2.4. CD spectra of (KOG)_{10} homotrimer in PBS buffer before and after incubation at 5°C. 37
Figure 2.5. Comparison of structures of phosphate and Tris buffer. 38
Figure 2.6. Circular Dichroism spectra in Tris buffer. 38
Figure 2.7. CD spectra of (POrG)_{10} in phosphate buffer. 39
Figure 2.8. CD spectra of (PKG)_{10} in phosphate buffer. 39
Figure 2.9. CD spectra of (KOG)_{10} peptide taken with a 1 hour interval of 60 scans. 40
Figure 2.10. CD spectra and thermal unfolding of (ROG)\textsubscript{10} peptide in 10mM phosphate buffer. ........................................... 40

Figure 2.11. CD spectra and thermal unfolding of (EOG)\textsubscript{10} peptide in 10mM phosphate buffer. ........................................... 42

Figure 2.12. CD spectra (EOG)\textsubscript{10} peptide in 10mM Tris buffer. 42

Figure 2.13. CD spectra (EOG)\textsubscript{10} peptide in 10mM Tris buffer with cations. 43

Figure 2.14. CD spectra of (DOG)\textsubscript{10} in PBS buffer. ........................................... 43

Figure 2.15. CD spectra of (PDG)\textsubscript{10} in PBS buffer. ........................................... 44

Figure 2.16. CD spectra of (PEG)\textsubscript{10} in PBS buffer. ........................................... 44

Figure 3.1. CD thermal unfolding study of binary and trimeric system with \((PKG)_{10}GY\). .................................................. 53

Figure 3.2. \(A_{\text{PEG}}B_{\text{POR}}\) in phosphate buffer. ........................................... 55

Figure 3.3. \(A_{\text{DOG}}B_{\text{POR}}\) in phosphate buffer. ........................................... 56

Figure 3.4. CD of collagen peptide \((DOG)B_{\text{POR}}C_{\text{POG}}\) in phosphate buffer. ........................................... 56

Figure 3.5. \(A_{\text{EOG}}B_{\text{POR}}\) in phosphate buffer (a) Thermal transition. ........................................... 58

Figure 3.6. CD Thermal unfolding of \((EOG)B_{\text{POR}}C_{\text{POG}}\) in phosphate buffer. ........................................... 58

Figure 3.7. DSC spectra of \((EOG)B_{\text{POR}}C_{\text{POG}}\) with different ratio of \((POR)_{10}\). ........................................... 60

Figure 3.8. \((DOG)B_{\text{PKG}}C_{\text{POG}}\) thermal transition comparison in phosphate buffer and PBS. ........................................... 61

Figure 3.9. \((EOG)B_{\text{PKG}}C_{\text{POG}}\) thermal transition comparison in phosphate buffer and PBS. ........................................... 62

Figure 3.10. \((OG)B_{\text{POR}}C_{\text{POG}}\) thermal transition comparison in phosphate buffer and PBS. ........................................... 62

Figure 3.11. \((DOG)B_{\text{POR}}C_{\text{POG}}\) thermal transition comparison in phosphate buffer and PBS. ........................................... 63

Figure 3.12. DSC spectra of Group 1 ABC helices along with homotrimer \((POG)_{10}\). ........................................... 64

Figure 3.13. Sedimentation velocity data analysis of \((DOG)_{10}:(PKG)_{10}GY:(POG)_{10}\) system. ........................................... 65
Figure 3.14. Sedimentation equilibrium analysis of (DOG)_{10}:(PKG)_{10}GY:(POG)_{10} system. .................................................. 66

Figure 3.15. Sedimentation velocity data analysis of (DOG)_{10}:(PKG)_{10}GY system. ................................. 68

Figure 4.1. A_{(KOG)}B_{(PDG)} in phosphate buffer. ................................................................. 76

Figure 4.2. A_{(KOG)}B_{(POG)} in PBS. ................................................................. 76

Figure 4.3. A_{(PDG)}B_{(POG)} in PBS. ................................................................. 77

Figure 4.4. Thermal unfolding of collagen peptide (KOG)_{10}*(PDG)_{10}*(POG)_{10} in 10mM sodium phosphate buffer as monitored by CD. .................................................. 78

Figure 4.5. A_{(KOG)}B_{(PEG)} in phosphate buffer. ................................................................. 79

Figure 4.6. A_{(PEG)}B_{(POG)} in PBS. ................................................................. 80

Figure 4.7. Thermal unfolding of collagen peptide (KOG)_{10}*(PEG)_{10}*(POG)_{10} in 10mM sodium phosphate buffer as monitored by CD. .................................................. 80

Figure 4.8. A_{(OrOG)}B_{(PDG)} in 10 mM phosphate buffer. ................................................................. 81

Figure 4.9. A_{(OrOG)}B_{(POG)} in PBS. ................................................................. 82

Figure 4.10. Thermal unfolding of collagen peptide (OrOG)_{10}*(PDG)_{10}*(POG)_{10} in 10mM sodium phosphate buffer as monitored by CD. .................................................. 82

Figure 4.11. A_{(OrOG)}B_{(PEG)} in phosphate buffer. ................................................................. 84

Figure 4.12. Thermal unfolding of collagen peptide (OrOG)_{10}*(PEG)_{10}*(POG)_{10} in 10mM sodium phosphate buffer as monitored by CD. .................................................. 84

Figure 4.13. A_{(KOG)}B_{(PDG)}C_{(POG)} thermal transition comparison in phosphate buffer and PBS. ........................................................................ 86

Figure 4.14. A_{(KOG)}B_{(PEG)}C_{(POG)} thermal transition comparison in phosphate buffer and PBS. ........................................................................ 86

Figure 4.15. A_{(OrOG)}B_{(PDG)}C_{(POG)} thermal transition comparison in phosphate buffer and PBS. ........................................................................ 87

Figure 4.16. A_{(OrOG)}B_{(PEG)}C_{(POG)} thermal transition comparison in phosphate buffer and PBS. ........................................................................ 87

Figure 4.17. DSC spectra of Group 2 ABC helices. ........................................................................ 89
Figure 5.1. UV-Vis spectra of complete Au Nanoshell. ................................. 98
Figure 5.2. TEM Images of Au nanoshells at different fabrication stages. .............. 99
Figure 5.3. SERS spectra of the WC peptide from 400 cm⁻¹ to 1800 cm⁻¹. ............. 104
Figure 5.4. SERS spectra of Phe-Cys from 400 cm⁻¹ to 1800 cm⁻¹. ....................... 105
Figure 5.5. SERS spectra of Tyr-Cys from 400 cm⁻¹ to 1800 cm⁻¹. .................... 106
Figure 5.6. SERS and normal Raman spectra of the FC peptide. ......................... 108
Figure 5.7. Raman and SERS spectra of the YC peptide. ................................. 110
Figure 5.8. Raman and SERS spectra of the WC peptide. ................................ 112
Figure 5.9. Comparison of empirically predicted (red) and directly measured experimental (black) spectra of penetratin. .............................. 116
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>amino acid analysis</td>
</tr>
<tr>
<td>APTMS</td>
<td>3-aminopropyltrimethoxysilane</td>
</tr>
<tr>
<td>AUC</td>
<td>analytical ultracentrifugation</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>COL29A1</td>
<td>collagen XXIX chain α1</td>
</tr>
<tr>
<td>COL6A5</td>
<td>collagen VI chain α5</td>
</tr>
<tr>
<td>DIEA</td>
<td>N,N-diisopropylethylamine</td>
</tr>
<tr>
<td>D</td>
<td>diffusion coefficient</td>
</tr>
<tr>
<td>DSC</td>
<td>differential scanning calorimetry</td>
</tr>
<tr>
<td>EDT</td>
<td>1,2-ethane dithiol</td>
</tr>
<tr>
<td>FACIT</td>
<td>fibril-associated collagens with interrupted triple helices</td>
</tr>
<tr>
<td>FMOC</td>
<td>fluorenylmethyloxycarbonyl</td>
</tr>
<tr>
<td>FQ</td>
<td>fused quartz</td>
</tr>
<tr>
<td>HBTU</td>
<td>O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate</td>
</tr>
<tr>
<td>HOBr</td>
<td>1-hydroxybenzotriazole hydrate</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>MALDI</td>
<td>matrix assisted laser desorption ionization</td>
</tr>
<tr>
<td>MRE</td>
<td>molar residue ellipticity</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
</tbody>
</table>
PVP  poly-vinyl pyridine
S    sedimentation coefficient
SERS surface enhanced Raman spectroscopy
T_m  melting temperature
TEM  transmission electron microscopy
TEOS tetraethyl orthosilicate
TFA  trifluoroacetic acid
TOF  time of flight
Tris tris(hydroxymethyl)aminomethane
Chapter 1. Introduction

Charged Pair Hydrogen Bonding Interactions in Synthetic Collagen Heterotrimers

Collagen is an abundant naturally occurring protein in mammals; about 20-30% of proteins in animals are collagens. They are the major components of cartilage, skin, ligaments, bone, and tendons. "Collagen" is a Greek word, meaning glue. Initially it was used to describe the components of connective tissues that can form gelatin by boiling. There are at least 28 natural human collagens reported to date. Recently, Lee claimed a new type of COL29A1. However, there had been arguments that the new COL29A1 was actually COL6A5, which was also newly discovered. Based on the superstructure it forms, collagen has several subfamilies, such as fibrillar collagens, Fibril-Associated Collagens with Interrupted Triple Helices (FACIT), network-forming collagens, anchoring fibrils collagen, transmembrane collagens and others. Some of the biological superstructures formed by collagen triple helices are shown in figure1.1. Collagens not only provide elasticity and tensile strength to tissues but also help to form critical components of the extracellular matrix. They have applications in medical and cosmetic areas.
1.1 Collagen Structure Characteristics

Collagens are right handed triple helical structures formed by three left handed polyproline II helices. The unique composition of collagen is that it is mainly composed of a tripeptide repeat (X-Y-Gly). The smallest amino acid, glycine (abbreviated as Gly or G), must be at every third position to ensure the formation of the closed packed structure. The structures also have to have high content of imino acids, proline (abbreviated as Pro or P) and its modified variant hydroxyproline (abbreviated as Hyp or O). One of the major driving forces for triple helix formation of collagen is the inter-chain hydrogen bonds. These occur between the N-H group of glycine and the C=O group of the amino acids in the X positions from the adjacent chain. In order for the hydrogen bonds to form, the...
neighboring chains have to overcome the steric effects to be close to each other. So the three chains of the triple helix have to have a small rotation and translation to avoid the steric clashes. The triple helical structure prefers imino acids in both the X and Y positions and glycine at the third position. This structure was proposed by Ramachandran and subsequently by Rich and Crick as shown in Figure 1.2.\textsuperscript{1,11} The dihedral angles in the X and Y positions in each chain have to be such that all three chains can be in close proximity to each other. The dihedral angles in the X and Y positions in each chain have to be such that all three chains can be in close proximity to each other.

**Figure 1.2.** Ball and stick diagrams showing two projections of the currently accepted triple helical structure for collagen with one inter-chain hydrogen bond per tripeptide. The sequence shown is (Gly-Pro-Hyp)_3 and each chain in the triple helix has a different color ribbon drawn through the backbone.\textsuperscript{1} (Figure and caption are taken from reference # 1.)
Y position of the backbone of (PPG)_{10} are similar to those typically found in imino acids, meaning that high content of imino acids in collagen are essential because they can form triple helices without large strain. And the high content of imino acids enhances the thermal stability of the triple helix. The most cited model of a triple helix from x-ray fiber diffraction data shows that the three left-handed polyproline II chains staggered by one residue from each other as shown in Figure 1.3. Each of the cross-section shown in the shaded area is composted of Gly-Hyp-Pro, and each of the amino acid belongs to a different chain of the triple helix. This ensures that each of the cross-sections has a Gly to guarantee the closely packed structure.

```
collagen chain #
Pro Hyp Gly Pro Hyp Gly Pro Hyp Gly  1
Pro Hyp Gly Pro Hyp Gly Pro Hyp Gly  2
Pro Hyp Gly Pro Hyp Gly Pro Hyp Gly  3
```

**Figure 1.3.** The staggered alignment of individual chains in a triple helix collagen model peptide (Gly-Pro-Hyp)₃, the cross-section of a Gly-Pro-Hyp unit (shaded area) as it is aligned in this collagen model peptide shown below. Each residue belongs to a different chain of the triple helix. (Caption is taken from reference # 14.)

### 1.2 Benefits of Study Synthetic Collagen-like Peptides

Natural collagens have been used to understand the interactions of specific biofunctions of collagen and biomolecules. However, they pose some difficulties to characterize natural collagen molecules. One is that collagens are not very soluble in buffers under physiological condition. Second, it is difficult to express and purify
recombinant collagen using standard bacterial expression systems. Third, it is not easy to obtain shorter collagen fragments; Even though shorter fragments could be obtained, they always have already lost their triple helical structure. In order to understand, on a molecular level, the factors that affect the folding process and stability, synthetic collagen-like peptides have been used as a surrogate to facilitate the understanding of the factors that affect the stability of the triple helical structures, folding processes, and the biofunctions of collagen such as binding of collagen to proteins. The final goal of this work is to develop biofunctional materials based on the above knowledge to use for extracellular matrix and tissue engineering.

1.3 Homotrimeric Study of Stability of Collagen-like Peptide Systems

Among the 28 human collagens, about 90% of them are type I, II, III and IV. Type II and III are homotrimers, in which all three chains are the same. Type I and IV are heterotrimers, in which two or even three chains are different. To understand the structure and stability of native collagen, using collagen peptides to study both of the homotrimers and heterotrimers are equally important. However, studies have been focused on homotrimers, such as associated peptides or branched peptides, mainly because they are easy to form stable triple helices. Attempts to study heterotimers have been limited due to the difficulty of making stable triple helices.
1.3.1 Homotrimeric Associated Triple Helical Collagen-like Peptides

Associated triple helices are triple helical structures formed by non-covalent interactions. Since the structure of collagen was proposed in 1954 by Ramachandran and in 1955 by Rich and Crick, studies have been focused on collagen model peptides in order to further understand the structure and function of collagen. The first model collagen peptide synthesized using solid phase peptide synthesis was (PPG)$_{10}$ by Sakakibara. \(^{18}\) Thermal unfolding studies found the melting temperature ($T_m$) was 41°C. \(^{19}\) Studies showed that (P$_O$R$_G$)$_{10}$ (R refers to the hydroxyl group of the hydroxyproline is at UP conformation) formed triple helices with $T_m$ of 69°C. This indicates that hydroxyproline at the Y positions can stabilize triple helix. \(^{20}\) Many studies had been done to try to explain this stabilization. In 1994, a breakthrough was reported by Brodsky’s group with high resolution X-ray analysis of peptide (Pro-Hyp$_R$-Gly)$_{10}$-Pro-Hyp$_R$-Ala-(Pro-Hyp$_R$-Gly)$_5$. \(^{21}\) In this peptide, R represents the steric conformation of hydroxyl group at $\gamma$ position in hydroxyproline. It not only provided the first high resolution structure of a collagen triple helix, but also proved that hydroxyl groups of Hyp formed hydrogen bonds with surrounding water molecules. However, when Rains reported that (Pro-fPro$_R$-Gly)$_{10}$ (fPro$_R$ refers to fluoro- replacing the hydroxyl group of the hydroxyproline) formed a triple helix with a $T_m$ of over 90°C, \(^{19,22}\) the water bridge model was in question because fPro$_R$ had a low ability to form hydrogen bonds. The alternative explanation of how the Hyp at the Y position can stabilize the triple helix may be the inductive effects of the electronegative oxygen of the hydroxyl group, which favors the trans conformation of hydroxyproline.
(Hyp\textsuperscript{R}). The formation of a stable triple helix relates to the conformation of hydroxyproline. Researchers found that not only (Hyp\textsuperscript{S}-Pro-Gly)\textsubscript{10} and (Pro-Hyp\textsuperscript{S}-Gly)\textsubscript{10} can not form triple helices,\textsuperscript{23} but also (Hyp\textsuperscript{R}-Pro-Gly)\textsubscript{10} can not form a triple helix.\textsuperscript{24} To explain the substitution effect of hydroxyproline, Zagari proposed propensity-based rules of the pyrrolidine ring pucker for the stability of collagen triple helix derived from x-ray data of various collagen model peptides.\textsuperscript{25,26} They found that the puckering should be “up” at the Y position and “down” at the X position. It is confirmed that Hyp\textsuperscript{R} and fPro\textsuperscript{R} prefer the up form and the Hyp\textsuperscript{S} and fPro\textsuperscript{S} prefer the down form, and Pro can take both forms as illustrated in Figure 1.4.\textsuperscript{27} But people found some exceptions to this rule. For example,

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure14.png}
\caption{Conformational preferences of imino acids. A) and B) show the most populated conformations of proline residues. C) and D) indicate the most populated conformations of 4R-Hyp and 4S-Hyp, respectively.\textsuperscript{12} (Figure and caption are taken from reference #12)}
\end{figure}
(Hyp$^{5}$-Pro-Gly)$_{10}$ should form a stable triple helix, but it did not. And (Hyp$^{R}$-Hyp$^{R}$-Gly)$_{10}$ formed a thermally stable triple helix. A crystal structure showed that the consecutive Hyp$^{R}$ residues were in the up-up conformation.$^{27,28}$

1.3.2 Homotrimeric Branched Triple Helical Collagen-like Peptides

Branched triple helices are another class of triple helix composed of collagen-like peptides. They are three peptide strands covalently linked via a C-terminal branch. It is a mimic of type III collagen in which the C-terminus is linked by disulfide bonds. As shown in Figure 1.5, the triple helical peptide is composed of a C-terminal branch linked by two lysine (abbreviated as Lys or K) residues, each chain has a collagen sequence and a N-terminal tripeptide (Gly-Pro-Hyp)$_{n}$ repeats.$^{29}$ There are three methods reported to make branched triple helical peptides: solid-phase branching, solution-phase cross linking and chemoselective ligation. The advantage of using branched peptides is that the enhanced thermal stability allows mechanistic investigation of collagen mediated activities, such as cell activation and signaling. The other advantage is it could be used to construct heterotrimeric collagen like peptides to facilitate the understanding of collagen structures.$^{29}$

1.3.3 Brodsky’s Host-guest Amino Acid Propensity Study

Using host-guest approaches to study the tendency of amino acids to form a collagen triple helix were first applied by Brodsky in late 90’s.$^{30-32}$ It has been used to evaluate the propensity of an amino acid to form a collagen triple helix at specific positions
(X or Y). It provides us a way to evaluate how the amino acids and their positions affect the stability of the triple helices and gives us guidelines for designing collagen-like peptides. In host-guest studies, the sequence acetyl-(GPO)$_3$-GXY-(GPO)$_4$GG-CONH$_2$ was used, in

\[
\text{Gly-Pro-Hyp}, \quad \text{Gly-X-Y}, \quad \text{Ahx}
\]

which the middle GXY was the guest triplet. Nineteen natural amino acids were selected to be placed at the X or Y positions to study how the substitution affects the stability of the triple helix. The host peptide was (POG)$_5$G. Eight repeats of the triplet were selected to ensure the thermal stability of the guest would not be masked if a more stable host peptide was used. Studies showed that (POG)$_{10}$ formed a very stable triple helix with a melting temperature ($T_m$) of 61°C while (POG)$_5$ was very unstable with a $T_m$ of about 2°C.

**Figure 1.5.** General schemes of branched, triple-helical peptides. The triple helical peptide is composed of a C-terminal branch generated from two Lys residues, one native collagen sequence per chain, and N-terminal Gly-Pro-Hyp repeats. Ahx is 6-aminohexanoic acid. (Figure and caption are taken from reference #29.)
Acetylation of the N-terminus and amidation of the C-terminus were used to eliminate the destabilization that charges at both ends of the helix would cause. Studies have shown peptides without terminal protection have lower thermal stabilities. \(^{35,36}\) Host-guest studies of all the natural amino acids at the X and the Y positions showed that all of the peptides form stable triple helices, but with decreased stability. Table 1 (a) shows the host-guest study results of the amino acids substituting at the X position in the guest region as Gly-X-Hyp. \(^{32}\) The host peptide had a \(T_m\) of 47.3°C. When replacing Pro in the guest X position by any other amino acid, \(T_m\) dropped. Pro at the X position formed the most stable triple helix. All the charged residues such as glutamic acid (abbreviated as Glu or E), Lys, arginine (abbreviated as Arg or R) and aspartic acid (abbreviated as Asp or D) formed stable triple helices but with their \(T_m\) decreased accordingly. The stabilities of the triple helices formed by nonpolar residues were in the middle. The most unstable triple helices were formed by aromatic residues and Gly. Table 1.1 (b) shows the host-guest study results of the amino acids substituting at the Y position in the guest region as Gly-Pro-Y. \(^{32}\) The host peptide had a \(T_m\) of 47.3°C. With replacing Hyp in the guest Y position by the other amino acids, the \(T_m\) dropped with the exception of Arg with \(T_m\) of 47.2°C. The aromatic residues formed the most unstable triple helices, while the other residues, such as charged, nonpolar and other residues were with mixed stabilities, especially the charged residue Asp, as it had a \(T_m\) with an almost 13°C decrease. Overall, when the residues were substituting at the X positions, they didn’t show much difference with the host peptide, but when the residues were substituted at the Y positions, \(T_m\) showed that the triple helices
became much less stable. Apparently, the X and Y positions were not equivalent. This relates to the differences between the inter-chain interactions and accessibility to solvent of the two positions. It was shown that the X positions were highly exposed to solvent, while the Y positions were less accessible to solvent due to their proximity to the neighboring chains.

Table 1.1. Melting Temperature ($T_m$) of host-guest peptides with 20 amino acids’ substitutions at the guest region as (a) Gly-X-Hyp, and (b) Gly-Pro-Y. (Table and caption are taken from reference # 32.)

<table>
<thead>
<tr>
<th></th>
<th>Gly-X-Hyp</th>
<th>$T_m$ (°C)</th>
<th>Gly-Pro-Y</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>47.3</td>
<td></td>
<td>Hyp</td>
<td>47.3</td>
</tr>
<tr>
<td>Glu$^b$</td>
<td>42.9</td>
<td></td>
<td>Arg</td>
<td>47.2</td>
</tr>
<tr>
<td>Ala</td>
<td>41.7</td>
<td></td>
<td>Met</td>
<td>42.6</td>
</tr>
<tr>
<td>Lys$^b$</td>
<td>41.5</td>
<td></td>
<td>Ile</td>
<td>41.5</td>
</tr>
<tr>
<td>Arg$^a$</td>
<td>40.6</td>
<td></td>
<td>Gin</td>
<td>41.3</td>
</tr>
<tr>
<td>Cin</td>
<td>40.4</td>
<td></td>
<td>Ala</td>
<td>40.9</td>
</tr>
<tr>
<td>Asp$^a$</td>
<td>40.1</td>
<td></td>
<td>Val</td>
<td>40.0</td>
</tr>
<tr>
<td>Leu$^c$</td>
<td>39.0</td>
<td></td>
<td>Glu$^b$</td>
<td>39.7</td>
</tr>
<tr>
<td>Val</td>
<td>38.9</td>
<td></td>
<td>Thr</td>
<td>39.7</td>
</tr>
<tr>
<td>Met</td>
<td>38.6</td>
<td></td>
<td>Cys</td>
<td>37.7</td>
</tr>
<tr>
<td>Ile</td>
<td>38.4</td>
<td></td>
<td>Lys$^b$</td>
<td>36.8</td>
</tr>
<tr>
<td>Asn</td>
<td>38.3</td>
<td></td>
<td>His</td>
<td>35.7</td>
</tr>
<tr>
<td>Ser$^d$</td>
<td>38.0</td>
<td></td>
<td>Ser$^e$</td>
<td>35.0</td>
</tr>
<tr>
<td>His</td>
<td>36.5</td>
<td></td>
<td>Asp</td>
<td>34.0</td>
</tr>
<tr>
<td>Thr</td>
<td>36.2</td>
<td></td>
<td>Gly</td>
<td>32.7</td>
</tr>
<tr>
<td>Cys</td>
<td>36.1</td>
<td></td>
<td>Leu$^e$</td>
<td>31.7</td>
</tr>
<tr>
<td>Tyr</td>
<td>34.3</td>
<td></td>
<td>Asn</td>
<td>30.3</td>
</tr>
<tr>
<td>Phe$^c$</td>
<td>33.5</td>
<td></td>
<td>Tyr</td>
<td>30.2</td>
</tr>
<tr>
<td>Gly$^c$</td>
<td>33.2</td>
<td></td>
<td>Phe$^d$</td>
<td>28.3</td>
</tr>
<tr>
<td>Trp</td>
<td>31.9</td>
<td></td>
<td>Trp</td>
<td>26.1</td>
</tr>
</tbody>
</table>

1.3.4 Heterotrimeric Collagen-like Peptides

Several groups have reported heterotrimeric peptide systems, but mainly their triple helices were formed with branched peptides, in which the desired chain compositions and alignment were forced through covalent linkage between peptides as shown in Figure 1.5. However, these approaches required a more complicated synthetic
methodology and the presence of the covalently bound region may have unintended consequences on the structures. Bailey reported making heterotrimers of (Pro-Pro-Gly)$_{10}$ and (Pro-Hyp-Gly)$_{10}$ via non-covalent interactions. They could obtain heterotrimers by mixing the two peptides with different ratios and then separate the heterotrimers from homotrimers by HPLC. However, the system was not optimally controllable.\(^{44}\)

Recently, our group reported a method for preparing heterotrimeric collagen-like helices utilized electrostatic interactions.\(^{45-47}\) We found that mixing three peptides which include a positively charged peptide, a negatively charged peptide, and a neutral peptide resulted in the formation of highly stable ABC heterotrimers. Among the ABC heterotrimers reported, (DOG)$_{10}$*(PKG)$_{10}$*(POG)$_{10}$ was found to have the highest thermal stability with a melting temperature of 65°C.\(^{46}\) Studies on homotrimers suggested that replacement of proline in the X position or hydroxyproline in the Y position of the X-Y-Gly repeat resulted in thermal destabilization\(^{7}\) (except in the cases of several unnatural amino acids, especially fluoroproline\(^{19, 22}\)). In the case of the (DOG)$_{10}$*(PKG)$_{10}$*(POG)$_{10}$ heterotrimer, this was not observed despite 20 such replacements. We believe that the unexpected stability of this system arises from salt-bridged hydrogen bonding between negatively charged aspartate and positively charged lysine.

In this work, we further explore the role of charge-paired hydrogen bonding in heterotrimeric collagen stabilization. First we adopted the host-guest approach to evaluate the tendency of triple helix formation of the amino acid: ornithine (abbreviated as Orn or
Or), and compared it with Lys containing host-guest peptides in the literature. Then we studied the tendency of homotrimeric triple helix formation of the charged peptides: (XOG)$_{10}$ and (PYG)$_{10}$. We found that positively charged homotrimeric helices, which were typically unstable, could be formed in the presence of phosphate anion. This part of the work is discussed in chapter 2. Then we explore the effect of side chain length on charged-pair stabilization. This is done by examining oppositely charged pairs of amino acids which have similar chemical functionality but differing side chain length. Lysine (K) or ornithine (Or) are paired with glutamic acid (E) or aspartic acid (D). The stability was measured for all four possible pairs: K/D, K/E, Or/D and Or/E. Of these, K/D pairs were found to be the most stable and Or/D pairs least stable. Then we examine the effect of salt concentration on the stability of the formed triple helices. If electrostatics plays a major role in the stabilization of these triple helices, increased salt concentration should result in reduced stability. We find that helix stability is reduced but not eliminated. This work is discussed in chapter 3. Finally, we examine the positional effect of the charged amino acids. In natural collagen there is a strong predisposition for negatively charged amino acids to be in the X position and positively charged amino acids to be in the Y position.\textsuperscript{1,32} In the second group of peptide studies we intentionally reverse this predisposition. It is found that while the thermal stability of these systems was similar to those from group 1, the triple helical stability and compositional purity was significantly reduced. This work is discussed in chapter 4. All these triple helical systems were studied by thermal unfolding monitored by circular dichroism (CD) and differential scanning calorimetry (DSC).
1.4 Peptide and System Design

Generally, the design of collagen like peptides includes Gly at every third position. The X and Y positions have to contain a large number of the imino acids, proline and hydroxyproline, to ensure the triple helix formation. When the triple helix is formed, the three peptide chains have to be staggered by one residue from each other in order for each cross section to have one Gly, one residue from the X position of one chain and one residue from the Y position of another chain. The goal was to design some charged collagen-like peptides, which will form stable triple helices by electrostatic interactions of the sidechains. Shown in Table 1.2 are the peptides we designed, synthesized and characterized for this study. Among them, twelve are new peptides. The amino acid, ornithine, which has one less methylene group on its side chain than lysine, was adopted in most of the sequences. We would like to utilize it to study how the side chain length affects the hydrogen bonding formed by charged pairs. Peptides 1, 2 and 3 are host-guest peptides, in which Or is at the X position of peptide 2 and at the Y position of peptide 3 of the guest peptides. Peptide 1 is used as the control. Peptides 5, 7 and 8 all have the charged amino acids at the Y positions with Pro at the X position. Peptides 9-13 all have the charged amino acids at the X positions, with Hyp at the Y position. This design will enable us to apply different combinations, such as (D/K, D/Or or K/D, Or/D) to study not only the effects of the length but also the position of the amino acids to the stability of the triple helices. Peptide 14 is designed for analytical ultracentrifugation (AUC) experiments. For
all of the peptides, the N-termini are acetylated and C-termini are amidated to avoid the charge effects on the stability of the triple helix.

<table>
<thead>
<tr>
<th>#</th>
<th>Triple Helix Composition $^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(POG)$_8$G</td>
</tr>
<tr>
<td>2</td>
<td>(POG)$_3$(POrG)(POG)$_4$G</td>
</tr>
<tr>
<td>3</td>
<td>(POG)$_2$(OrOG)(POG)$_6$G</td>
</tr>
<tr>
<td>4</td>
<td>(POG)$_{10}$</td>
</tr>
<tr>
<td>5</td>
<td>(POrG)$_{10}$</td>
</tr>
<tr>
<td>6</td>
<td>(PKG)$_{10}$</td>
</tr>
<tr>
<td>7</td>
<td>(PDG)$_{10}$</td>
</tr>
<tr>
<td>8</td>
<td>(PEG)$_{10}$</td>
</tr>
<tr>
<td>9</td>
<td>(OrOG)$_{10}$</td>
</tr>
<tr>
<td>10</td>
<td>(KOG)$_{10}$</td>
</tr>
<tr>
<td>11</td>
<td>(EOG)$_{10}$</td>
</tr>
<tr>
<td>12</td>
<td>(DOG)$_{10}$</td>
</tr>
<tr>
<td>13</td>
<td>(ROG)$_{10}$</td>
</tr>
<tr>
<td>14</td>
<td>(PKG)$_{10}$GY</td>
</tr>
</tbody>
</table>

$a.$ N- and C- termini acetylated and amidated respectively.

By utilizing the positively charged, negatively charged and neutral peptides, eight ABC triple helical systems are formed and characterized. They are divided into two groups as shown in Table 1.3. Group 1 is composed of negatively charged collagen-like peptides (XOG)$_{10}$ (X=Glu, Asp), positively charged collagen-like peptides (PYG)$_{10}$ (Y=Lys, Orn) and (POG)$_{10}$. This group places positive and negative charges in the positions most commonly observed in nature. Group 2 is composed of positively charged collagen-like peptides (XOG)$_{10}$ (X=Orn, Lys), negatively charged collagen-like peptides (PYG)$_{10}$ (Y=Glu, Asp) and (POG)$_{10}$. In group 2, the positively and negatively charge amino acids are swapped from their naturally preferred positions.
### Table 1.3. ABC heterotrimeric systems

<table>
<thead>
<tr>
<th>Group 1</th>
<th>X(-)Y(+)</th>
<th>#</th>
<th>Charge in neutral pH</th>
<th>Abbreviation</th>
<th>Peptide composition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A(DOG)B(PKG)C(POG)</td>
<td>1 (DOG)<em>{10}:(PKG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(EOG)B(PKG)C(POG)</td>
<td>1 (EOG)<em>{10}:(PKG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(DOG)B(POrG)C(POG)</td>
<td>1 (DOG)<em>{10}:(POrG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(EOG)B(POrG)C(POG)</td>
<td>1 (EOG)<em>{10}:(POrG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>X(+)Y(-)</td>
<td>#</td>
<td>Charge in neutral pH</td>
<td>Abbreviation</td>
<td>Peptide composition</td>
</tr>
<tr>
<td></td>
<td>A(KOG)B(PDG)C(POG)</td>
<td>1 (KOG)<em>{10}:(PDG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(KOG)B(PEG)C(POG)</td>
<td>1 (KOG)<em>{10}:(PEG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(OrOG)B(PDG)C(POG)</td>
<td>1 (OrOG)<em>{10}:(PDG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A(OrOG)B(PEG)C(POG)</td>
<td>1 (OrOG)<em>{10}:(PEG)</em>{10}:(POG)_{10} = 1:1:1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 1.5 Peptide Synthesis and Purification

All the peptides are N terminally acetylated and C terminally amidated to reduce interfering electrostatic interactions. Peptides were synthesized using solid phase peptide synthesis with standard FMOC chemistry on an Advanced Chemtech Apex 396 peptide synthesizer using a 0.15 mmol scale. The coupling reagents for the synthesis were HBTU, HOBr, and DIEA using 4, 4, and 6 molar equivalents with respect to the amino acid loading. The acetylation was performed with a large excess of acetic anhydride in the presence of DIEA in dichloromethane. The cleavage and deprotection were performed using a mixture of trifluoroacetic acid (TFA): triisopropylsilane : water = 38 : 1 : 1 by volume. Yields were found to be significantly improved by double coupling all prolines. Peptides were purified by reverse-phase high performance liquid chromatography (HPLC) with a C-18 column and characterized by MALDI-TOF MS. HPLC and MS spectra are available in the Appendices.
1.6 Characterization Methods

Three characterization methods were used to understand the factors that affect the stability of the heterotrimeric systems. Circular dichroism (CD) was used to evaluate the triple helicity of the collagen peptide systems and study the thermal unfolding behavior of these systems. Differential scanning calorimetry (DSC) was explored and first used for these systems to obtain the thermodynamic parameters which affect the stability of the triple helices. Analytical ultracentrifugation (AUC) was explored and first used to study these systems to evaluate the composition in the heterotrimeric systems.

1.6.1 Circular Dichroism Spectroscopy (CD)

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light due to the asymmetry of a molecule. "Far-UV" spectra region (190-250 nm) is used to determine the secondary structure of protein. The unique CD spectra of triple helical structure can be easily identified. It is similar to polyproline II spectra. Compared to polyproline II, which has a positive peak at around 228 nm and a negative peak at 206 nm, collagen has its positive peak between 220 nm to 225 nm, negative peak between 200 to 204 nm as shown in Figure 1.6 (a). Thermal unfolding studies usually are performed to differentiate polyproline II from collagen. The temperature of the solution is increased with a very slow rate to be close to its equilibration state, usually at 0.1-0.2 C/min to a certain degree to monitor the change of the ellipticity of the wavelength of the maximum in the CD spectra. The thermal
unfolding study of polyproline II usually is linear, meaning the constant decrease of the optical rotation, while in collagen; the thermal unfolding study shows a cooperative, sharp change of the optical rotation. The sharp transition usually gives the melting temperature \( T_m \), at which half of the triple helix is folded and half is unfolded. Figure 1.6 (b) is a CD thermal unfolding spectra of a collagen peptide in 10 mM phosphate buffer.

![Figure 1.6](image)

**Figure 1.6.** CD of collagen triple helix. (a) CD spectra, and (b) Thermal unfolding spectra.

In this work, CD measurements were obtained from a Jasco J-810 spectrometer with a Peltier temperature control system. A quartz cuvette with a path length of 0.1 cm was used. Peptide solutions in phosphate buffer (10 mM sodium phosphate), PBS buffer (0.15 M NaCl, 10 mM sodium phosphate) or tris buffer (10 mM) were evaluated. Thermal studies were performed by monitoring the decrease of the ellipticity at the wavelength of the maximum of the CD spectra. For evaluating the host-guest peptides, the thermal studies use a 6°C/hour heating rate in the 0°C to 80°C temperature ranges to be consistent with the literature. For all the other cases, the thermal studies use a 10°C/hour heating rate in the
5°C to 95°C temperature ranges. The MRE (Molar Residue Ellipticity) was calculated as:

$$[\theta] = (\theta \times m) / (c \times l \times n_r)$$

in which $\theta$ is the ellipticity in mdeg, $m$ is the molecular weight in g/mol, $c$ is the concentration in mg/ml, $l$ is the path length of the cuvette in mm and $n_r$ is the number of amino acid residues in the peptide. The melting temperature ($T_m$) was obtained as the midpoint of the transition. The minimum of the 1st derivative of the thermal unfolding spectra shows the steepest slope of the transition and is referred as the $T_m$.

### 1.6.2 Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry (DSC) is a method to determine the thermodynamic data for thermally-induced transitions. It is a direct measurement of the intrinsic thermal properties of the system and provides us fundamental thermodynamic information of the process. Figure 1.7 is a typical layout of a DSC instrument. Two identical cells labeled as sample (S) and reference (R) hold protein and buffer solution respectively. The top covers of both cells are capped and held at elevated pressure to prevent bubble formation during heating. The system measures the heat flow that needs to be applied or withdrawn from the sample in order to maintain the same temperature as the reference cell. The output from the DSC experiment is a thermogram of the excess heat capacity ($C_p$, sample minus reference) as a function of temperature. After buffer subtraction, concentration normalization, and baseline correction, the enthalpy of the protein denaturation change can be directly obtained by the integration of the measurement curve which is the change of molar heat capacity verse temperature.
In our work, experiments were performed on a VP-DSC from Microcal LLC. (Northampton, MA). Peptide solution of 0.067 mM in phosphate buffer or PBS buffer was dialyzed against 10 mM phosphate buffer or PBS buffer at 5°C. The buffer solution was changed every 12 hrs for at least three times. The last buffer was used to fill both the sample and the reference cells to obtain the buffer baselines. In the PBS buffer case, the scanning was from 5-80°C at 6°C/hr until at least two subsequent scans were overlapping with each other. In the phosphate buffer case, the scanning was from 10-90°C at 10°C/hr. The last scan was used as the buffer scan. The dialyzed peptide solution was degassed and then loaded into the sample cell during the cooling cycle at around 20°C. The first scan was
used for the data analysis. The subsequent scans showed the high reversibility with about (80%) recovery. Data analysis was performed using the Microcal software with Origin 7 program. For each sample, the appropriate buffer baseline was subtracted from the sample scan and then normalized to its monomer concentration. Next the progress baseline was subtracted. The progress baseline was obtained by estimating the pre- and post-transition baseline, which was calculated by estimating the portion of folded and unfolded materials at certain temperature from the area under the curve using the software the system provided.

In the case when two melting transitions were observed, the peak corresponding to T_m of 68°C was assigned to (POG)_{10} homotrimers. From the individual (POG)_{10} study, we could get the molar heat capacity C_p equaled to 101.08 kcal/mol. The heat capacity of the (POG)_{10} peak in the heterotrimeric mixture was obtained by subtracting the sample data with the buffer baseline and integrating the area of peak at 68°C, here we call it C_p'.

% of (POG)_{10} in the mixture = C_p'(cal) / 101.08 (kcal/mol)
% of newly formed triple helix = 1 - % of (POG)_{10}

The concentration of newly formed triple helix = the corrected heterotrimer concentration based on AAA results * % composition of newly formed triple helix.

Using the calculated concentration of the newly formed triple helix, the subtracted buffer baseline could be normalized. After baseline progression, the enthalpy of the newly formed triple helix at T_m (\Delta H_{cal}^{T_m}) could be obtained by the integration of the area of the peak. At the T_m, \Delta G equals zero. Using the T_m obtained from the maximum of the
endothermic peak, the $\Delta S^{\text{m}}$ was calculated by: $\Delta S^{\text{m}} = \Delta H / T_m$. Here, for simplicity, our thermodynamic results were based on the assumption that $\Delta C_p$ equaled zero. Hence, the $\Delta H$ and $\Delta S$ would be constant in the temperature range (Although from the literature, we know that the $\Delta C_p$ is probably not equal to zero, but a constant with positive sign.\(^{50}\)). From this assumption, standard $\Delta G^0$ was calculated by: $\Delta G^0 = \Delta H - 298 \times \Delta S$. \(^{51}\)

1.6.3 Amino Acid Analysis (AAA)

Since the accurate quantification of sample purity and concentration is one of the most important factors to affect the interpretation of DSC data, Amino acid analysis was used to determine the peptide concentration. The post-DSC sample was saved and sent to the Protein Core Lab (Baylor College of Medicine) for quantitative amino acid analysis. Peptide concentrations were calculated from glycine content.

1.6.4 Analytical Ultracentrifugation (AUC)

Analytical ultracentrifugation was invented by Theodor Svedberg in 1923 to study colloids and proteins. He won the Nobel Prize in chemistry in 1926 for this research. The introduction of a vacuum into the system by Edward Greydon Pickels reduced the friction generated at high speed and enabled one to maintain constant speed, which later on made the commercialization of this instrument possible. \(^{52}\) Ultracentrifugation can be used to either measure the movement of macromolecules in a solution with a very high centrifugal field or to evaluate the distribution of a sample in an equilibrium state with relative lower speed. Two types of experiments can be performed by AUC, sedimentation
velocity and sedimentation equilibrium. Sedimentation velocity is a hydrodynamic technique. Shown in Figure 1.8 (b) is the sedimentation velocity cell setup. It needs a two-section cell, the bottom cell is filled with solvent, and the top cell is filled with sample.

Figure 1.8 (a) shows the sample concentration vs. the radial distance. The concentration of the sample was initially uniform through the cell, after the rotor speeded up (60,000 rpm), the solute molecules move to the bottom of the cell with an appreciable rate. Eventually all the solute molecules will settle to the bottom of the cell. During this process, there is a transition zone, in which there has a concentration distribution. The concentration is different with the distance from the axis of rotor. The transition zone is called the boundary. Shown in Figure 1.9 is the data generally collected from sedimentation velocity experiments. It shows 10 scans collected from a Beckman XL-I using absorbance optical system. With time, the boundaries are moving. Sedimentation velocity is used to collect and analyze the movement of the boundary and get information such as purity,
sedimentation coefficient (s), the diffusion coefficient (D) and the distribution of the sample if it is not pure.\textsuperscript{52}

![Diagram showing sedimentation equilibrium](image)

**Figure 1.9.** Sedimentation velocity data.\textsuperscript{53} (Figure and caption are taken from reference # 53.)

Sedimentation equilibrium is a thermodynamic technique. It uses a relative low speed and measures the concentration distribution at equilibrium. Shown in Figure 1.8 (d) is the sedimentation equilibrium cell setup. It needs a six-channel cell, the top cell is filled with sample, and the bottom cell is filled with the corresponding solvent. Usually they are the same sample with different concentrations. Figure 1.8(c) shows the sample concentration vs. the radial distance of sedimentation equilibrium. With a relatively low speed, the sedimentation of the solute is slow and it is counterbalanced by the diffusion due to the concentration gradient created by the sedimentation. At the beginning, the concentration decreases at the meniscus and increases at the bottom of the cell due to the
sedimentation. Then, the back diffusion at the bottom of the cell opposes the process of sedimentation. Eventually, after certain time, equilibrium is reached and no more concentration changes with time. Shown in Figure 1.10 is the schematic representation of the sedimentation equilibrium.

![Figure 1.10](image)

**Figure 1.10.** Schematic representation of sedimentation equilibrium. The flow of solute due to sedimentation (black arrows) increases with radial distance. This process is balanced at equilibrium by the reverse flow from diffusion (open arrows), which increases with concentration gradient. At equilibrium, the resulting concentration distribution is exponential with the square of the radial position. (Figure and caption are taken from reference # 54.)
References


Chapter 2. Homotrimeric Collagen-like Peptide Systems

2.1 Introduction

After the structure of collagen had been elucidated by Ramachandran and subsequently by Rich and Crick (1950s), homotrimeric collagen-like peptides have been extensively studied to understand the structure and the stability of the collagen peptides. (Pro-Hyp-Gly)$_{10}$ and (Pro-Pro-Gly)$_{10}$ were among them. It was found that with the replacement of proline at the X position with other amino acids, the stability of the triple helices decreased. Thermal studies found that the melting temperature ($T_m$) of (Pro-Pro-Gly)$_{10}$ and (Pro-Hyp$_R$-Gly)$_{10}$ were 41°C and 69°C respectively, which indicated that hydroxyproline at the Y position could stabilize triple helices. However, when the proline was replaced with hydroxyproline at the X position, (Hyp$_S$-Pro-Gly)$_{10}$ could not form a triple helix. These results showed that the conformation and position of amino acids affected the stability of the collagen triple helix. Brodsky’s host-guest peptides provided a tool to evaluate the amino acid propensity for collagen triple helix formation at specific positions (X or Y). It can quantitatively evaluate the triple helical propensity of amino acids. To facilitate our study of heterotrimeric collagen systems, we evaluated the homotrimeric triple helical tendency of each peptide. First, we designed and synthesized three host-guest peptides 1, 2 and 3 as shown in Table 2.1. These peptides were used to evaluate the amino acid propensity of ornithine compared to lysine. Second,
five new peptides with sequences such as $(X\cdot O\cdot G)_{10}$ (X = D, E, Or, or K and $(P\cdot Y\cdot G)_{10}$, and Y = Or, K, D, E) were designed and synthesized. We evaluated the triple helical propensities of these peptides in phosphate buffer at neutral pH. To our surprise, both $(KOG)_{10}$ and $(OrOG)_{10}$ showed weak positive peaks in the range of 220-225 nm, which indicated formation of a triple helix. Since these peptides were positively charged at neutral pH, we did not expect them to form triple helices due to the repulsion of the positively charged amino groups. Very likely the formation of the triple helices was due to the divalent negatively charged phosphate ions in the solution, which acted as a bridge to shield the repulsion of the charged amino groups and link them together. This hypothesis was supported when we used tris buffer to replace the phosphate buffer and no triple helices were observed for both peptides. The other seven peptides did not show triple helix formation by CD studies.

### 2.2 Design of Peptides

As shown in Table 2.1, peptide 1, 2 and 3 were designed to evaluate the amino acid propensity of ornithine. Peptide 1 was used as a control and compared to the model peptide of Brodsky’s host-guest peptide systems. Peptide 2 replaced hydroxyproline by ornithine at the Y position in the guest tri-peptide unit to compare it with its counterpart, in which the Y position was replaced by lysine. Peptide 3 replaced proline by ornithine at the X position in the guest tri-peptide unit to compare it with its counterpart, in which the X position was replaced by lysine. Seven new peptides shown in Table 2.1 were designed
to replace X or Y in the (XYG)$_{10}$ repeat with positively charged amino acid residues such as Asp (D), Glu (E), or negatively charged residues such as Orn (Or), Lys (K) or Arg (R) to facilitate the systematic studies of our heterotrimeric systems.

Table 2.1. Peptide sequence, abbreviation and melting temperature of homotrimers.

<table>
<thead>
<tr>
<th>#</th>
<th>Triple Helix Composition $^a$</th>
<th>T$_m$ (°C) in Phosphate Buffer</th>
<th>T$_m$ (°C) in PBS Buffer</th>
<th>T$_m$ (°C) in Tris Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(POG)$_n$G</td>
<td>b</td>
<td>48.5</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>(POG)$_3$(POrG)(POG)$_n$G</td>
<td>-</td>
<td>37.5</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>(POG)$_3$(OrOG)(POG)$_n$G</td>
<td>-</td>
<td>43.5</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>(POG)$_{10}$</td>
<td>67.5</td>
<td>67.5</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>(POrG)$_{10}$</td>
<td>none</td>
<td>none</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>(PKG)$_{10}$</td>
<td>none</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>(PDG)$_{10}$</td>
<td>-</td>
<td>none</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>(PEG)$_{10}$</td>
<td>-</td>
<td>none</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>(OrOG)$_{10}$</td>
<td>27.5</td>
<td>25.5</td>
<td>none</td>
</tr>
<tr>
<td>10</td>
<td>(KOG)$_{10}$</td>
<td>34$^c$</td>
<td>31$^c$</td>
<td>none</td>
</tr>
<tr>
<td>11</td>
<td>(EOG)$_{10}$</td>
<td>-</td>
<td>none</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>(DOG)$_{10}$</td>
<td>-</td>
<td>none</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>(ROG)$_{10}$</td>
<td>-</td>
<td>none</td>
<td>-</td>
</tr>
</tbody>
</table>

a. N- and C- termini acetylated and amidated respectively. b. Indicates data not collected. c. After 24 hour equilibration.

2.3 Characterization

All peptides in Table 2.1 were evaluated by CD and for peptides that had thermal transitions; the peptides were also evaluated by CD and DSC.

2.3.1 Circular Dichroism (CD)

Host guest peptides 1, 2 and 3: CD spectra and thermal unfolding experiments were performed with 1mg/ml peptides in phosphate buffered saline. For thermal unfolding experiments, the solution of the peptides was incubated at 5°C for over 48 hrs
prior to the thermal unfolding study. The temperature was increased from 5°C to 85°C with a heating rate of 6°C/ hr.

(XYG)₁₀ peptides: CD spectra and thermal unfolding were performed with peptide concentrations of 0.2 mM in 10 mM phosphate buffer. For thermal unfolding, the temperature was increased from 5°C to 95°C with a heating rate of 10°C/ hr.

2.3.2 Differential Scanning Calorimetry (DSC)

Host guest peptides 1, 2 and 3: DSC experiments were performed on a VP-DSC from Microcal. LLC. (Northampton, MA). 0.128 mM Peptide solutions in 10 mM PBS buffer were dialyzed against 10 mM PBS buffer at 5°C, the buffer solution was changed every 12 hours for at least three times. The last buffer was used to fill both the sample and the reference cells to obtain the buffer baselines. The thermal unfolding was performed with 5-85°C of 6°C/hr heating rate. The last scan was used as the buffer scan. The dialyzed peptide solution was degassed and then loaded into the sample cell during the cooling cycle at around 10°C. The first scan was used for the data analysis. Subsequent scans showed high reversibility with about 80% recovery.

2.4 Results and Discussion

Eleven new peptides were synthesized, their propensities of triple helix formation were tested and the corresponding melting temperatures (Tₘ) are listed in Table 2.1.
2.4.1 Host-guest Peptides:

Host-guest peptide 1, 2, 3 were designed to evaluate amino acid ornithine's propensity of triple helix formation. It was completed by placing ornithine at the X or Y position in the guest region and compared them with the corresponding lysine containing host-guest peptides. Shown in Figure 2.1 are the CD spectra, thermal transition and 1st derivative of the thermal transitions of the three host-guest peptides. In Table 2.2, the T_m s of the Lys containing host-guest peptides were compared with the T_m s of the Orn containing host-guest peptides. Two conclusions can be drawn from these data. First,

ornithine has the similar triple helical propensity to lysine. Second, the stabilities of ornithine containing peptides follow the same trend as the stabilities of the lysine containing peptides. When ornithine or lysine replaces proline at the X position, T_m decreases by nearly 5.4°C. When ornithine or lysine replaces hydroxyproline at the Y position, T_m decreases by nearly 11°C. This is in agreement with Brodsky’s observation that replacing hydroxyproline at the Y position with other amino acids had a greater impact on the stability of the triple helix than replacing proline at the X position with
Table 2.2. Melting temperature of Ornithine containing peptides compared to Lysine.

<table>
<thead>
<tr>
<th>#</th>
<th>Peptide</th>
<th>$T_m$ (°C) of Lys Containing Peptides $^5$</th>
<th>$\Delta T_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(POG)$_3$G</td>
<td>48.5</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>(POG)$_3$(POrG)(POG)$_4$G</td>
<td>43.5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>(POG)$_3$(OrOG)(POG)$_4$G</td>
<td>37</td>
<td>1</td>
</tr>
</tbody>
</table>

other amino acids. $^5$ Thermodynamic parameters involved in the thermal unfolding processes of the three ornithine containing host-guest peptides were obtained by DSC experiments. DSC spectra are shown in Figure 2.2 and their thermodynamic data are summarized in Table 2.3. The melting temperatures of the three peptides agreed with those obtained from CD very well. With the replacement of proline at the X position in the guest tripeptide unit, both the enthalpy and entropy decreased. Similarly, with the

Figure 2.2. DSC spectra of host guest homotrimers: (POG)$_3$G (in black), (POG)$_3$(OrOG)(POG)$_4$G (in red), (POG)$_3$(POrG)(POG)$_4$G (in green).
replacement of hydroxyproline at the Y position in the guest tripeptide unit, both the enthalpy and entropy decreased dramatically, which agreed with the known fact that hydroxyproline at the Y position stabilizes the triple helix. ³, ⁷, ⁸

Table 2.3. Thermodynamic parameters obtained from DSC.

| #  | Abbreviation        | \( T_m \) \(^\circ\)C \)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>from calorimetry</td>
<td>from CD</td>
</tr>
<tr>
<td>1</td>
<td>(POG)₃G</td>
<td>49.0</td>
</tr>
<tr>
<td>2</td>
<td>(POG)₃(OrOG)(POG)₄G</td>
<td>42.8</td>
</tr>
<tr>
<td>3</td>
<td>(POG)₃(POrG)(POG)₄G</td>
<td>38</td>
</tr>
</tbody>
</table>

*Thermodynamic results are based on concentration calculated from mass.

2.4.2 (XYG)₁₀ Homotrimers:

2.4.2.1 Positively Charged (XYG)₁₀ Systems:

Unlike in our previous reports that showed charged peptides could not form homotrimers, ⁹ both of the positively charged peptides (OrOG)₁₀ and (KOG)₁₀ were found to form triple helices at neutral pH in phosphate buffer. (OrOG)₁₀ can readily form a triple helix at room temperature, with a thermal unfolding temperature of 27°C, while (KOG)₁₀ only formed a triple helix after incubation at 5°C for over 24 hours with a \( T_m \) of 34°C. The CD spectra, thermal transition and 1ˢᵗ derivative of the thermal transitions of these two peptides are shown in Figure 2.3 and Figure 2.4 respectively. One possible explanation of this observation is that the phosphate anions in the solution interact with the positively charged primary amine of ornithine or lysine. This interaction may quench
Figure 2.3. CD spectra of (OrOG)$_{10}$ homotrimer in PBS and phosphate buffer. (a) Circular Dichroism spectra of (OrOG)$_{10}$ at 5°C. (b) Thermal transition of (OrOG)$_{10}$. (c) 1st derivative of the thermal transition of (OrOG)$_{10}$.

Figure 2.4. CD spectra of (KOG)$_{10}$ homotrimer in PBS buffer before and after incubation at 5°C. (a) Circular Dichroism spectra of (KOG)$_{10}$ at 5°C. (b) Thermal transition of (KOG)$_{10}$ monitored at 218 nm w/o incubation and 222 nm after incubation. (c) 1st derivative of the thermal transition of (KOG)$_{10}$ after five days incubation in PBS buffer.

the charge repulsion and perhaps also bridge the two lysine residues together. To test this, these peptides were prepared in tris buffer, which was positively charged at neutral pH (shown in Figure 2.5). Under similar conditions, neither peptide was observed to form a triple helix (shown in Figure 2.6). This meant that there was a specific interaction between the divalent phosphate anion and the positively charged amino group. This interaction between phosphate and amino was not, however, a general phenomenon as the
Figure 2.5. Comparison of structures of phosphate and Tris buffer. (a) Phosphate buffer is composed of $H_2PO_4^-$ (62%) and $HPO_4^{2-}$ (38%) at neutral pH. (b) Tris: tris(hydroxymethyl)aminomethane, $pK_a = 8.06$.

other amine containing peptides (POrG)$_{10}$ and (PKG)$_{10}$ did not display this behavior (Figure 2.7 and Figure 2.8). This was likely due to the fact that neither of these peptides contained hydroxyproline and this absence significantly destabilized them. Additionally, amino acids in the X position have been shown to be more solvent accessible comparing

Figure 2.6. CD spectra in Tris buffer. (a) (OrOG)$_{10}$ at 5°C, CD spectra were taken with newly prepared and after 72 hrs incubation samples. (b) (KOG)$_{10}$ at 5°C, CD spectra were taken with newly prepared sample, after 24 hr incubation, 48 hrs incubation and 1 week incubation.

to amino acids in the Y position. This may facilitate the interaction of phosphate anions with the amino side chain of ornithine or lysine. Since (KOG)$_{10}$ did not form a triple
helix readily, we performed an interval CD scan experiment to monitor the formation of the triple helix. The instrument was programmed to take a CD spectrum every hour for

![CD spectra of (POrG)₁₀ in phosphate buffer](image1)

**Figure 2.7.** CD spectra of (POrG)₁₀ in phosphate buffer. (a) Circular Dichroism spectra of (POrG)₁₀ at 5°C. (b) Thermal transition of (POrG)₁₀.

![CD spectra of (PKG)₁₀ in phosphate buffer](image2)

**Figure 2.8.** CD spectra of (PKG)₁₀ in phosphate buffer. (a) Circular Dichroism spectra of (PKG)₁₀ at 5°C. (b) Thermal transition of (PKG)₁₀ at 225nm.

60 hours at 5°C. The experiments were performed in both phosphate and PBS buffer. As shown in Figure 2.9, in both of the buffers, the maximum of the spectra showed a graduate distinct peak shift from 218 nm to 222 nm after incubation for 24 hrs. The
thermal unfolding of the incubated sample showed a transition with melting temperature of 34°C in phosphate buffer and 31°C in PBS buffer. Since Arg is also positively charged at neutral pH, (ROG)io was also tested for homotrimeric helix formation. CD spectra (Figure 2.10) showed a weak positive at 220nm. After about 60 hrs, the peak shifted to

![CD spectra](image)

**Figure 2.9.** CD spectra of (KOG)io peptide taken with a 1 hour interval of 60 scans. (a) in 10mM phosphate buffer at pH = 7.4. (b) in 10mM PBS buffer (phosphate buffer with 150mM NaCl) at pH=7.4.

![CD spectra and thermal unfolding](image)

**Figure 2.10.** CD spectra and thermal unfolding of (ROG)io peptide in 10mM phosphate buffer. (a) Thermal unfolding at 220nm. (b) Interval scans taken with a 1 hour interval of 60 scans. (c) Thermal unfolding at 221nm.

221nm, but both of the thermal unfolding at 220nm and 221nm were linear, which indicated that (ROG)io did not form a triple helix at the tested condition. It is still not
clear why (KOG)$_{10}$ and (OrOG)$_{10}$ can form triple helices in phosphate buffer, but
(ROG)$_{10}$ cannot.

2.4.2.2 Negatively Charged (XYG)$_{10}$ Systems:

Since negatively charged phosphate anions can act as bridges to shield the
positively charged amine groups in (KOG)$_{10}$ and (OrOG)$_{10}$, we would like to find out
whether similar effects exist among the negatively charged peptides with divalent cations
to help induce the formation of triple helix. In order to evaluate the effects of positive
cations on the formation of the homotrimeric triple helices of (EOG)$_{10}$ peptide, Tris
buffer was used to avoid the interference of phosphate anions. We tested (EOG)$_{10}$ with
Ca$^{2+}$, Ni$^{2+}$ and Cu$^{2+}$. CD spectra of peptide (EOG)$_{10}$ in 10 mM phosphate buffer at
neutral pH had a weak positive peak at 218nm. It was noticed that no peak shift was
observed after 72 hours incubation at 5°C, which was consistent with the results after 6
more days’ incubation. The thermal unfolding monitored at 218 nm showed a linear
transition (Figure 2.11 and Figure 2.12). In the case of Ca$^{2+}$, the ratio of peptide to Ca$^{2+}$
of 1 to 5 and 1 to 10 were tested. Figure 2.12 showed CD spectra of (EOG)$_{10}$ in tris buffer
with and without Ca$^{2+}$. As we can see, after incubation of three and seven days, no
peak shift was observed. Figure 2.13 are the CD spectra of (EOG)$_{10}$ peptide with Ni$^{2+}$
and Cu$^{2+}$. Similar to Ca$^{2+}$, (EOG)$_{10}$ did not show a peak shift or sign of triple helix
formation with either Ni$^{2+}$ or Cu$^{2+}$. And in the case of Cu$^{2+}$, precipitation was observed
and the CD spectra were taken from the supernatant. The remainders of the negatively
Figure 2.11. CD spectra and thermal unfolding of (EOG)$_{10}$ peptide in 10mM phosphate buffer. (a) CD spectra of (EOG)$_{10}$ peptide has a maximum at 218nm. (b) Thermal unfolding at 218 nm.

Figure 2.12. CD spectra (EOG)$_{10}$ peptide in 10mM Tris buffer. (a) CD spectra of (EOG)$_{10}$ peptide after 3 days and 7 days incubation. (b) CD spectra of (EOG)$_{10}$ peptide after 3 days and 7 days incubation with Ca$^{2+}$ (1:5).

charged peptides were not observed to form homotrimeric triple helices under any tested conditions. The CD spectra, thermal transitions of all the other three peptides are show in Figures 2.14-2.16.
Figure 2.13. CD spectra (EOG)$_{10}$ peptide in 10mM Tris buffer with cations. (a) (EOG)$_{10}$ peptide with Cu$^{2+}$ (1:5) after 1 day incubation. (b) (EOG)$_{10}$ peptide with Ni$^{2+}$ (1:10) after 1 day incubation.

Figure 2.14. CD spectra of (DOG)$_{10}$ in PBS buffer. (a) Circular Dichroism spectra of (DOG)$_{10}$ at 5 °C. (b) Thermal transition of (DOG)$_{10}$.

2.5 Conclusion

In this chapter, homotrimeric collagen-like peptides were studied to facilitate the heterotrimeric systems that will be discussed in later chapters. First, we studied ornithine containing host-guest collagen-like peptides to evaluate the amino acid propensity of
ornithine at the X or Y position by comparing them with their counterpart of lysine containing similar peptides. The $T_m$ of ornithine containing peptides showed a similar trend as compared to the lysine containing host-guest peptides. Both CD and DSC studies showed that ornithine had a similar triple helical propensity to lysine. The stabilities of ornithine containing peptides followed the same trend as the stabilities of the lysine containing peptides. The reason that peptide 3 was more stable than peptide 2 was due to
the known fact that hydroxyproline at the Y position stabilizes triple helix. Next, we studied (XYG)$_{10}$ homotrimeric collagen like peptides. To our surprise, and unlike in our previous reports that showed the charged peptides could not form homotrimers, both the positively charged peptides (OrOG)$_{10}$ and (KOG)$_{10}$ were found to form triple helices at neutral pH in 10 mM phosphate buffer. We showed that phosphate ions in the buffer could be the major driving force to induce the formation of the charged peptide to form triple helices by testing the system in tris buffer which did not have divalent anions. Negatively charged peptide (EOG)$_{10}$ was also tested by adding cations such as Ca$^{2+}$, Ni$^{2+}$ and Cu$^{2+}$, and none of them showed formation of triple helices. Other peptides such as (ROG)$_{10}$, (PEG)$_{10}$, (PDG)$_{10}$ did not form triple helices at the tested condition.
References


Chapter 3. Study of Heterotrimeric Collagen-like Peptide Systems: Group 1 ($A_{X(-)}B_{Y(+)}C_{O}$)

3.1 Introduction

Examinations of homotrimeric and heterotrimeric collagen-like peptides are equally important to understand the structure and stability of native collagen. However, studies have been focused on homotrimers, mainly because they are easy to form stable triple helices. Attempts to study heterotrimers are limited due to the difficulty of making stable triple helices. There were some reported stable heterotrimers, but most of the triple helices were formed by branched peptides, in which the desired chain composition and alignment were forced through covalent linkage between peptides. Although these approaches yielded desired staggering and very stable heterotrimeric triple helices, they required more complicated synthetic methodologies and the presence of the covalently bound region might have unintended consequences on the structure. Recently our group reported a method for preparing heterotrimeric collagen-like helices which utilized electrostatic interactions to organize their self-assembly. We found that mixing three peptides that included a positively charged peptide, a negatively charged peptide, and a neutral peptide resulted in the formation of highly stable ABC heterotrimers. Among the ABC heterotrimers reported, ($DOG)_{10}^*(PKG)_{10}^*(POG)_{10}$ was found to have the highest thermal stability with a melting temperature of 65°C. Studies on homotrimers suggested
that the replacement of proline from the X position or hydroxyproline from the Y position of the X-Y-Gly repeat should result in thermal destabilization \(^{10}\) (except in the cases of several unnatural amino acids, especially fluoroproline \(^{11,12}\)). In the case of the \((\text{DOG})_{10}*(\text{PKG})_{10}*(\text{POG})_{10}\) heterotrimer, this was not observed despite 20 such replacements. We believe that the unexpected stability of this system arises from salt-bridged hydrogen bonding between the negatively charged aspartate and the positively charged lysine.

When three peptides (one positively charged, one negatively charged, and one neutral under neutral pH) are mixed together, there are ten possible combinations of triple helices. \(^9\) For each A•B•C triple helical system, we systematically studied the three possible homotrimers (A•A•A, B•B•B, C•C•C) by examining each peptide in isolation. Next we examined the six possible binary heterotrimers (A•A•B, A•A•C, B•B•A, B•B•C, C•C•A, C•C•B) by mixing pairs of peptides. Only after completing this survey we are able to mix all three peptides and make conclusions about the composition of the three way mixture. Thermal transitions that do not correspond to the homotrimers or binary heterotrimers can be assigned to an A•B•C heterotrimer. Ideally most of the homotrimers and binary heterotrimers either do not form or have low stability. When this happens, A•B•C heterotrimers are easy to identify and characterize.

Group 1 are heterotrimers composed of \([\text{X(-)OG}]_{10}, [\text{PY(+)}\text{G}]_{10}\) and \((\text{POG})_{10}\) as shown in Table 3.1. Heterotrimetric systems \((\text{EOG})_{10}*(\text{PKG})_{10}*(\text{POG})_{10}\) and \((\text{DOG})_{10}*(\text{PKG})_{10}*(\text{POG})_{10}\) had been studied and reported before. \(^8\) We studied the other
two heterotrimeric systems in a similar way. Each of the heterotrimeric system can possibly form ten triple helices, three homotrimers, six AAB type binary heterotrimers and one ABC heterotrimer. The homotrimeric study of each of the individual peptide showed that only \((\text{POG})_{10}\) form triple helices under neutral pH. And its \(T_m\) was distinguishable. 

Due to the limitation of our analytical system, we could not differentiate between an AAB heterotrimer and ABB heterotrimer. Transitions assigned to this type of heterotrimer might be pure AAB or pure ABB or a mixture of both. For the remainder of this work we simply call these systems AAB despite the ambiguity acknowledged above. The AAB types of heterotrimers were evaluated by mixing the respective peptides together in 1:1 ratio and observing their thermal transition with and without preheating. In the case of preheating, we mixed the two peptides together, increased temperature to 85°C, and then slowly cooled it to allow the disrupted molecules to rearrange to their most thermodynamically stable structure. Then we studied their thermal unfolding behavior by CD and DSC. In all evaluated systems, the AAB helix displayed a lower \(T_m\) as comparing to the ABC heterotrimers. After combining and evaluating all the possibilities, we could identify the ABC heterotrimers for most of the systems.

<table>
<thead>
<tr>
<th>#</th>
<th>Charge in neutral pH</th>
<th>Abbreviation</th>
<th>Peptide composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>X(-)Y(+)</td>
<td>(A_{(\text{DOG})B_{(\text{PKG})C_{(\text{POG})}}})</td>
<td>((\text{DOG})<em>{10};(\text{PKG})</em>{10};(\text{POG})_{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A_{(\text{EOG})B_{(\text{PKG})C_{(\text{POG})}}})</td>
<td>((\text{EOG})<em>{10};(\text{PKG})</em>{10};(\text{POG})_{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A_{(\text{DOG})B_{(\text{POG})C_{(\text{POG})}}})</td>
<td>((\text{DOG})<em>{10};(\text{POG})</em>{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(A_{(\text{EOG})B_{(\text{POG})C_{(\text{POG})}}})</td>
<td>((\text{EOG})<em>{10};(\text{POG})</em>{10}=1:1:1)</td>
</tr>
</tbody>
</table>
3.2 Characterization Methods

Three characterization methods were implemented for all four systems to understand the factors that affected the stability of the heterotrimeric systems. Circular dichroism (CD) was used to evaluate the triple helicity of the collagen peptide systems. Thermal unfolding studies by CD helped us to identify the ABC heterotrimer species and learned about their thermal stabilities. Thermal unfolding studies by differential scanning calorimetry (DSC) were used to obtain the thermodynamic parameters which related to the stability of the triple helices. Analytical ultracentrifugation (AUC) was used to evaluate the composition of the heterotrimeric systems.

3.2.1 Circular Dichroism (CD)

AB binary systems: stock aqueous solution of 2 mM of peptide A and B were mixed with 1:1 ratio and led to a final concentration of the total peptide in the solution of 0.2 mM in 10 mM sodium phosphate buffer or PBS buffer.

ABC systems: stock aqueous solution of 2 mM of peptide A, B and C were mixed with 1:1:1 ratio and led to a final concentration of the total peptide in the solution of 0.2 mM in 10 mM sodium phosphate buffer or PBS buffer.

In the case of non-preheating studies, CD spectra were taken right after the solution was prepared. If the CD spectra showed a positive peak between 220-225 nm, a thermal unfolding study would be followed. About 450 ul of peptide solution in a 0.1cm path length of quartz cuvette was used for the study. The solution was incubated at 5°C for 15 minutes
and then with a 10°C/hr heating rate, the temperature was increased to 95°C while monitoring the change of the ellipticity of the maximum peak at around 225 nm. The 1st derivative was obtained, and the minimum of the peak was referred as the melting temperature (T_m).

In the case of preheating, samples were heated to 85°C, slowly cooled to room temp or 5°C with a cooling rate of 1°C/min, and then incubated at the final temperature for at least 12 hours before the thermal unfolding study.

3.2.2 Differential Scanning Calorimetry (DSC)

All DSC experiments were performed using preheated samples of ABC systems. an aliquot of 2 mM of peptide stock solution of A, B and C were mixed with 1:1:1 ratio and led to a final concentration of the total peptide in the solution of 0.2 mM in 10 mM sodium phosphate buffer. The sample was heated in a water bath at about 90°C for 30 minutes. Then it was cooled to room temperature and incubated for about 12 hours. Next it was loaded to a dialysis cassette (Thermo Scientific, 2K MWCO) and dialyzed against 10 mM sodium phosphate buffer. The buffer was changed every 12 hours for three times. The last buffer was saved and it was used to fill both the sample and the reference cells to obtain the buffer baseline. After three buffer scans (if the last two scans were overlapping with each other), about 0.5 ml degassed dialyzed sample was loaded into the sample cell to start the thermal unfolding study during the cooling cycle at around 15°C. The temperature was increased from 10 to 90°C with a 10°C/hr heating rate. In most cases, the subsequent scans showed high reversibility with about 80% recovery. The last buffer scan was
subtracted from the first sample scan to eliminate any interference of the buffer. Then the concentration was normalized with monomer peptide concentration obtained from amino acid analysis, which was done after the DSC experiments. After the progression of the baseline, enthalpy was obtained by integration of the area above the baseline. In the cases of two transitions of the heterotrimeric systems, the thermodynamic parameters were obtained as described in chapter 1.

3.2.3 Analytical Ultracentrifugation (AUC)

In order to cover a wide range of peptide concentration, a new peptide was designed to incorporate an aromatic amino acid to the peptide chain without interrupting the triple helical structures. Studies have shown that with tyrosine at the C terminus, the triple helical structure was not interrupted. The sequence of the new peptide was (PKG)$_{10}$GY that had a sequence similar to (PKG)$_{10}$ with a glycine as spacer and tyrosine at the C terminus. Before the AUC studies, CD experiments were performed to ensure the introduction of tyrosine in the sequence did not destroy the triple helical structure. Two systems were tested. One was ABC system: (EOG)$_{10}$•(PKG)$_{10}$GY•(POG)$_{10}$. And the other was binary system: (EOG)$_{10}$•(PKG)$_{10}$GY. The comparison of CD spectra of the two systems with the similar systems without tyrosine is shown in Figure 3.1. $T_m$ of the four systems are summarized in Table 3.2. With the introduction of tyrosine into the system, the triple helical structures were not interrupted and the thermal stability of the systems was maintained.
Figure 3.1. CD thermal unfolding study of binary and ABC systems with (PKG)$_{10}$GY. (a) Thermal unfolding spectra comparison. (b) 1$^{st}$ derivative of the thermal unfolding of (a).

Table 3.2. Comparison of $T_m$s from CD thermal unfolding studies.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Peptide composition</th>
<th>$T_m$ (°C) in Phosphate Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>EK</td>
<td>(EOG)$<em>{10}$:(PKG)$</em>{10}$=1:1</td>
<td>37</td>
</tr>
<tr>
<td>EKy</td>
<td>(EOG)$<em>{10}$:(PKG)$</em>{10}$GY=1:1</td>
<td>36</td>
</tr>
<tr>
<td>EKO</td>
<td>(EOG)$<em>{10}$:(PKG)$</em>{10}$:(POG)$_{10}$=1:1:1</td>
<td>55</td>
</tr>
<tr>
<td>EKyO</td>
<td>(EOG)$<em>{10}$:(PKG)$</em>{10}$GY:(POG)$_{10}$=1:1:1</td>
<td>55</td>
</tr>
</tbody>
</table>

Sedimentation velocity and sedimentation equilibrium experiments were performed on a Beckman Optima XL-A system with an An-60-Ti rotor equipped with absorbance optics. Both data and Monte Carlo analyses were performed with UltraScan version 9.9 (for windows). In all AUC experiments, 10 mM sodium phosphate buffer at pH 7.4 was used as the solvent in the reference channels. Partial specific volume of all samples was estimated from the peptide sequences using UltraScan.

Sedimentation velocity experiments used double-sector aluminum centerpieces with sample absorbance in the range of 0.9-0.4 OD at 274 nm or 230 nm. The system was spun with 60,000 rpm, and the temperature was kept at 5 °C through the whole run.
Sedimentation equilibrium experiments used six-channel epon/charchol centerpiece. Multiple loading concentrations of peptides in the range from 0.1-0.9 OD at 230 nm and 274 nm were used to ensure a large loading concentration of the peptides. Five speeds were used based on the estimated molecular weight of corresponding monomer and trimers with the δ value of 1-4 using UltraScan software.

3.2.3.1 (DOG)_{10}:(PKG)_{10}GY:(POG)_{10}

Sedimentation velocity was performed with 60,000 rpm at 5°C with a radial continuous mode and 0.001 cm step size. Absorbance of sample of 0.75 OD at 274 nm was tested. Enhanced Van Holde-Weischet method implemented in UltraScan was used for the data analysis.

Sedimentation equilibrium experiments for this system were performed as following: nine samples were loaded in the 3 cells each with 3 channels. Six samples of absorbance of 0.1 to 0.8 OD at 274 nm and 3 samples with absorbance of 0.25, 0.47 and 0.75 OD at 230 nm were loaded to the 9 channels. Five speeds (40k, 45K, 50K, 55K and 60K rpm) of the rotor were used and scans at 230 nm and 274 nm were collected in radial step mode with 0.001 cm step size.

3.2.3.2 (DOG)_{10}:(PKG)_{10}GY

Sedimentation velocity was performed with 60,000 rpm at 5°C of a radial continuous mode with 0.001 cm step size. Absorbance of sample of 0.66 OD at 274 nm was tested. 150 scans were collected. Enhanced Van Holde-Weischet method implemented in Ultrascan was used for the data analysis.
3.3 Result and Discussion

3.3.1 (DOG)$_{10}$*(POrG)$_{10}$*(POG)$_{10}$

From the homotrimeric study, we knew that (DOG)$_{10}$ did not form a triple helix and its thermal transition at 225 nm was linear. The T$_{m}$ of (POG)$_{10}$ is around 68°C. The thermal transition studies of (POG)$_{10}$ and (POrG)$_{10}$ mixture with heating and without heating only showed the transition of (POG)$_{10}$ homotrimer as shown in Figure 3.2. It indicated that

![Figure 3.2. A(POG)B(POrG) in phosphate buffer. (a) Thermal transition. (b) 1st derivative of the thermal transition.](image)

(POG)$_{10}$ and (POrG)$_{10}$ did not interact with each other. The thermal transition studies of (DOG)$_{10}$ and (POrG)$_{10}$ mixture with heating and without heating showed that the two peptides formed an AAB type of triple helix with a T$_{m}$ of about 19°C shown in Figure 3.3. It was reported that peptides (DOG)$_{10}$ and (POG)$_{10}$ mixture did not interact with each other.

Figure 3.4 are the thermal transitions of the (DOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture. Without heating, the thermal study showed two well-separated transitions, one
corresponded to the T_m (19°C) of the AAB heterotrimer and the other corresponded to the T_m of (POG)_10 homotrimer. After heating and allowing time for incubation, one new

\[ \text{Figure 3.3. } A_{(DOG)B(POG)} \text{ in phosphate buffer. (a) Thermal transition. (b) 1st derivative of the thermal transition} \]

transition was emerged with a T_m of 47.5°C and a very weak transition at 68°C. We think that the new transition with T_m of 47.5°C corresponded to the transition of the A_{(DOG)B(POG)C(POG)} heterotrimer. The thermal transition data are summarized in Table 3.3.

\[ \text{Figure 3.4. CD of collagen peptide } A_{(DOG)B(POG)C(POG)} \text{ in phosphate buffer. (a) Thermal transition. (b) 1st derivative of the thermal transition.} \]
Table 3.3. Possible heterotrimers in (DOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture

<table>
<thead>
<tr>
<th>Type of Heterotrimer</th>
<th>Peptides</th>
<th>Phosphate buffer</th>
<th>PBS buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Thermal History</td>
<td>$T_m$ (°C)</td>
</tr>
<tr>
<td>ABC</td>
<td>$A_{(DOG)B_{(POrG)}}C_{(POG)}$</td>
<td>Preheat w/o heat</td>
<td>47.5(s), 69.5(w)</td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>$A_{(DOG)B_{(POrG)}}$</td>
<td>Preheat w/o heat</td>
<td>19</td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>$A_{(POrG)B_{(POG)}}$</td>
<td>Preheat w/o heat</td>
<td>67</td>
</tr>
</tbody>
</table>

3.3.2 (EOG)$_{10}$*(POrG)$_{10}$*(POG)$_{10}$

From the homotrimeric study, we knew that (EOG)$_{10}$ and (POrG)$_{10}$ did not form homotrimeric helices as their thermal transitions at 225nm were linear. The $T_m$ of (POG)$_{10}$ is around 68°C. The thermal transition studies of (POG)$_{10}$ and (POrG)$_{10}$ mixture with heating and without heating only showed the transition of (POG)$_{10}$. The thermal transition studies of (EOG)$_{10}$ and (POrG)$_{10}$ mixture with heating and without heating showed that the two peptides form AAB type of triple helix with a $T_m$ of 28°C shown in Figure 3.5. It was reported that the (EOG)$_{10}$ and (POG)$_{10}$ mixture did not interact with each other.  

Figure 3.6 shows the thermal transition of the (EOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture. Without heating, the thermal study showed two well-separated transitions, one corresponded to the $T_m$ (28°C) of the AAB heterotrimer and the other corresponded to the $T_m$ of (POG)$_{10}$ homotrimer. After heating and allowing time for incubation, one new transition was emerged with $T_m$ of 46.5°C. We believed it was the $A_{(EOG)B_{(POrG)}}C_{(POG)}$ heterotrimer. The thermal transition data are summarized in Table 3.4.
Figure 3.5. $A_{(EOG)}B_{(POG)}$ in phosphate buffer (a) Thermal transition. (b) 1$^{st}$ derivative of the thermal transition.

Figure 3.6. CD Thermal unfolding of $A_{(EOG)}B_{(POG)}C_{(POG)}$ in phosphate buffer. (a) Thermal transition. (b) 1$^{st}$ derivative of the thermal transition.

Table 3.4. Possible heterotrimers in (EOG)$_{10}$, (POrG)$_{10}$ and (POG)$_{10}$ peptide mixture.

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Phosphate buffer</th>
<th>PBS buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thermal History</td>
<td>$T_m$ (°C)</td>
</tr>
<tr>
<td>Heterotrimer (ABC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{(EOG)}B_{(POG)}C_{(POG)}$</td>
<td>Preheat</td>
<td>46.5(s), 42</td>
</tr>
<tr>
<td></td>
<td>w/o heat</td>
<td>28, 69.5(w)</td>
</tr>
<tr>
<td></td>
<td>28(s),</td>
<td></td>
</tr>
<tr>
<td>ABB(AAB) $A_{(EOG)}B_{(POG)}$</td>
<td>Preheat</td>
<td>51(w)</td>
</tr>
<tr>
<td></td>
<td>w/o heat</td>
<td>27.5</td>
</tr>
</tbody>
</table>
3.3.3 Effects of Different Ratio of \((\text{POrG})_{10}\) on the ABC System

We noticed that in the \((\text{EOG})_{10} \cdot (\text{POrG})_{10} \cdot (\text{POG})_{10}\) system, CD spectra of the preheated sample showed two peaks, one was \(A_{(\text{EOG})}B_{(\text{POrG})}C_{(\text{POG})}\) peak, the other was \((\text{POG})_{10}\) peak. Since we used 1:1:1 ratio, if we still had free \((\text{POG})_{10}\) that could form homotrimers in the solution, it meant that we could have residual \((\text{EOG})_{10}\) or \((\text{POrG})_{10}\) in the solution as well. And one of them would be the limiting reagent. As we used mass to calculate the concentration, it may not be very accurate and it was possible that one of the reagents was less than the other two. Here, we assumed that \((\text{POrG})_{10}\) was the limiting reagent. If we increase the amount of \((\text{POrG})_{10}\) in the solution, it should be able to use the free \((\text{EOG})_{10}\) residue and \((\text{POG})_{10}\) in the solution and form the \(A_{(\text{EOG})}B_{(\text{POrG})}C_{(\text{POG})}\) heterotrimer. Hence, the peak of \(A_{(\text{EOG})}B_{(\text{POrG})}C_{(\text{POG})}\) transition should increase and the peak of \((\text{POG})_{10}\) transition should decrease. When all of the free \((\text{EOG})_{10}\) residue were used up, the areas of the peak of \(A_{(\text{EOG})}B_{(\text{POrG})}C_{(\text{POG})}\) and \((\text{POG})_{10}\) would not change any more with addition of \((\text{POrG})_{10}\). We tested it by doubling the amount of \((\text{POrG})_{10}\) and tripling the amount of \((\text{POrG})_{10}\) in the system and used DSC to monitor the change of the two peaks (shown in Figure 3.7). As we expected, when the amount of \((\text{POrG})_{10}\) in the system was increased, the \(A_{(\text{EOG})}B_{(\text{POrG})}C_{(\text{POG})}\) peak intensity increased and the intensity of peak \((\text{POG})_{10}\) decreased. Table 3.5 summarized the data. Two conclusions can be drawn from this. One is that \((\text{POrG})_{10}\) is the limiting reagent. The other is that concentration from mass is not very accurate. This is the reason that why we need to use amino acid analysis result to calculate the concentration of peptides for the thermodynamic studies.
3.3.4 Effect of Salts on Helix Stability

In our heterotrimeric systems, one of the major driving forces to form ABC heterotrimeric collagen systems was the electrostatic interactions of the ion pairs formed. 

### Table 3.5. Comparison of the area of \( A_{\text{EOG}B\text{POG}C\text{POG}} \) peak with different ratio of \( \text{POG}_{10} \) from DSC spectra.

<table>
<thead>
<tr>
<th>Ratio of ( \text{POG}_{10} )</th>
<th>Peak</th>
<th>Area</th>
<th>( T_m )</th>
<th>( A_{\text{EOG}B\text{POG}C\text{POG}} ) / ( \text{POG}_{10} ) Peak</th>
<th>Area</th>
<th>( T_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{EOG}<em>{10} : \text{POG}</em>{10} : \text{POG}_{10} = 1:1:1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area=23674.37</td>
<td>Area=6094.50</td>
<td>3.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_m =46.34 )</td>
<td>( T_m =67.35 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{EOG}<em>{10} : \text{POG}</em>{10} : \text{POG}_{10} = 1:2:1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area=35603.32</td>
<td>Area=3637.18</td>
<td>9.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_m =46.23 )</td>
<td>( T_m =68.17 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{EOG}<em>{10} : \text{POG}</em>{10} : \text{POG}_{10} = 1:3:1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area=39223.86</td>
<td>Area=2713.38</td>
<td>14.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_m =46.25 )</td>
<td>( T_m =68.43 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
by negatively charged amino acids and positively charged amino acids. To further prove
the existence of this ion pair interaction and its effects on the stabilities of the series of
ABC heterotrimeric collagen-like peptides, we increased the concentration of salts by
adding sodium chloride to the phosphate buffer and then compared the stabilities of the
heterotrimers in 10 mM phosphate buffer and PBS (phosphate buffer saline) which was
composed of 10 mM sodium phosphate buffer in addition of 150 mM NaCl. Shown in
Figure 3.8-3.11 are the CD comparison of the Group 1 ABC helices in phosphate buffer
and PBS. With the addition of NaCl in phosphate buffer, except
(DOG)_{10}^{*}(PorG)_{10}^{*}(POG)_{10}
system, both the peak intensities and T_{m} for the other three
systems decreased. Table 3.6 shows the comparison of the thermal transitions of the four
systems. While the T_{m} of (POG)_{10} homotrimer was unaffected, the T_{m} decreased for all
ABC systems in the presence of NaCl. The addition of NaCl was also found to have a

![Figure 3.8. A_{1}(DOG)B_{1}(PKG)C_{(POG)} thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.](image)
Figure 3.9. $A_{(EOG)B(PKG)C(POG)}$ thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

detrimental impact on the compositional purity. In phosphate buffer, the ABC triple helices in $(EOG)_{10}(PKG)_{10}(POG)_{10}$ system showed only a single transition, after adding salt, in PBS buffer, the intensity of the ABC helix in the system decreased and a significant quantity of $(POG)_{10}$ homotrimer was found. $T_m$ of ABC helix of $(EOG)_{10}(PKG)_{10}(POG)_{10}$ decreased by 6.0°C. The salt affected the other two systems of

Figure 3.10. $A_{(EOG)B(POG)C(POG)}$ thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.
Figure 3.11. $A_{(DOG)B_{(POG)C_{(POG)}}}$ thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

(DOG)$_{10}$*(POG)$_{10}$*(POG)$_{10}$ and (EOG)$_{10}$*(POG)$_{10}$*(POG)$_{10}$ in the same fashion, with $T_m$ decreasing by 3.0°C and 4.0°C respectively. This could be simply explained by the fact that the stability of the (POG)$_{10}$ homotrimer was not decreased by the addition of salts while the ABC heterotrimers were. The exception to this was (DOG)$_{10}$*(PKG)$_{10}$*(POG)$_{10}$ which was found to be pure even in the presence of NaCl with the $T_m$ decreasing by 4.5°C.

Table 3.6. Summary of the $T_m$ of ABC heterotrimers in phosphate buffer and PBS.

<table>
<thead>
<tr>
<th>Triple Helix Composition</th>
<th>$T_m$(°C) in Phosphate Buffer</th>
<th>$T_m$(°C) in PBS</th>
<th>Thermal destablization</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 • (POG)$_{10}$</td>
<td>67.5</td>
<td>67.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DOG)$<em>{10}$*(PKG)$</em>{10}$*(POG)$_{10}$</td>
<td>65</td>
<td>60.5</td>
<td>4.5</td>
</tr>
<tr>
<td>(EOG)$<em>{10}$*(PKG)$</em>{10}$*(POG)$_{10}$</td>
<td>57</td>
<td>51, 69</td>
<td>6.0</td>
</tr>
<tr>
<td>(DOG)$<em>{10}$*(POG)$</em>{10}$*(POG)$_{10}$</td>
<td>47.5, 68</td>
<td>44.5, 68.5</td>
<td>3.0</td>
</tr>
<tr>
<td>(EOG)$<em>{10}$*(POG)$</em>{10}$*(POG)$_{10}$</td>
<td>46, 69.5</td>
<td>42, 67</td>
<td>4.0</td>
</tr>
</tbody>
</table>

3.3.5 Thermodynamic Analysis

DSC measurements of the four ABC collagen-like peptide systems were performed under the same conditions as used for CD analysis. Only pre-heated samples were examined. The thermodynamic parameters were determined by direct integration of the
DSC curve. Figure 3.12 showed the DSC thermal unfolding curves for group 1 helices including \((\text{POG})_{10}\). Our thermodynamic results were based on the simplification that the difference of the heat capacity of the folded and unfolded peptides \(\Delta C_p\) was zero. Hence, the \(\Delta H\) and \(\Delta S\) would be constant with respect to temperature. Shown in Table 3.7 is the summary of the \(\Delta H\), \(\Delta S\) and \(\Delta G^\circ\) of unfolding for the five systems. For all the systems, the \(T_m\) from DSC matched very well with the \(T_m\) from CD. From the molecular simulations we observed that for all helices in group 1, the distance of the hydrogen bonding formed by the salt bridges were maintained at approximately 2.7\(\text{Å}\). The enthalpy of \((\text{DOG})_{10}*(\text{PKG})_{10}*(\text{POG})_{10}\) was much higher than that found in the other heterotrimeric systems, which was in agreement with its high melting temperature. Our expectation was that the replacement of the conformationally restricted, cyclic imino acids proline and hydroxyproline with more flexible, charged, amino acids that can form electrostatic
hydrogen bonds would have simultaneously increased the $\Delta H$ and $\Delta S$, but this was not observed to be the case. With the exception of $(\text{DOG})_{10}^* (\text{PKG})_{10}^* (\text{POG})_{10}$ whose values were nearly the same as $3^* (\text{POG})_{10}$, all the triple helices were found to have a $\Delta H$ and $\Delta S$ lower than $(\text{POG})_{10}$ homotrimer.

### Table 3.7. Thermodynamic parameters obtained from DSC (in phosphate buffer)

<table>
<thead>
<tr>
<th>Triple Helix Composition</th>
<th>$T_m$ (°C) from DSC</th>
<th>$T_m$ (°C) from CD</th>
<th>$\Delta H_{\text{cal}}$ (kcal/mol)</th>
<th>$\Delta S$ (cal/mol/K)</th>
<th>$\Delta G^\circ$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\text{POG})<em>{10}^* (\text{POG})</em>{10}^* (\text{POG})_{10}$</td>
<td>67.7</td>
<td>67.5</td>
<td>101</td>
<td>297</td>
<td>12.7</td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\text{DOG})<em>{10}^* (\text{PKG})</em>{10}^* (\text{POG})_{10}$</td>
<td>64.4</td>
<td>65</td>
<td>102</td>
<td>302</td>
<td>11.9</td>
</tr>
<tr>
<td>$(\text{EOG})<em>{10}^* (\text{PKG})</em>{10}^* (\text{POG})_{10}$</td>
<td>56.5</td>
<td>57</td>
<td>85.5</td>
<td>260</td>
<td>8.2</td>
</tr>
<tr>
<td>$(\text{DOG})<em>{10}^* (\text{PorG})</em>{10}^* (\text{POG})_{10}$</td>
<td>48.8</td>
<td>47.5</td>
<td>55.7</td>
<td>173</td>
<td>4.1</td>
</tr>
<tr>
<td>$(\text{EOG})<em>{10}^* (\text{PorG})</em>{10}^* (\text{POG})_{10}$</td>
<td>46.3</td>
<td>46</td>
<td>61.8</td>
<td>194</td>
<td>4.1</td>
</tr>
</tbody>
</table>

#### 3.3.6 Using Analytical Ultracentrifugation to Study the Composition of the Heterotrimeric Systems

**3.3.6.1 $(\text{DOG})_{10}:(\text{PKG})_{10}GY:(\text{POG})_{10}$**

Sedimentation velocity experiment was used to evaluate the heterogeneity of the system. Shown in Figure 3.13 is the 2-Dimensional spectrum analysis of friction ratio ($f/f_0$)

![Figure 3.13](image)

**Figure 3.13** Sedimentation velocity data analysis of $(\text{DOG})_{10}:(\text{PKG})_{10}GY:(\text{POG})_{10}$ system. (a) 2-D spectrum analysis. (b) Residues.
vs. MW and relative concentration. There were three components in the system. The detailed information of each of the component is summarized in Table 3.8. The major component was the species with MW of 9,473 and the expected molecular weight for the ABC heterotrimer is 8,744.

Table 3.8. Summary of sedimentation velocity data with Monte Carlo analysis

<table>
<thead>
<tr>
<th>Component</th>
<th>Molecular Weight (95% confidence)</th>
<th>Relative concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component1</td>
<td>9473 (9366, 9572)</td>
<td>78.0%</td>
</tr>
<tr>
<td>Component2</td>
<td>1928 (1844, 2014)</td>
<td>15.8%</td>
</tr>
<tr>
<td>Component3</td>
<td>13939 (12857, 15339)</td>
<td>6.2%</td>
</tr>
</tbody>
</table>

Equilibrium data was fitted to different modes and the appropriate mode was chosen based on the visual inspection of the residue with random distribution and on the best statistics. It was found that 1-component, ideal model was the best model tested for this system. Shown in Figure 3.14 is the residue and overlay of the fitting. The molecular
weight found was 8077 with 95% confidence limits of 8014 (low), 8140 (high). The expected MW was 8744.

Although velocity and equilibrium experiments did not give the same MW for the system, they were in the +\(-10\)% range. To further understand this equilibration process, wavelength scans need to be performed to get a globally fitted extinction value for this peptide. Only after the wavelength scan experiments, should we be able to convert optical density data to molar concentration at the measured wavelength. And then we would be able to get the data associated to the equilibration.

3.3.6.2 (DOG)$_{10}$:(PKG)$_{10}$GY

From CD experiment, we observed that this system could form a triple helix, but the composition of the triple helix was unclear. As one of the chain was positively charged, the other was negatively charged, it was easy to think that the triple helix could be the triplet of this AB unit. Another hypothesis was that the ions in the solution could stabilize the extra charge if it formed an AAB or ABB type of triple helix. The third possibility was that it could be a hexamer which composed of AAB and ABB type of heterotrimer with a 1:1 ratio, and the net charge of this hexamer was neutral. All above mentioned structures could give us a triple helical signal by CD experiments. Shown in Table 3.9 is the expected molecular mass for some of the combinations of A and B. AUC experimental was an ideal tool to help us elucidate the composition of the triple helix. Sedimentation velocity experiment was performed for this system. Shown in Figure 3.15 is the 2-Dimensional spectrum analysis of friction ratio ($f/fo$) vs. MW and relative concentration. There were
Table 3.9. Expected molecular weight of different combinations.

<table>
<thead>
<tr>
<th></th>
<th>Trimer:</th>
<th>Tetramer:</th>
<th>Hexamer:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dimer: AB</td>
<td>AAB/ABB</td>
<td>2AB</td>
</tr>
<tr>
<td></td>
<td>6013.6</td>
<td>8924/9115</td>
<td>12027.2</td>
</tr>
</tbody>
</table>

three components in the system. The detailed information of each of the component is summarized in Table 3.10. The major component was the species with MW of 1.67×10^4 and this corresponded to the MW of hexamer with an expected MW of about 18,000. The other minor component with higher relative concentration could be the AAB or ABB trimer. More experiments need to be done to confirm the composition of the system, such as velocity experiment with a different concentration (230nm). Sedimentation equilibrium experiment will be also helpful to get the MW, the equilibrium constant and thermodynamic parameters to understand this equilibration process.

Figure 3.15. Sedimentation velocity data analysis of (DOG)_{10}:(PKG)_{10}GY system. (a) 2-D spectrum analysis. (b) Residues.
Table 3.10. Summary of sedimentation velocity data with Monte Carlo analysis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Molecular Weight (95% confidence)</th>
<th>Relative Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component 1</td>
<td>16,707 (16,445, 16,941)</td>
<td>75.2%</td>
</tr>
<tr>
<td>Component 2</td>
<td>8,872 (6,850, 11,130)</td>
<td>18.6%</td>
</tr>
<tr>
<td>Component 3</td>
<td>29,630 (27,899, 30,596)</td>
<td>6.2%</td>
</tr>
</tbody>
</table>

3.4 Conclusion

In this chapter we examined four ABC triple helix forming compositions. Of these four systems, all were shown to be able to form good ABC heterotrimeric helices. Two were reported for the first time. (DOG)$_{10}$(PKG)$_{10}$(POG)$_{10}$ was found to be the most thermally stable ABC type triple helices and (EOG)$_{10}$(POrG)$_{10}$(POG)$_{10}$ was found to be the least thermally stable ABC type triple helices. Increased salt concentration was found to destabilize all of the observed ABC helices while having no impact on (POG)$_{10}$ homotrimer, which proved the hypothesis that the electrostatic interactions are one of the major driving forces to form the stable ABC heterotrimers. DSC revealed that (DOG)$_{10}$(PKG)$_{10}$(POG)$_{10}$ had thermodynamic values comparable to that of (POG)$_{10}$ homotrimer, all other helix forming systems had lower $\Delta H$, $\Delta S$ and $\Delta G^\circ$ of unfolding. Sedimentation velocity and sedimentation equilibrium experiments were performed for (DOG)$_{10}$(PKG)$_{10}$GY and (DOG)$_{10}$(PKG)$_{10}$GY:(POG)$_{10}$ systems. The preliminary data confirmed the ABC heterotrimeric composition of (DOG)$_{10}$(PKG)$_{10}$GY:(POG)$_{10}$ system. While the composition of (DOG)$_{10}$(PKG)$_{10}$GY might be a hexamer. Further AUC
experiments need to be done to confirm the results and clarify the composition of the hexamer.
References


Chapter 4. Study of Heterotrimeric Collagen-like Peptide

Systems: Group 2 \((A_X(+), B_Y(-), C_0)\)

4.1 Introduction

Recently our group reported a method for preparing heterotrimeric collagen-like helices that utilized electrostatic interactions to organize their self-assemblies.\(^1\) We found that mixing three peptides that included a positively charged peptide, a negatively charged peptide, and a neutral peptide resulted in the formation of highly stable ABC heterotrimers. Among the ABC heterotrimers reported, \((\text{DOG})_{10}(\text{PKG})_{10}(\text{POG})_{10}\) was found to have the highest thermal stability with a melting temperature of 65°C.\(^2\) We believe that the inter-chain interaction of the collagen triple helices formed by the X residues in one chain and the Y residues in the adjacent chain. Apart from the hydrogen bonding and van der waals interactions, the driving force to stabilize the heterotrimeric collagen helices was the N-O salt bridges formed by the negatively charged \(\text{O}^-\) from glutamic acid (abbreviated as Glu or E) or aspartic acid (abbreviated as Asp or D) and positively charged \(\text{N}^+\) from the lysine (abbreviated as Lys or K). The unusual high stability of the \(A_{(\text{DOG})}B_{(\text{PKG})}C_{(\text{POG})}\) heterotrimer could be the spatial matching of the \(\text{O}^-\) on the side chain of the D and \(\text{N}^+\) on the side chain of the K. To further prove this hypothesis, we designed and synthesized two ornithine (abbreviated as Orn or Or) containing host-guest peptides: \((\text{POG})_3(\text{POrG})_4\) \((\text{POG})_4G\) and \((\text{POG})_3(\text{OrOG})(\text{POG})_4G\). In the structures of the two above-mentioned
host-guest peptides, either the X or Y positions were occupied by Orn which is an amino acid with the same amine functional group but one methylene less on its side chain than lysine. In Chapter 2, we studied the amino acid propensity of ornithine by host-guest peptide studies with these peptides and compared them with lysine host-guest peptides reported by Brodsky’s group. Then, in Chapter 3, we explored the effect of side chain length on the stabilization of the charged pairs. In this chapter, we examined the positional effect of the charged amino acids. In natural collagen there is a strong predisposition for negatively charged amino acids to be in the X positions and positively charged amino acids to be in the Y positions.\textsuperscript{4,5} Here, we intentionally reversed this predisposition. Four new peptides: \((\text{OrOG})_{10}\), \((\text{KOG})_{10}\), \((\text{PEG})_{10}\) and \((\text{PDG})_{10}\) were designed, synthesized and characterized. Comparing to peptides in Group 1, the amino acids in the X and the Y positions from these peptides were swapped. Now positively charged amino acids are in the X positions and negatively charged amino acids are in the Y positions. Studies found that Asp had an unfavorable effect in the Y positions,\textsuperscript{6} Glu prefers to be in the X position, and Lys prefers to be in the Y position.\textsuperscript{4,5} We would like to use these peptides and apply the similar approach as Group 1 to construct ABC heterotrimeric systems with positively charged collagen-like peptides \((\text{XOG})_{10}\) \((X = \text{Or, K})\), negatively charged collagen-like peptides \((\text{PYG})_{10}\) \((Y = \text{E, D})\), and \((\text{POG})_{10}\) in neutral pH. By studying these systems, we would like the following questions to be answered: Is the D-K pair the ideal ion pair to form the stable ABC heterotrimeric triple helix? Does the amino acid propensity and position of amino acids in the peptide chain \((X \text{ or } Y)\) play roles?
Group 2 were heterotrimers composed of \([X(+)OG]_{10}, [PY(-)G]_{10}\) and \((POG)_{10}\) as shown in Table 4.1. Each of the heterotrimeric system can possibly form ten triple helices, three homotrimers, six AAB type binary heterotrimers and one ABC heterotrimer. The homotrimeric study of each of the individual peptide showed that \((POG)_{10}, (OrOG)_{10}\) and \((KOG)_{10}\) formed triple helices under neutral pH. And their \(T_m\)s were distinguishable. As we mentioned in previous chapter, for the AAB type binary systems, due to the limitation of our analytical system, we simply called these systems AAB systems. They were evaluated by mixing the respective peptides together in 1:1 ratio and observing their thermal transition with and without preheating. In the case of preheating, two peptides were mixed, and the temperature was increased to 85°C, and then cooled slowly to allow the disrupted molecules to rearrange to their most thermodynamically stable structure. Their thermal unfolding behaviors were studied by CD and DSC. In all of the systems evaluated the AAB helix displayed a lower \(T_m\) as comparing to the ABC heterotrimers. After combining and evaluating all the possibilities, for most of the systems we could identify the ABC heterotrimer.

**Table 4.1. Group 2 ABC heterotrimeric systems**

<table>
<thead>
<tr>
<th>Charge in neutral pH</th>
<th>Abbreviation</th>
<th>Peptide composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 2 X(+)Y(-)</td>
<td>(A_{(KOG)}B_{(PDG)}C_{(POG)})</td>
<td>((KOG)<em>{10}:(PDG)</em>{10}:(POG)_{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td>(A_{(KOG)}B_{(PEG)}C_{(POG)})</td>
<td>((KOG)<em>{10}:(PEG)</em>{10}:(POG)_{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td>(A_{(OrOG)}B_{(PDG)}C_{(POG)})</td>
<td>((OrOG)<em>{10}:(PDG)</em>{10}:(POG)_{10}=1:1:1)</td>
</tr>
<tr>
<td></td>
<td>(A_{(OrOG)}B_{(PEG)}C_{(POG)})</td>
<td>((OrOG)<em>{10}:(PEG)</em>{10}:(POG)_{10}=1:1:1)</td>
</tr>
</tbody>
</table>
4.2 Characterization Methods

Using similar methods to Group 1 systems, we used CD and DSC to characterize the four systems in Group 2. Circular dichroism (CD) was used to evaluate the triple helicity of the collagen peptide systems. The studies of thermal unfolding by CD helped us to identify the ABC heterotrimer species and learn about their thermal stabilities. Thermal unfolding by differential scanning calorimetry (DSC) was used to understand the thermodynamic parameters which relate to the stability of the triple helices.

In the cases of two peaks for the heterotrimeric systems, the thermodynamic parameters were obtained as described in chapter 1.

4.3 Result and Discussion

4.3.1 (KOG)$_{10}$*(PDG)$_{10}$*(POG)$_{10}$

From our studies on homotrimers we knew that (PDG)$_{10}$ did not form a homotrimeric triple helix. However, unlike peptides from group 1, (KOG)$_{10}$ could form a homotrimer after incubation at 5°C for over 24 hours. Shown in Figure 4.1 is the thermal transition studies of mixture of (KOG)$_{10}$ and (PDG)$_{10}$. Without preheating the mixture showed a transition with a $T_m$ of 14°C. After preheating, the same mixture showed a transition with a $T_m$ of 21°C. We believe both of these transitions were AAB heterotrimers. While it was not entirely clear what was responsible for the difference in the transition temperature before and after preheating we often saw that the melting temperatures after
preheating were slightly higher. The system before this annealing step

![Figure 4.1](image1.png)

**Figure 4.1.** $A_{(KOG)B_{(PDG)}}$ in phosphate buffer. (a) Thermal transition at 225nm. (b) 1st derivative of the thermal transition.

may contained some imperfections in which the packing of the triple helix was offset by one or more amino acid triplets and that these staggered, or sticky ended, triple helices were able to find higher stability, blunt ended structures after heating. The analysis of this system was further complicated by pre-heated mixtures of $(KOG)_{10}$ with $(POG)_{10}$ shown in Figure 4.2. The thermal study without heating showed only the transition of $(POG)_{10}$

![Figure 4.2](image2.png)

**Figure 4.2.** $A_{(KOG)B_{(PDG)}}$ in PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.
homotrimer with $T_m$ of 67°C. The thermal study with heating showed two transitions: besides the $(POG)_{10}$ homotrimers transition, there appeared a new weak transition with $T_m$ of 50°C. It could be an additional AAB type of heterotrimer. The thermal transition studies of $(PDG)_{10}$ and $(POG)_{10}$ mixture with heating and without heating showed only the transition of $(POG)_{10}$ homotrimers with $T_m$ of 67°C (Figure 4.3). Shown in Figure 4.4

**Figure 4.3.** $A_{(PDG)B(POG)}$ in PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

is the thermal unfolding study of the $(KOG)_{10}$, $(PDG)_{10}$ and $(POG)_{10}$ peptide mixture in phosphate buffer. Without preheating, the thermal study showed a very strong transition of $(POG)_{10}$ homotrimer with $T_m$ at 67°C and another peak with $T_m$ of 17.5°C which could be assigned to the $(KOG)_{10}$ / $(PDG)_{10}$ AAB helix. Markedly different from what was seen in group 1, after heating and allowing time for equilibration, the transition of $(POG)_{10}$ homotrimer was still the major peak. A new peak emerged with a $T_m$ of 53°C that could be tentatively assigned to an ABC heterotrimer. However, the driving force for the formation
of the ABC helix was clearly not sufficient to effectively out compete the (POG)_{10} homotrimer. The thermal transition data are summarized in Table 4.2.

![Figure 4.4](image.png)

**Figure 4.4.** Thermal unfolding of collagen peptide (KOG)_{10}(PDG)_{10}(POG)_{10} in 10mM sodium phosphate buffer as monitored by CD. (a) ellipticity at 225nm. (b) 1st derivative of (a).

<table>
<thead>
<tr>
<th>Type of Heterotrimer</th>
<th>Peptides</th>
<th>Phosphate buffer</th>
<th>Thermal History</th>
<th>PBS buffer</th>
<th>Thermal History</th>
<th>T_m (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>A_{KOG}B_{PDG}C_{POG}</td>
<td>Preheat 20,35,53,68</td>
<td>w/o heat 17.5, 68</td>
<td>Preheat</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>A_{KOG}B_{PDG}</td>
<td>Preheat 21</td>
<td>w/o heat 14</td>
<td>Preheat</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>A_{KOG}B_{POG}</td>
<td>Preheat ND</td>
<td>w/o heat ND</td>
<td>Preheat</td>
<td>67.5</td>
<td></td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>A_{PDG}B_{POG}</td>
<td>Preheat ND</td>
<td>w/o heat ND</td>
<td>Preheat</td>
<td>67.5</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.2.** Heterotrimers in (KOG)_{10}, (PDG)_{10} and (POG)_{10} peptide mixture.

### 4.3.2 (KOG)_{10}(PEG)_{10}(POG)_{10}

From the homotrimeric studies, (KOG)_{10} formed a triple helix after incubation at 5°C for over 24 hours. The T_m of its homotrimer was 34°C in phosphate buffer and 31°C in PBS buffer. The thermal transition studies of (KOG)_{10} and (PEG)_{10} mixture with heating...
and without heating in phosphate buffer showed that the two peptides formed AAB type of triple helix with $T_m$ of about 33°C in phosphate buffer (Figure 4.5). As mentioned earlier, the thermal transition studies of $(KOG)_{10}$ and $(POG)_{10}$ mixture was a little complicated. The preheated mixtures showed an additional AAB type of heterotrimer with a $T_m$ of 50°C. Figure 4.6 shows the thermal study of $(PEG)_{10}$ and $(POG)_{10}$ mixture with only the $(POG)_{10}$

![Figure 4.5](image)

**Figure 4.5.** $A_{(KOG)B(PEG)}$ in phosphate buffer. (a) Thermal transition. (b) 1st derivative of the thermal transition.

transition. Shown in Figure 4.7 is the thermal transition of the $(KOG)_{10}$, $(PEG)_{10}$ and $(POG)_{10}$ peptide mixture in phosphate buffer. Without heating, the thermal study showed two well-separated transitions with $T_m$ of 31.5°C and 68°C. They were from the transitions of the $(KOG)_{10}$ and $(POG)_{10}$ homotrimers. After heating and allowing time for incubation, one new transition was emerged with $T_m$ of 43.5°C. We think that the new transition with $T_m$ of 43.5°C corresponded to the transition of the $A_{(KOG)B(PEG)C(POG)}$ heterotrimer. The thermal transition data are summarized in Table 4.3.
Figure 4.6. $A_{(PEG)}B_{(POG)}$ in PBS (a) Thermal transition. (b) 1st derivative of the thermal transition.

Figure 4.7. Thermal unfolding of collagen system (KOG)$_{10}$•(PEG)$_{10}$•(POG)$_{10}$ in 10 mM sodium phosphate buffer monitored by CD. (a) ellipticity at 225nm. (b) 1st derivative of (a).

Table 4.3. Heterotrimers from (KOG)$_{10}$, (PEG)$_{10}$ and (POG)$_{10}$.

<table>
<thead>
<tr>
<th>Type of Heterotrimer</th>
<th>Peptides</th>
<th>Phosphate buffer Thermal History</th>
<th>Thermal History</th>
<th>Tm (°C)</th>
<th>PBS buffer Thermal History</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>(KOG)$<em>{10}$•(PEG)$</em>{10}$•(POG)$_{10}$</td>
<td>Preheat 43.5</td>
<td>Preheat</td>
<td>28, 40,</td>
<td>w/o heat 31.5, 68</td>
<td>w/o heat 28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>w/o heat 31.5, 68</td>
<td></td>
<td></td>
<td>w/o heat 67</td>
<td></td>
</tr>
<tr>
<td>AAB</td>
<td>(KOG)$<em>{10}$, (PEG)$</em>{10}$</td>
<td>Preheat 33.5</td>
<td>Preheat</td>
<td>28</td>
<td>w/o heat 33.5</td>
<td>27.5</td>
</tr>
<tr>
<td>AAB</td>
<td>(PEG)$<em>{10}$, (POG)$</em>{10}$</td>
<td>Preheat ND</td>
<td>Preheat</td>
<td>66.5</td>
<td>w/o heat ND</td>
<td>67.5</td>
</tr>
</tbody>
</table>
4.3.3 (OrOG)_{10} \cdot (PDG)_{10} \cdot (POG)_{10}

From the homotrimeric study, (OrOG)_{10} and (POG)_{10} formed homotrimeric triple helices and (PDG)_{10} did not form a triple helix. The thermal transition studies of (OrOG)_{10} and (PDG)_{10} mixture with heating and without heating in phosphate buffer showed a transition with a $T_m$ of 18°C (Figure 4.8). This transition might correspond to one of the AAB type of triple helix. Unlike in the (KOG)_{10} and (PDG)_{10} mixture, the AAB

![Graph](image)

**Figure 4.8.** A_{(OrOG)}B_{(PDG)} in 10 mM phosphate buffer. (a) Thermal transition at 225nm. (b) 1\textsuperscript{st} derivative of the thermal transition.

heterotrimer formed by (OrOG)_{10} and (PDG)_{10} was more stable. Shown in Figure 4.9 are the thermal transitions of (OrOG)_{10} and (POG)_{10} mixture. It was a little complicated. The thermal study without heating showed two transitions; one had a $T_m$ of 28°C, and the other had a $T_m$ of 67°C. They were from (OrOG)_{10} and (POG)_{10} homotrimers. The thermal study with heating also showed two transitions, with the (OrOG)_{10} transition disappearing, a new transition with a $T_m$ of 48.5°C emerged, and the (POG)_{10} transition was still there. The new transition could be the AAB type heterotrimer. Figure 4.10 shows the thermal
Figure 4.9. $A_{(OrOG)}B_{(POG)}$ in PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

Figure 4.10. Thermal unfolding of collagen peptide $(OrOG)_{10}(PDG)_{10}(POG)_{10}$ in 10mM sodium phosphate buffer monitored by CD. (a) ellipticity at 225nm. (b) 1st derivative of (a).
	ransition of the $(OrOG)_{10}$, $(PDG)_{10}$ and $(POG)_{10}$ peptide mixture in phosphate buffer. Without heating, the thermal study showed a very strong transition of $(POG)_{10}$ homotrimer with a $T_m$ at 68°C. After heating and allowing time for incubation, the $(POG)_{10}$ homotrimer transition was still the major transition and two new transitions emerged with $T_m$ of 33°C and 50°C. As the situation was too complicated, it was very hard to identify AAB and ABC heterotrimers. Apparently, the ABC heterotrimer here was not very stable
or not easy to form mainly due to the very unstable (PDG)10 that was reported by Brodsky's host-guest collagen peptide study. This indicated that the stability of each individual chain also played role in stabling the ABC heterotrimers. The thermal transition data are summarized in Table 4.4.

Table 4.4. Possible heterotrimers in (OrOG)io, (PDG)io and (POG)io peptide mixture.

<table>
<thead>
<tr>
<th>Type of Heterotrimer</th>
<th>Peptides</th>
<th>Phosphate buffer</th>
<th>PBS buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Thermal History</td>
<td>Tm (°C)</td>
</tr>
<tr>
<td>ABC</td>
<td>(OrOG)io*(PDG)io*(POG)io</td>
<td>Preheat</td>
<td>33, 52, 68, 18, 68</td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>(OrOG)io*(PDG)io</td>
<td>Preheat</td>
<td>18</td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>(OrOG)io*(POG)io</td>
<td>Preheat</td>
<td>ND</td>
</tr>
</tbody>
</table>

4.3.4 (OrOG)io*(PEG)io*(POG)io

From the homotrimeric study, we knew that (PEG)io did not form a triple helix as its thermal transition at 225nm was linear. (OrOG)io and (POG)io formed homotrimeric triple helices, their Tm's were 28°C and 68°C in PBS buffer respectively. The thermal transition studies of (OrOG)io and (PEG)io mixture with heating and without heating showed that the two peptides formed AAB type of triple helix with Tm of 33°C in phosphate buffer (Figure 4.11). As mentioned earlier, the thermal transition studies of (OrOG)io and (POG)io mixture was a little complicated. The thermal study of (PEG)io and (POG)io mixture only showed the (POG)io transition. Shown in Figure 4.12 is the
Figure 4.11. $A_{(OrOG)}B_{(PEG)}$ in phosphate buffer. (a) Thermal transition. (b) 1st derivative of the thermal transition.

Figure 4.12. Thermal unfolding of collagen peptide $(OrOG)_{10}$*(PEG)$_{10}$*(POG)$_{10}$ in 10mM sodium phosphate buffer monitored by CD. (a) ellipticity at 225nm. (b) 1st derivative of (a). thermal transition of the $(OrOG)_{10}$, (PEG)$_{10}$ and (POG)$_{10}$ peptide mixture. Without heating, the thermal study showed two well-separated transitions, they were from the transition of the $(OrOG)_{10}$ and (POG)$_{10}$ homotrimers. After heating and allowing time for incubation, one new transition emerged with a $T_m$ of 50°C. We think that the new transition with $T_m$ of
50°C corresponded to the transition of the $A(OrOG)B_{(PEG)}C_{(POG)}$ heterotrimer. The thermal transition data are summarized in Table 4.5.

**Table 4.5.** Possible heterotrimers in $(OrOG)_{10}$, $(PEG)_{10}$ and $(POG)_{10}$ peptide mixture.

<table>
<thead>
<tr>
<th>Type of Heterotrimer</th>
<th>Peptides</th>
<th>Thermal History</th>
<th>$T_m$ (°C)</th>
<th>PBS buffer</th>
<th>Therma l History</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>$A(OrOG)B_{(PEG)}C_{(POG)}$</td>
<td>Preheat w/o heat</td>
<td>50</td>
<td>Preheat</td>
<td>42, 68.5</td>
<td></td>
</tr>
<tr>
<td>AAB (ABB)</td>
<td>$A(OrOG)B_{(PEG)}$</td>
<td>Preheat w/o heat</td>
<td>35</td>
<td>Preheat</td>
<td>28.5</td>
<td></td>
</tr>
</tbody>
</table>

**4.3.5 Effect of Salts on Helix Stability**

In our heterotrimeric systems, one of the major driving forces to form an ABC heterotrimeric collagen system was the electrostatic interactions of the ion pairs formed by the negatively charged amino acids and the positively charged amino acids. To further prove the existence of this ion pair interaction and its effects on the stabilities of the series of ABC heterotrimeric collagen-like peptides, we increased the concentration of salts by adding sodium chloride to the phosphate buffer and then compared the stabilities of the heterotrimers in 10 mM phosphate buffer and PBS (phosphate buffer saline) which was composed of 10 mM sodium phosphate buffer in addition to 150 mM NaCl. Shown in Figures 4.13-4.16 are the CD comparisons of the Group 2 ABC helices in phosphate buffer and PBS. While the $T_m$ of $(POG)_{10}$ homotrimer was not altered, the $T_m$ decreased for all ABC systems in the presence of NaCl. The largest decrease was $(OrOG)_{10}*(PEG)_{10}*(POG)_{10}$ whose $T_m$ was lowered by 7.5°C. The addition of NaCl was
also found to have a detrimental impact on compositional purity as ABC triple helices in

![Graphs showing thermal transition comparison in phosphate buffer and PBS.](image)

**Figure 4.13.** $A_{(KOG)}B_{(PDG)}C_{(POG)}$ thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

![Graphs showing thermal transition comparison in phosphate buffer and PBS.](image)

**Figure 4.14.** $A_{(KOG)}B_{(PEG)}C_{(POG)}$ thermal transition comparison in phosphate buffer and PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

all four systems. These systems showed only two transitions in simple phosphate buffer, however, they were all found to have a significant quantity of (POG)$_{10}$ homotrimer and other new peaks emerged in the low temperature region in the presence of salts. This could be simply explained by the fact that the stability of (POG)$_{10}$ homotrimer was not effected
by the salts while the ABC heterotrimers were destabilized. The exception to this was
(OrOG)_{10}*(PEG)_{10}*(POG)_{10} that was found to have only stronger (POG)_{10} peak in the
presence of NaCl. Table 4.6 is the summary of the Tm's in phosphate and PBS buffer.

Figure 4.15. A(OrOG)B(PDG)C(POG) thermal transition comparison in phosphate buffer and
PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.

Figure 4.16. A(OrOG)B(PEG)C(POG) thermal transition comparison in phosphate buffer and
PBS. (a) Thermal transition. (b) 1st derivative of the thermal transition.
Table 4.6. Summary of the $T_m$ of ABC heterotrimers in phosphate buffer and PBS.

<table>
<thead>
<tr>
<th>Triple Helix Composition</th>
<th>$T_m$ (°C) in Phosphate Buffer</th>
<th>$T_m$ (°C) in PBS</th>
<th>Thermal destabilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 \cdot (POG)_{10}$</td>
<td>67.5</td>
<td>67.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(KOG)<em>{10} \cdot (PDG)</em>{10}$</td>
<td>53, 69</td>
<td>41, 53, 67</td>
<td>-</td>
</tr>
<tr>
<td>$(KOG)<em>{10} \cdot (PEG)</em>{10} \cdot (POG)_{10}$</td>
<td>44</td>
<td>28, 40, 67</td>
<td>4.0</td>
</tr>
<tr>
<td>$(OrOG)<em>{10} \cdot (PDG)</em>{10} \cdot (POG)_{10}$</td>
<td>52.5, 69</td>
<td>36, 40, 68.5</td>
<td>-</td>
</tr>
<tr>
<td>$(OrOG)<em>{10} \cdot (PEG)</em>{10} \cdot (POG)_{10}$</td>
<td>49.5</td>
<td>42, 68.5</td>
<td>7.5</td>
</tr>
</tbody>
</table>

4.3.6 Thermodynamic Analysis

DSC measurements of the four ABC collagen-like peptide systems were performed under the same conditions as used for CD analysis. Only pre-heated samples were examined. The thermodynamic parameters were determined by direct integration of the DSC curve. Figure 4.17 shows the DSC thermal unfolding curves for group 2 helices. Our thermodynamic results were based on the simplification that the difference of the heat capacity of the folded and unfolded peptides ($\Delta C_p$) was zero. Hence, the $\Delta H$ and $\Delta S$ would be constant with respect to the temperature. Shown in Table 4.7 is the summary of the $\Delta H$, $\Delta S$ and $\Delta G^\circ$ of unfolding for the four systems. For all systems, the $T_m$s from DSC matched very well with the $T_m$s from CD. Due to the complexity of the systems containing $(PDG)_{10}$, we were unable to calculate the thermodynamic parameters for these two systems. Our expectation was that the replacement of the conformationally restricted, cyclic imino acids proline and hydroxyproline with more flexible, charged, amino acids could form electrostatic hydrogen bonds and would have simultaneously increased the $\Delta H$ and $\Delta S$, but this was not observed to be the case. As we mentioned in the analysis of the different heterotrimers, we knew that $A_{(KOG)}B_{(PDG)}C_{(POG)}$ could not form a clean triple helix as
Figure 4.17. DSC spectra of Group 2 ABC helices.

the \( \text{A(DOG)} \text{B(PKG)} \text{C(POG)} \) system at the same experimental conditions. Even though
\( \text{A(KOG)} \text{B(PEG)} \text{C(POG)} \) triple helices formed under the same condition as the
\( \text{A(EOG)} \text{B(PKG)} \text{C(POG)} \), the \( T_m \), \( \Delta H \) and \( \Delta S \) of \( \text{A(KOG)} \text{B(PEG)} \text{C(POG)} \) were much smaller than
that of \( \text{A(EOG)} \text{B(PKG)} \text{C(POG)} \). It meant that the position of the amino acid and the stabilities
of the corresponding peptide monomers also played roles in the stability of the
heterotrimeric triple helices.

Table 4.7. Thermodynamic parameters obtained from DSC (in phosphate buffer)

<table>
<thead>
<tr>
<th>Triple Helix Composition</th>
<th>( T_m(°C) ) from DSC</th>
<th>( T_m(°C) ) from CD</th>
<th>( \Delta H_{\text{cal}} ) (kcal/mol)</th>
<th>( \Delta S ) (cal/mol/K)</th>
<th>( \Delta G^\circ ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\text{POG})<em>{10} \cdot (\text{POG})</em>{10} \cdot (\text{POG})_{10})</td>
<td>67.7</td>
<td>67.5</td>
<td>101</td>
<td>297</td>
<td>12.7</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((\text{KOG})<em>{10} \cdot (\text{PDG})</em>{10} \cdot (\text{POG})_{10})</td>
<td>52.4</td>
<td>53</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>((\text{KOG})<em>{10} \cdot (\text{PEG})</em>{10} \cdot (\text{POG})_{10})</td>
<td>45.1</td>
<td>44</td>
<td>56.9</td>
<td>179</td>
<td>3.6</td>
</tr>
<tr>
<td>((\text{OrOG})<em>{10} \cdot (\text{PDG})</em>{10} \cdot (\text{POG})_{10})</td>
<td>51.9</td>
<td>52</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>((\text{OrOG})<em>{10} \cdot (\text{PEG})</em>{10} \cdot (\text{POG})_{10})</td>
<td>49</td>
<td>49.5</td>
<td>52.8</td>
<td>164</td>
<td>3.9</td>
</tr>
</tbody>
</table>
4.4 Conclusion

In this chapter we examined four ABC collagen triple helical forming systems. For these four systems, we maintained the same amino acid composition as the peptides in group 1, but the sequence positions of the charged amino acids were swapped. Here basic amino acids occupied the X positions and acidic amino acids occupied the Y positions. Triple helices containing lysine were destabilized by 12°C and 14°C. In contrast, triple helices containing ornithine were stabilized by 4°C and 3°C. The change in stability did not, however, tell the entire story. While the compositional purity of the glutamic acid containing helices was quite good after preheating, group 2 ABC triple helices containing aspartic acid were only formed as a minority component of what was seen to be a complicated mixture. Interestingly the helix compositional purity was not directly related to the thermal stability of the ABC triple helix. For example, while \((\text{KOG})_{10}\text{PEG}_{10}\text{POG}_{10}\) cleanly formed with a \(T_m\) of 43.5°C, \((\text{KOG})_{10}\text{PDG}_{10}\text{POG}_{10}\) formed only as a minority component despite a higher melting point of 53°C. Similarly, \((\text{OrOG})_{10}\text{PEG}_{10}\text{POG}_{10}\) was the majority component with a \(T_m\) of 50°C while \((\text{OrOG})_{10}\text{PDG}_{10}\text{POG}_{10}\) had a higher \(T_m\) of 52°C yet is one of four species present in a complicated mixture. Thermodynamic analysis results showed that the position of the amino acids and the stabilities of the corresponding peptide monomers also played roles in the stability of the heterotrimeric triple helices.
References


Chapter 5. Aromatic Amino Acids Providing Characteristic Motifs in the Raman and SERS Spectroscopy of Peptides

5.1 Introduction

Surface Enhanced Raman Spectroscopy (SERS) has been an active research area since its discovery in the 1970s, with important applications ranging from surface chemistry to biological chemistry and biomedical analysis. Much of the interest in SERS focuses on making use of the large SERS enhancements provided by the substrate to detect organic and biological molecules down to single molecule detection limits. SERS of many amino acids, peptides, and proteins acquired with various SERS active substrates have been reported. In general, however, the SERS spectra of biological molecules such as proteins and DNA have been difficult to reproduce and interpret, and many SERS spectra reported differ significantly from their respective Raman spectra. In many cases it is not uncommon to find the appearance of new modes or the disappearance of modes characteristic of the normal Raman spectrum in the SERS spectrum. Additionally, frequency shifts between modes in normal Raman and SERS spectra as large as ~20 cm\(^{-1}\) have been reported. Most of the observed differences between Raman and SERS spectra have been attributed to differences in

selection rules and/or chemical effects resulting from interactions between the adsorbate molecules and the SERS substrate. However, limitations in our understanding of these differences have severely limited the adoption of SERS as a biodetection and characterization method.

Nanoshells belong to a class of metallodielectric nanoparticles with a surface plasmon resonance which is tunable from visible to infrared wavelengths. With their excellent chemical stability and highly reproducible electromagnetic enhancements, nanoshells have been shown to be useful for SERS detection of DNA, as well as small molecules such as para-mercaptobenzoic acid and alkanethiols. In this work, a series of three simple peptides designed specifically for this Raman-SERS spectroscopy study, consisting of a single cysteine bound to a single aromatic amino acid, were synthesized and characterized: tyrosine-cysteine (YC), phenylalanine-cysteine (FC) and tryptophan-cysteine (WC). Normal Raman spectra were obtained for these peptides in solution. The peptides were also bound to the surfaces of Au nanoshells via the thiol moiety of the cysteine, and SERS spectra were obtained. The Raman and SERS spectra were quantitatively compared to each other to determine differences in relative Raman cross section of the aromatic amino acids, peak position and linewidth of the observed Stokes modes.

Several considerations were taken into account in the design of the peptides used in this study. For the three peptides synthesized, the number and local molecular environment of the aliphatic C-H bonds are quite similar, so these modes should all have
similar Raman cross sections. The Stokes mode amplitudes of the aromatic constituents of the peptides can therefore be compared to the amplitudes of the aliphatic C-H stretch region to obtain their relative Raman cross sections.\textsuperscript{30} Cysteine was selected to provide a covalent interaction between the Au nanoshell substrate and the peptide molecules through thiol-Au bond formation. Such a covalent linkage presumably orients all the three dipeptides similarly on the nanoshell surface, so the same surface selection rule can be applied to modulate the SERS activity of their aliphatic C-H bonds. The use of relatively high concentration of dipeptide solution (0.3 mM) was to ensure a high surface coverage of the peptide adsorbate molecules on the nanoshell surface for SERS and should also displace nonspecific adsorbates such as chemical impurities that may be found in the sample solutions. Both high substrate coverage and the displacement of unwanted adsorbed impurities are important for obtaining high spectral reproducibility in SERS. The three aromatic amino acids were selected since the Raman features of proteins are typically dominated by the vibrational modes involving the amide backbone and the aromatic amino acid side chains.\textsuperscript{31} Identifying the spectral characteristics of aromatic amino acid residues in both normal Raman and SERS, including their relative Raman cross-sections in these two modalities, provide an important knowledge base for our understanding and prediction of the SERS spectra of proteins.

To our knowledge, this work represents the first quantitative comparison of the Raman and SERS characteristics of the Stokes modes for each individual aromatic amino acid, as well as a cross comparison of the spectral features of these three aromatic amino
acids. The spectra of the YC and FC peptide were then used to predict the normal Raman and SERS spectra of penetratin, a 19 amino acid cell-penetrating peptide. Excellent agreement between the predicted and experimental Raman and SERS spectra for penetratin indicates that the aromatic peptide spectra can be used as an empirical “basis set” to predict and assign the Raman spectral features of large biological peptides and proteins, correlating with the aromatic amino acid content of the biomolecule.

5.2 Peptide Design and Synthesis

Since the aromatic amino acids have their characteristic Raman modes, it would be easier to identify and characterize them. We chose to use aromatic amino acids such as phenylalanine, tyrosine, and tryptophan as one of the building blocks. In the mean time, we would like the peptide to form a covalent bond on the gold nanoshell surface by the Au-S bond. We use amino acid cysteine, which has a thiol group on its side chain, as the other building block. To avoid undesired interaction of the termini of the peptides with the gold surface, all the C and N termini are acetylated and amided.

5.2.1. Dipeptides and Penetratin Peptide

To facilitate our study of using gold nanoshells as substrate of SERS of biomolecules, we started with the smallest and simplest peptides---dipeptides. As shown in Table 5.1, FC, YC and WC are the three dipeptides. We also designed a longer peptide, penetratin, which is a modified sequence of the actual penetratin peptide. We used two
glycines at the N-termini as spacers to have cysteine in the sequence in order to attach the peptide on the Au nanoshell surface by Au-S bond.

### Table 5.1. Peptide Sequences

<table>
<thead>
<tr>
<th>Peptide Name</th>
<th>Peptide Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptophan-Cysteine</td>
<td>WC</td>
</tr>
<tr>
<td>Tyrosine-Cysteine</td>
<td>YC</td>
</tr>
<tr>
<td>Phenylalanine-Cysteine</td>
<td>FC</td>
</tr>
<tr>
<td>Penetratin</td>
<td>CGGRQIKIWFQNRRMKWKK</td>
</tr>
</tbody>
</table>

<sup>a</sup> All peptides were N-terminally acetylated and C-terminally amidated.

#### 5.2.2. Peptide Synthesis

The peptide sequences are given in Table 5.1. Both the C terminal and N terminal ends of all the peptides were capped to reduce interfering electrostatic interactions. Peptides were synthesized using solid phase peptide synthesis with standard FMOC chemistry. Except the three dipeptides, all the other peptides were synthesized on an Advanced Chemtech Apex 396 peptide synthesizer on a 0.15 mmol scale. The three peptides were manually synthesized on a 0.3 mmol scale. The coupling reagents for the synthesis were HBTU, HOBt, and DIEA using 4, 4, and 6 molar equivalents with respect to the amino acid loading. The acetylation was performed with a large excess of acetic anhydride in the presence of DIEA in dichloromethane. The cleavage and deprotection were performed using a mixture of trifluoroacetic acid (TFA): triisopropylsilane : EDT : water = 30 : 1 : 1 : 1 by volume. Peptides were purified by reverse phase high performance liquid chromatography (HPLC) with a C-18 column and characterized by MALDI-TOF MS.
5.3 Synthesis and Characterization of Au Nanoshells

5.3.1 Synthesis of Au Nanoshells

Au nanoshells were made following a published procedure. Briefly, tetraethyl orthosilicate (TEOS), NH₄OH and ethanol were mixed together and stirred overnight to form SiO₂. SiO₂ was then functionalized with 3-aminopropyltrimethoxysilane (APTMS). The functionalized SiO₂ was mixed with Au colloid to form a “seed” with approximately a 30% surface coverage of gold nanoparticles on the SiO₂ surface. The seed solution was mixed with HAuCl₄ and formalaldehyde was added as a reducing agent to form a complete gold layer on the SiO₂.

5.3.2 Characterization of Au Nanoshells by UV-visible and TEM

UV-VIS was used to evaluate the plasmon resonance and, along with TEM, the completeness of the Au shell formation on the nanoparticles. Shown in Figure 5.1 is the UV-VIS spectra of a batch of complete Au nanoshell, the peak at 785nm shown is the plasmon resonance of the Au nanoshell, the narrowness of the peak indicates it has a complete Au layer on the surface. Shown in Figure 5.2 are the TEM images of the Au nanoshells at different fabrication stages. The size of the SiO₂ core is about 110-120 nm in diameter shown in the top left. The bottom left is a complete Au nanoshell.
5.4 Raman and SERS Measurements of Peptides and Peptide-nanoshell Conjugates

5.4.1 Conjugation of Cysteine Containing Peptides on Au Nanoshells

Fused Quartz (FQ) was cleaned by overnight incubation in freshly prepared piranha solution (H₂O₂: H₂SO₄ = 1: 3) followed by a thorough rinse with Milli-Q deionized water and subsequently dried with a gentle flow of N₂. The cleaned fused quartz was then immersed in a 1% poly-vinyl pyridine (PVP) ethanol solution for 24 hours, subsequently rinsed with ethanol and dried with N₂ gas before being used to
Figure 5.2. TEM Images of Au nanoshells at different fabrication stages. **Top left:** SiO$_2$ core (scale bar: 50 nm). **Top right:** Seed, which has about 25-30% coverage of Au nanoparticles on the SiO$_2$ core surface (scale bar: 20 nm). **Bottom left:** Complete gold nanoshell (scale bar: 20 nm). **Bottom right:** Incomplete Au nanoshell, in which the Au nanoparticles on the SiO$_2$ did not coalesce to form a complete Au layer (scale bar: 20 nm).

Immobilize analyte-bound nanoshells. Conjugation of nanoshell and peptide samples was performed by mixing 0.5 ml Au nanoshell solution (~10 pM) with respective peptide solutions (final peptide concentration after mixing was 0.3 mM). These were allowed to
mix at room temperature for 24 hours. The Au nanoshell solution was centrifuged and the supernatant removed. The pellet was then dispersed in deionized water. The centrifugation and resuspension was repeated two additional times to ensure removal of any remaining non-specifically bound peptide. After final resuspension, 5 µl of the peptide conjugated Au nanoshell suspension was deposited on the PVP functionalized fused quartz substrate and dried under ambient.

### 5.4.2 Raman Spectra Measurement of Peptides

Except for the FC peptide, where ~1% TFA was added to increase its solubility, all samples were prepared using purified water (18 MΩ cm⁻¹, Millipore). Normal Raman spectra were taken on SpectRIM substrates (SigmaAldrich) made of Teflon-coated stainless steel. A near-saturated solution of 50 to 100 µl of the peptide was deposited onto the substrates. Raman measurements were performed in a closed and vapor-saturated chamber to minimize solvent evaporation during spectral acquisition. A 63X water immersion lens was used for spectral acquisition. This corresponds to a sampling spot of 3 µm x 30 µm in the InVia Renishaw MicroRaman system with a 1200 lines/mm grating. The integration time of the normal Raman spectra ranged from 60-100 s with a laser excitation wavelength of 785 nm with a laser power of 200 mW measured before the microscope objective.
5.4.3 SERS Spectra Measurement of Peptide-nanoshell Conjugates

All the SERS spectra were obtained using the same instrument as the normal Raman spectra, but with a 50X objective. The laser power was set at 200 μW, 1000 times less power than that used for the normal Raman spectra, to ensure that no photo-induced modifications in the sample occurred within our data acquisition time (50-100 s).

5.4.4 Raman and SERS Spectra Data Analysis

Unless stated otherwise, all spectra were solvent and instrument background subtracted and baseline corrected with a multiple point linear curve fitting baseline correction. Peak fitting was performed using Grams AI software (Grams/AI 8.0, Thermo Scientific). The entire peak fitting results reported here were the average of multiple measurements from the same peptide and gold nanoshell conjugation. For each vibrational mode, the same fitting region was used for the multiple measurements, but no constraints were imposed on the peak position and peak width during fitting. The peak fitting region for the C-H antisymmetric stretching mode centered near 2940 cm\(^{-1}\) was the same for all peptides shown in Table 5.2. The ratio of integrated peak area of each vibrational mode and the aliphatic C-H anti-symmetric stretching mode was obtained, and is referred to as the relative Raman cross section. It is worthy of noting that such a relative Raman cross-section is independent of experimental parameters used for spectral acquisition such as spectral integration time and power, number of molecules at the probe
volume, quality of laser focusing, etc. thus enabling cross-comparison of Raman and SERS activities of different amino acid residues.

Table 5.2. Raman and SERS spectra peaking fitting parameters

<table>
<thead>
<tr>
<th>Approximate Peak Center</th>
<th>Peak Fitting Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>2945</td>
<td>2889.8  2998.6</td>
</tr>
<tr>
<td>1003, 1033</td>
<td>975.7   1061.4</td>
</tr>
<tr>
<td>644</td>
<td>607.3   673.6</td>
</tr>
<tr>
<td>833, 853 doublet</td>
<td>800.1   908.1</td>
</tr>
<tr>
<td>760</td>
<td>743.7   776.4</td>
</tr>
<tr>
<td>880</td>
<td>860.3   897.2</td>
</tr>
<tr>
<td>1012</td>
<td>992.0   1028.0</td>
</tr>
<tr>
<td>1551</td>
<td>1529.1  1570.4</td>
</tr>
</tbody>
</table>

5.4.5 SERS Spectra Reproducibility Analysis

One of the challenges of using SERS to study biomolecules is the reproducibility of the multiple measures. Here we use cross-correlation coefficient to evaluate the reproducibility of the SERS spectra for each conjugate. Prior to reproducibility analysis, all spectra were processed with a Savitzky-Golay second derivative method (window size of 39 data points with 2\textsuperscript{nd} order polynomial). Correlation coefficients between all non-identical spectral pairs (i \neq j) in the same dataset were determined from the data.\textsuperscript{33}
\[ P_{i,j} = \frac{\sum_{k=1}^{w} (I_i(k) - \bar{I}_i)(I_j(k) - \bar{I}_j)}{\sigma_i \sigma_j} \] (1)

where \( i, j \) is the index of the spectra in the data matrix, \( k \) is the wavenumber index of the individual spectra, \( I \) is the spectral intensity, \( W \) is the spectral range, and \( \sigma \) is the standard deviation of the \( i^{th} \) spectrum. Once the correlation coefficients \( P_{ij} \) are calculated, \( \Gamma \), the average of the off-diagonal correlation coefficients, can then be determined:

\[ \Gamma \equiv \frac{2 \sum_{i=1}^{N} \sum_{j=i+1}^{N} P_{i,j}}{N(N-1)} \] (2)

\( \Gamma \) thus defined is an easily determined and very useful parameter for quantitative assessment of spectral reproducibility. \( \Gamma \) varies between 0 and 1, where 1 is the case of identical spectra and 0 the case of completely uncorrelated spectra. \( \Gamma \) as defined in Equation 2 was used to evaluate the reproducibility of multiple measurements.

### 5.5 Results and Discussions

#### 5.5.1 Reproducibility of SERS Spectra

A series of vertically displaced SERS spectra of the tryptophan-cysteine (WC) peptide bound to nanoshell substrates, acquired at various positions on the same substrate and also on different substrates, is shown in Figure 5.3 (a). In each sampled region, the same modes appear with extremely high reproducibility with only a slight variation in
amplitude for some of the higher wavenumber modes. The normalized second derivative of the SERS spectra in Figure 5.3 (a) are shown in Figure 5.3 (b), displayed without offset. The almost complete overlap in features observed in the second derivative spectra indicate that the differences in the SERS spectra are primarily intensity and/or baseline

![Figure 5.3. SERS spectra of the WC peptide from 400 cm$^{-1}$ to 1800 cm$^{-1}$: (a) Offset SERS spectra from eight different measurements. (b) Second derivative of SERS spectra with calculated cross correlation coefficient $\Gamma$ of 0.95.](image)
variations and not due to variations in peak positions or in relative peak intensities. \( T \) was calculated as a cross-correlation between all pairs of spectra and an average \( T \) of 0.95 was obtained, a quantitative metric of the high degree of reproducibility observed. Similar results were found with SERS spectra acquired with the two other peptides, also demonstrating excellent reproducibility (Figure 5.4 and Figure 5.5). In addition to their

Figure 5.4. SERS spectra of FC from 400 cm\(^{-1}\) to 1800 cm\(^{-1}\): (a) Stacking of SERS spectra with seven measurements; (b) Second-derivative of SERS spectra: cross correlation coefficient is 0.99.
Figure 5.5. SERS spectra of YC from 400 cm\(^{-1}\) to 1800 cm\(^{-1}\): (a) Stacking of SERS spectra with seven measurements; (b) Second-derivative of SERS spectra: cross correlation coefficient is 0.86.

Spectral reproducibility, the SERS spectra of the three aromatic peptides exhibit a high degree of similarity when compared to their respective normal Raman spectra. Together, this constitutes convincing evidence that the spectral features in the SERS spectra originate from the unmodified peptide bound to the nanoparticle surface, and not from chemical contaminants or photo-induced degradation products. If the latter scenario were
the case, the spectra would show a significantly higher degree of spatial and/or temporal variability. Both SERS substrate reproducibility and the high packing density of adsorbates on the substrate permits the use of relatively low laser powers in this study. It should be noted that for higher laser powers, above 1mW, for example, spectral reproducibility drops dramatically, presumably due to photo-induced chemical reactions.

5.5.2 Phenylalanine-Cysteine (FC) peptide

The normal Raman (in black) and SERS (in red) spectra for the FC peptide are shown in Figure 5.6(a). Spectral features of trifluoroacetic acid (TFA, in blue) can also be seen in the normal Raman spectrum, arising from the solvent mixture used. Covalent bonding between the Au surface and the peptide is evident from the disappearance of the S-H stretching peak at 2570 cm$^{-1}$ in the SERS spectrum.\textsuperscript{32} The peak at 1003 cm$^{-1}$ is the symmetric ring breathing mode (Fig. 5.6(c)) and the 1033 cm$^{-1}$ feature is the in-plane ring C-H bending mode (Fig. 5.6(d)).\textsuperscript{22, 34-36} The symmetric ring breathing mode is the strongest mode observed in both the SERS and the normal Raman spectra. In comparison to the C-H stretching modes, the Raman cross section of the 1003 cm$^{-1}$ and 1033 cm$^{-1}$ modes are slightly larger in SERS than in the normal Raman spectrum (Fig. 5.6 (b)). It is also interesting to compare the peak positions and linewidths of these two dominant spectral features in the normal Raman and SERS spectra. For both aromatic ring features, the differences of peak positions between normal Raman and SERS spectra are less than
Figure 5.6. SERS and normal Raman spectra of the FC peptide. (a) Comparison of FC Raman (black), SERS (red). Raman spectra of TFA in blue. Peaks denoted by * are contributed by TFA present in the solvent. The peak height in high wavenumber region is not proportional to the peak height in low wavenumber region for viewing clarity. (b) Comparison of the ratio of relative Raman cross section of FC peptide Raman and SERS at different Raman modes (relative to C-H stretch). (c) Symmetric breathing mode of phenylalanine (1003 cm⁻¹). (d) In-plane ring C-H bending mode of phenylalanine at 1033 cm⁻¹.
2 cm\(^{-1}\). However, both peaks are markedly broadened in the SERS spectrum, with the 1003 cm\(^{-1}\) peak broadened to \(\sim 1\) cm\(^{-1}\) (20%) and 1033 cm\(^{-1}\) broadened to \(\sim 4\) cm\(^{-1}\) (60%). This peak broadening suggests that there is a distribution of molecular conformations and/or orientations of the peptide on the Au nanoshell surface that sensitively affects these modes.

### 5.5.3 Tyrosine-Cysteine (YC) Peptide

The Raman and SERS spectra of the YC peptide are shown in Figure 5.7. Covalent binding of the peptide to the Au surface of the nanoshell can be inferred from the disappearance of the S-H stretch mode at 2570 cm\(^{-1}\) in the SERS spectrum.\(^{32}\) The peak at 644 cm\(^{-1}\) is the symmetric ring breathing mode (Fig. 5.7(c)). The spectral features at 833 cm\(^{-1}\) and 853 cm\(^{-1}\) are the characteristic tyrosine doublet Fermi resonance, consisting of the fundamental of the in-plane ring breathing mode (Fig. 5.7(d)) and the overtone of the out-of-plane ring bending mode (Fig. 5.7(e)).\(^{22,37}\) In comparison with the C-H stretching modes, all the primary Raman modes- 644 cm\(^{-1}\), 833 cm\(^{-1}\) and 853 cm\(^{-1}\)- are more enhanced but with different amplitudes (Fig. 5.7(b)). In the normal Raman spectrum, the 853 cm\(^{-1}\) peak has the highest relative Raman cross-section, while in the SERS spectrum the 833 cm\(^{-1}\) peak is more intense, with the intensity ratio of the (853 cm\(^{-1}\)) / (833 cm\(^{-1}\)) peaks varying from 1.58 in the Raman spectrum to 0.81 in the SERS spectrum. This peak ratio is considered to be an important indicator of the degree of "exposure" of the hydroxyl group on the phenol ring of the tyrosine, with a ratio above 1.
Figure 5.7. Raman and SERS spectra of the YC peptide. (a) Comparison of YC peptide Raman (in black) and SERS spectra (in red). The Y axis was offset at high wave number region for viewing clarity. (b) Comparison of the ratio of Raman and SERS at different Raman modes with respect to C-H stretching. (c) Symmetric ring breathing mode characteristic of tyrosine at 644 cm$^{-1}$. (d) In plane ring breathing mode of tyrosine at 833 cm$^{-1}$. (e) Out-of-plane ring bending mode of tyrosine at 853 cm$^{-1}$.
indicating that the hydroxyl group is exposed to its environment and below 1 corresponding to a “buried” hydroxyl group.\textsuperscript{38} Our Raman and SERS spectra suggest that the hydroxyl group of the phenol ring of the tyrosine is more sterically protected when the peptide is conjugated onto the Au nanoshell surface than when the molecule is unbound, as is the case for the Raman spectrum sample. The peak at 833 cm\textsuperscript{-1} in the normal Raman spectrum is red-shifted slightly, by 4 cm\textsuperscript{-1} to 829 cm\textsuperscript{-1} in SERS, while the peak at 853 cm\textsuperscript{-1} in normal Raman is red-shifted only 1 cm\textsuperscript{-1} to 852 cm\textsuperscript{-1} in SERS. Although no significant peak broadening is observed for the 644 cm\textsuperscript{-1} peak, the peak widths of the modes at 853 cm\textsuperscript{-1} and 833 cm\textsuperscript{-1} are increased in the SERS spectrum by 9 cm\textsuperscript{-1} (69\%) and 7 cm\textsuperscript{-1} (31\%) respectively.

5.5.4 Trpophan-Cysteine (WC) peptide

The Raman and SERS spectra of the WC peptide are shown in Figure 5.8. For this peptide as well, covalent interaction of the peptide with the Au nanoshell surface is evident from the disappearance of the S-H stretching mode at 2570 cm\textsuperscript{-1} in the SERS spectrum.\textsuperscript{32} The dominant Raman modes observed for the WC peptide appear at 760 cm\textsuperscript{-1}, 880 cm\textsuperscript{-1}, 1012 cm\textsuperscript{-1} and 1551 cm\textsuperscript{-1}. The peak at 760 cm\textsuperscript{-1} is the symmetric benzene and pyrrole in-phase breathing mode (Fig. 5.8(c)). The peak at 880 cm\textsuperscript{-1} is the indole ring vibration coupled to the NH bending mode (Fig. 5.8(d)). The peak at 1012 cm\textsuperscript{-1} is the symmetric benzene and pyrrole out-of-phase breathing mode (Fig. 5.8(e)) and the peak at
Figure 5.8. Raman and SERS spectra of the WC peptide. (a) Comparison of WC peptide Raman (blue) and SERS spectra (red). The peak height in the high wave number region is not proportional to the peak height in low wave number region for viewing clarity. (b) Comparison of the ratio of WC peptide Raman and SERS at different Raman modes divided by the C-H stretch. (c) Symmetric benzene/pyrrole in-phase breathing mode of tryptophan found at 760 cm$^{-1}$. (d) Indole ring vibration with NH bending mode of tryptophan found at 880 cm$^{-1}$. (e) Symmetric benzene / pyrrole out-of-phase breathing mode of tryptophan found at 1012 cm$^{-1}$. (f) Pyrrole C2-C3 stretching mode of tryptophan found at 1551 cm$^{-1}$.
1551 cm\(^{-1}\) is the pyrrole C2-C3 stretching mode (Fig. 5.8(f)).\(^{22}\) In comparison with the C-H stretching modes, the Raman aromatic modes are all significantly more enhanced with the 760 cm\(^{-1}\) and 1012 cm\(^{-1}\) than the 880 cm\(^{-1}\) and 1551 cm\(^{-1}\) mode. No significant peak shifts were observed for 760 cm\(^{-1}\), 1012 cm\(^{-1}\) and 1551 cm\(^{-1}\). However, the 880 cm\(^{-1}\) Raman peak was shifted nominally 5 cm\(^{-1}\) to 875 cm\(^{-1}\) in the SERS spectrum. The linewidth of the 1012 cm\(^{-1}\) spectral feature is increased significantly to 5 cm\(^{-1}\) (73%), while the other three peaks were broadened by 40-50% (7 cm\(^{-1}\)).

5.5.5 Cross Comparison of the Aromatic Raman Modes of the Peptides

Since all the peptides studied contained the same number of aliphatic C-H bonds, we compare the Raman and SERS activity of the aromatic modes by scaling them to the aliphatic features. Three observations can be made from the spectra shown in Figures 5.6-8. First, the 1003 cm\(^{-1}\) peak of the ring breathing mode in phenylalanine in the FC peptide is the most intense peak of all the aromatic modes in both SERS and Raman spectra. Second, the largest integrated areas for SERS peaks among all the observed aromatic spectral features are the peaks at 833 and 853 cm\(^{-1}\) from tyrosine in the YC peptide. Third, for both Raman and SERS, the ratio of the peak area (integral) relative to the C-H stretching mode peak area gives the relative Raman cross section (relRCS) of that peak:

\[
\text{relRCS}_i = \frac{A_{\text{mode}_i}}{A_{\text{C-H}}}. 
\]
The ratio of the relative Raman cross section of a vibrational mode in its SERS spectrum to its relative normal Raman cross section provides us with a measurement and gives us the relative (SERS/Raman) enhancement of the Stokes mode:

Relative Enhancement = \( \frac{\text{relRCS}_{\text{(SERS)}}}{\text{relRCS}_{\text{(normal Raman)}}} \).

### Table 5.3. Relative enhancements of the Stokes modes

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Wavenumber (cm(^{-1}))</th>
<th>Relative Raman Cross Section</th>
<th>Relative Enhancement (SERS/Normal)</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Normal Raman</td>
<td>SERS</td>
<td></td>
</tr>
<tr>
<td><strong>FC</strong></td>
<td>1003</td>
<td>5.92</td>
<td>6.99</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>1033</td>
<td>1.64</td>
<td>2.29</td>
<td>1.40</td>
</tr>
<tr>
<td><strong>YC</strong></td>
<td>644</td>
<td>0.59</td>
<td>1.26</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>833</td>
<td>1.63</td>
<td>6.53</td>
<td>5.23</td>
</tr>
<tr>
<td></td>
<td>853</td>
<td>2.58</td>
<td>6.92</td>
<td>2.68</td>
</tr>
<tr>
<td><strong>WC</strong></td>
<td>760</td>
<td>2.18</td>
<td>6.13</td>
<td>2.81</td>
</tr>
<tr>
<td></td>
<td>880</td>
<td>0.80</td>
<td>1.70</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td>1012</td>
<td>1.62</td>
<td>4.33</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>1551</td>
<td>2.60</td>
<td>5.02</td>
<td>1.93</td>
</tr>
</tbody>
</table>

For the three peptides studied, the relative enhancements of their observed Stokes modes are provided in Table 5.3, all the means and the standard errors are calculated with over seven spectra obtained at different sampling points at the same or different batches of nanoshell samples. Based on this analysis, the 833 cm\(^{-1}\) peak of the YC peptide is the most enhanced feature among all the aromatic modes observed in the three peptides studied. The large enhancement of this peak is very likely due to the combination of the
SERS enhancement experienced by all aromatic features in the YC peptide, and some small additional enhancement due to its more sequestered chemical environment, as discussed in the tyrosine-cysteine subsection, as an adsorbate for SERS relative to its solution phase conformation in the normal Raman sample.

5.5.6 Prediction of Raman and SERS spectra of Penetratin

Penetratin, a cell penetrating peptide,\textsuperscript{39,40} was chosen for two reasons: it is an important peptide in many biomaterials applications requiring cellular uptake, and it is a relatively short peptide with three aromatic residues: two tryptophans and one phenylalanine. To ensure the fidelity of our SERS spectrum, some modifications were made to the penetratin sequence as shown in Table 5.1. The thiol side chain of cysteine was used, as in our peptide studies, to link the peptide to the gold nanoshell surface. Additionally, two glycines were inserted as a spacer between the cysteine residue and the standard penetratin sequence. The molecular model of the peptide is shown in Figure 5.9(a). The phenylalanine (purple) and one tryptophan (green) residue are next to each other in the middle of the sequence, while the second tryptophan residue is located near the C-terminus of the peptide.

Since Raman features of aromatic amino acids are the dominant features in the Raman spectra of peptides and proteins, it should be possible to predict at least some of the major features of the SERS spectrum of more complex proteins and peptides using the spectra of the aromatic peptides as a conceptual basis set. In Figure 5.9, the
predicted (b) Raman and (c) SERS of penetratin, together with their experimental counterparts, are shown. The predicted Raman and SERS spectra were obtained by using

![Molecular model of penetratin peptide](image)

![Raman spectra](image)

![SERS spectra](image)

**Figure 5.9.** Comparison of empirically predicted (red) and directly measured experimental (black) spectra of penetratin. (a) Molecular model of penetratin peptide, including one phenylalanine (purple) and two tryptophan (green). (b) Raman spectra (peaks with contributions from TFA are denoted with *). (c) SERS spectra.

the normal Raman and SERS spectra of WC and FC. First, the two spectra were normalized with respect to the anti-symmetric C-H stretching mode. The predicted spectra shown are the sum of the FC peptide spectrum plus two times the WC peptide spectrum, corresponding directly to the aromatic amino acid content in the penetratin sequence. The spectra in Figure 5.9 show that the experimental spectra and the spectra predicted using only the aromatic peptides agree remarkably well with each other, both
for SERS and for normal Raman, despite the fact that only 3 of the 19 amino acids of the penetratin used in this study are aromatic. This clearly demonstrates the overwhelming dominance of the aromatic residues in the Raman and SERS spectra. In addition, both the SERS and Raman spectra of penetratin, the experimental data show significantly higher intensities in the amide I and III regions located at around 1650 cm$^{-1}$ and 1245 cm$^{-1}$ respectively, due to the fact that the amide backbone contributions are much larger for the penetratin molecule than for the constituent peptide spectra used to construct the predicted spectra. Compared to the normal Raman spectrum of penetratin, the peak ratio of the 1003 cm$^{-1}$ phenylalanine peak vs. the 1012 cm$^{-1}$ tryptophan peak in SERS is markedly higher, indicating that the phenylalanine features are more significantly enhanced than the tryptophan features in SERS.

5.6 Conclusion

SERS spectra obtained for three cysteine-containing aromatic peptides, phenylalanine-cysteine, tyrosine-cysteine, and tryptophan-cysteine, conjugated to Au nanoshell substrates show excellent reproducibility and close similarity to their respective solution-phase Raman spectra. The relative Raman and SERS cross-sections of the characteristic Stokes modes of the three aromatic amino acids were obtained. We have shown that the Raman and SERS spectra of penetratin, a 19 amino acid cell-penetrating peptide, can be reproduced very well, excepted in the backbone spectral regions, by using the appropriate Raman and SERS spectra of these aromatic peptides as an empirical
“basis set”. This study reveals that the spectral features of aromatic amino acid residues and protein backbones are the overwhelmingly dominant features in the Raman and SERS spectra of peptides/proteins, which greatly simplified the spectral interpretation. The spectral sensitivity to the amide bond makes Raman particularly suitable in studying protein conformational modification.
References


Chapter 6. Conclusions

It was a long journey to get to this point. For the past five years, I worked on two projects, which were quite different to each other but made my journey more interesting and challenging.

Synthetic collagen-like peptides were explored to study the stability of the ABC type heterotrimeric collagen systems, such as side chain length of charged amino acids, salt effects and position of amino acids. In our approach, two host-guest peptides were designed to study the triple helical propensity of amino acid--ornithine by putting ornithine at the X and Y position in the guest peptide region. The stability of these peptides with the lysine containing peptides was compared. It was found that ornithine had similar triple helical propensity as lysine.

Then five new peptides ((POrG)$_{10}$, (OrOG)$_{10}$, (KOG)$_{10}$, (PEG)$_{10}$ and (PDG)$_{10}$) were designed, synthesized and characterized. To our surprise, both (KOG)$_{10}$ and (OrOG)$_{10}$ showed triple helix formation with $T_m$ at 34°C and 27.5°C in phosphate buffer. It was discovered that the negatively charged phosphate bivalent ions acted as bridges and induced the formation of the triple helices for these two positively charged collagen-like peptide systems. Using cations, such as Ca$^{2+}$, Ni$^{2+}$ and Cu$^{2+}$, to induce the formation of triple helices with the negatively charged collagen-like peptides was not successful.

Eight systems that were composed of one positively charged peptide, one negative charged peptide and one neutral peptide under neutral pH were divided into two
groups. In addition to circular dichroism (CD), differential scanning colorimetric (DSC) was evaluated and explored in order to better understand the thermal unfolding processes of these systems.

In natural collagen there is a strong predisposition for negatively charged amino acids to be in the X positions and positively charged amino acids to be in the Y positions, for example, E at the X position and K at the Y position. In group 1, the four systems were all composed of peptides with charged amino acid at naturally preferred positions. The effects of the side chain length on the stability of the ABC collagen heterotrimer were evaluated. It was found that (DOG)$_{10}$(PKG)$_{10}$(POG)$_{10}$ was the most thermally stable ABC heterotrimer and (EOG)$_{10}$(POrG)$_{10}$(POG)$_{10}$ was the least thermally stable ABC heterotrimer.

In group 2, the position of the charged amino acids were intentionally reversed this predisposition. Comparing to peptides in Group 1, the amino acids in the X and the Y positions from these peptides were swapped. Now positively charged amino acids are in the X positions and negatively charged amino acids are in the Y positions. Although, in this group, pure ABC type triple helices were not easy to form, especially for the aspartic acid containing systems, interesting observations were obtained. First, high thermal stability does not necessarily mean that it is easier to form the pure ABC type heterotrimers. Second, DK pair is not always the ideal pair to form a stable triple helix. The stability of the triple helix also depends on the position of the amino acids and the stability of the individual peptide chain.
The thermal unfolding behaviors for all the eight systems were compared in both 10 mM phosphate buffer and 10 mM phosphate buffered saline. It was found that with the addition of salt, the stability of the systems decreased. Except the (DOG)$_{10}$•(PKG)$_{10}$•(POG)$_{10}$ system, for all the other seven systems, the ABC heterotrimers were destroyed to certain degree, a new or increased peak of (POG)$_{10}$ homotrimer started to emerge. It proved that electrostatic interaction was the major driving force of the formation of the ABC heterotrimers.

Since it was difficult to use mass spectrometry to directly identify the composition of the triple helix, analytical ultracentrifugation (AUC) was first explored with these systems. In order to cover a wide range of peptide concentration, a new peptide was designed to incorporate an aromatic amino acid to the peptide chain without interrupting the triple helical structures: (PKG)$_{10}$GY. CD experiments of the systems showed that with the introduction of tyrosine into the system, the triple helical structures were not interrupted and the thermal stability of the systems was maintained. For the ABC heterotrimeric systems tested, MW obtained from velocity and equilibrium experiments were in the +- 10% range. To further understand this equilibration process, wavelength scans need to be performed to get a globally fitted extinction value for this peptide. Only after the wavelength scan experiments, the optical density data can be converted to molar concentration at the measured wavelength. And then the parameters associated with the equilibration can be obtained. For the ABB heterotrimeric systems tested, the major component was shown to be the MW of hexamer. In order to confirm the composition of
the system, more experiments are needed, such as velocity experiment with a different concentration (230nm). Sedimentation equilibrium experiment will also be helpful to get the MW, the equilibrium constant and thermodynamic parameters to understand this equilibration process.

The project of surface enhanced Raman spectroscopy (SERS) of aromatic peptides was a collaboration project with Professor Naomi Halas. It was really a combination of two expertises. Professor Halas invented Au nanoshells and has done a lot of studies with Au nanoshells. Our lab is specialized in designing and synthesizing peptides. This project explored using Au nanoshells as SERS substrate to study peptides. Three aromatic dipeptides incorporated aromatic amino acid: phenylanaline, tyrosine, tryptophan and cysteine were designed and synthesized. The aromatic amino acids were used for easy identification of the peptides because they have their characteristic Raman peaks. Cysteine was used to attach the peptide on the Au nanoshell surface by covalent attachment. A simple and easy conjugation method was developed and implemented. Very reproducible SERS spectra were obtained. Then the spectra of the dipeptides were used to predict the Raman and SERS spectra of a biological important peptide, penetratin. Penetratin contains three aromatic residues, one phenylanaline and two tryptophans. The sequence of the penetratin peptide was modified to incorporate cysteine at the N-terminus in order to covalently attach it to the surface of Au nanoshells. The predicted Raman and SERS spectra agreed very well with the experimental spectra. This indicates that the spectral features of aromatic amino acid residues and protein backbones are the overwhelmingly
dominant features in the Raman and SERS spectra of peptides/proteins, which greatly simplify the spectral interpretation.
Appendices: Spectra
HPLC and MS (Chapter 2 and 3)

(a) HPLC and (b) MS of (KOG)\textsubscript{10}. Expected Mass: [M+H]\textsuperscript{+}: 3041.7, Mass found: 3041.7.

(a) HPLC and (b) MS of (OOG)\textsubscript{10}. Expected Mass: [M+H]\textsuperscript{+}: 2901.5, Mass found: 2901.8.
(a) HPLC and (b) MS of (PDG)$_{10}$. Expected Mass: [M+H]$^+$:2752.6, Mass found: 2753.5.

(a) HPLC and (b) MS of (PEG)$_{10}$. Expected Mass: [M+Na]$^+$:2913.2, Mass found: 2913.8.
(a) HPLC and (b) MS of (POrG)$_{10}$. Expected Mass: $[M]^+ = 2741.5$, Mass found: 2741.7.

(a) HPLC and (b) MS of (POG)$_8$. Expected Mass: $[M+Na]^+ = 2276.0$, Mass found: 2276.5.
(a) HPLC and (b) MS of \((\text{POG})_3\text{OrG}(\text{POG})_4\text{G}\). Expected Mass: \([\text{M+Na}]^+ : 2293.0\), Mass found: 2293.9.

(a) HPLC and (b) MS of \((\text{POG})_3\text{POrG}(\text{POG})_4\text{G}\). Expected Mass: \([\text{M}]^+ : 2254.1\), Mass found: 2255.5.
(a) HPLC and (b) MS of $\text{(ROG)}_{10}$. Expected Mass: $[M]^+:3102$, Mass found:3102.

(a) HPLC and (b) MS of $\text{(POG)}_{10} \text{GY}$. Expected Mass: $[M]^+:3322.2$, Mass found:3322.2.
HPLC and MS (Chapter 5)


(a) HPLC and (b) MS of YC peptide. Expected Mass: [M+Na]$^+$:347.4, Mass found:347.9; [M+K]$^+$:364.4, Mass found:364.01.