

**UV RESONANCE RAMAN SPECTROSCOPY STUDY OF PEPTIDE
CONFORMATIONAL TRANSITIONS**

by

Zhenmin Hong

B. S., Tsinghua University, 2003

M. S., Tsinghua University, 2006

Submitted to the Graduate Faculty of
The Dietrich School of Arts & Science in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH
THE DIETRICH SCHOOL OF ARTS & SCIENCES

This dissertation was presented

by

Zhenmin Hong

It was defended on

April 11th, 2014

and approved by

Lillian T. Chong, Associate Professor of Chemistry, Department of Chemistry

Michael A. Trakselis, Assistant Professor of Chemistry, Department of Chemistry

Ronald B. Wetzel, Professor, Department of Structural Biology

Dissertation Advisor: Sanford A. Asher, Distinguished Professor of Chemistry, Department
of Chemistry

Copyright © by Zhenmin Hong

2014

Reproduced with permission from:

Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. *J. Phys. Chem. Lett.* **2011**, *2*, 334-344

Hong, Z.; Ahmed, Z.; Asher, S. A. *J. Phys. Chem. B* **2011**, *115*, 4234-4243

Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. Handen, J. Lednev, I. K. *Chem. Rev.* **2012**

112, 2604-2628

Hong, Z.; Wert, J.; Asher, S. A. *J. Phys. Chem. B* **2013**, *117*, 7145-7156

UV RESONANCE RAMAN SPECTROSCOPY STUDY OF PEPTIDE CONFORMATIONAL TRANSITIONS

Zhenmin Hong, PhD

University of Pittsburgh, 2014

The conformational transition between α -helix-like conformations and the polyproline II conformation (now recognized by many as the conformation of unfolded peptides) is investigated here. We utilized UV resonance Raman spectroscopy together with circular dichroism and nuclear magnetic resonance spectroscopy to investigate the conformations of three polyalanine peptides and the impacts of salt bridge side chain interaction and external surfactants on the transitions between these conformations. We found that the macrodipole-terminal charge interactions typically affect the α -helix stability more strongly than the salt bridge side chain interactions do. The α -helix-turn- α -helix conformation can form in short peptides with ~ 20 residues. The arginine vibration band at $\sim 1170 \text{ cm}^{-1}$ was found to report on the guanidinium group hydration. Addition of anionic surfactants induces α -helix-like conformations in short cationic peptides through the formation of peptide-surfactant aggregates. The studies here highlight the crucial roles of hydrogen bonding, hydrophobic effect and electrostatic interactions in the peptide conformational transitions. In addition, the impact of sample self absorption on the observed resonance Raman intensities has also been theoretically investigated. In general, the Raman intensities increase as the excitation approaches resonance. However, narrow bandwidth impurity absorption can cause the observed Raman intensities to decrease.

TABLE OF CONTENTS

PREFACE.....	XXI
1.0 PROTEIN FOLDING.....	1
1.1 PROTEIN STRUCTURES	1
1.1.1 Dihedral angle and Ramachandran plot	2
1.1.2 Secondary structure	4
1.1.3 Tertiary structure	8
1.2 INTERACTIONS IN PROTEIN.....	8
1.2.1 Hydrogen bonding	9
1.2.2 Hydrophobic effect	10
1.2.3 Electrostatic interactions	13
1.3 PROTEIN FOLDING MECHANISM.....	14
1.3.1 Folding pathway	14
1.3.2 Hierarchical folding.....	16
1.3.3 Energy landscape theory	18
1.4 BIBLOGRAPHY	20
2.0 RAMAN SPECTROSCOPY	24
2.1 RAMAN SCATTERING POLARIZABILITY	24
2.1.1 Classical theory	24

2.1.2	Vibronic coupling theory	27
2.1.3	Time-dependent theory	30
2.2	VIBRATIONAL RESONANCE RAMAN SCATTERING.....	32
2.2.1	Resonance Raman scattering.....	32
2.2.2	Normal modes analysis.....	35
2.2.3	Duschinsky rotation.....	39
2.3	BIBLOGRAPHY	41
3.0	UV RESONANCE RAMAN STUDY OF PROTEIN SECONDARY STRUCTURES.....	43
3.1	ADVANTAGES OF UV RESONANCE RAMAN SPECTROSCOPY	44
3.1.1	High sensitivity.....	44
3.1.2	High selectivity	45
3.1.3	Excellent temporal resolution.....	47
3.2	CORRELATIONS BETWEEN PROTEIN SECONDARY STRUCTURES AND UV RESONANCE RAMAN SPECTRA.....	49
3.2.1	Resonance enhanced amide vibrations	49
3.2.2	Correlations between amide hydrogen bonding and UV resonance Raman spectrum.....	53
3.2.3	Correlations between protein secondary structures and UV resonance Raman spectra.....	58
3.2.4	Ramachandran ψ angle distribution and Gibbs free energy landscapes of peptide bond	64
3.2.5	Resonance enhanced protein side chain vibrations	66

3.3	BIBLIOGRAPHY	67
4.0	EXPERIMENTAL TECHNIQUES	71
4.1	UV RESONANCE RAMAN SPECTROSCOPY	71
4.1.1	UV resonance Raman spectroscopy instrumentation	71
4.1.2	Calculation of Raman scattering cross section	73
4.2	OTHER EXPERIMENTAL TECHNIQUES.....	73
4.2.1	Circular dichroism spectroscopy	73
4.2.2	Pulsed field gradient nuclear magnetic resonance spectroscopy	74
4.3	MATERIALS	75
4.3.1	Peptides.....	75
4.3.2	Purchased chemicals.....	76
4.3.3	Synthesized materials	77
4.4	BIBLIOGRAPHY	78
5.0	CD AND UV RESONANCE RAMAN INDICATE LITTLE ARG-GLU SIDE CHAIN α -HELIX PEPTIDE STABILIZATION	79
5.1	INTRODUCTION	80
5.2	EXPERIMENTAL SECTION.....	82
5.2.1	Materials.....	82
5.2.2	Circular dichroism measurements.....	82
5.2.3	UV resonance Raman spectra	82
5.3	RESULTS AND DISCUSSION	83
5.3.1	Temperature dependence of conformational transitions.....	83
5.3.2	Comparison between AEP and AP	89

5.3.3	pH and ionic strength dependence of AEP conformations	92
5.3.4	AEP forms α -helix-turn- α -helix conformation.....	102
5.4	CONCLUSIONS	103
5.5	APPENDIX.....	103
5.5.1	Derivation of eq. (7).....	103
5.5.2	Calculation of pKa difference (Δ pKa) in Table 5.2	105
5.6	ACKNOWLEDGEMENT	106
5.7	SUPPORTING INFORMATION	107
5.8	BIBLIOGRAPHY	109
6.0	UV RESONANCE RAMAN AND DFT STUDIES OF ARGININE SIDE CHAINS IN PEPTIDES: INSIGHTS INTO ARGININE HYDRATION.....	112
6.1	INTRODUCTION	113
6.2	EXPERIMENTAL SECTION.....	116
6.2.1	Materials.....	116
6.2.2	Synthesis of ethylguanidinium perchlorate and Ac-Arg-OMe perchlorate... ..	116
6.2.3	204 nm excited U.VRR spectra.....	117
6.2.4	DFT calculations	118
6.3	RESULTS AND DISCUSSIONS.....	119
6.3.1	Vibrational assignments.....	119
6.3.2	Calculated electronic ground state structures of gdn and EG	126
6.3.3	Calculated first electronic excited state structure of EG	129
6.3.4	Calculation of the U.VRR spectra of EG and EG-d.....	133

6.3.5	UVRR spectra of EG in water and in acetonitrile.....	138
6.3.6	Hydration of guanidinium group.....	141
6.3.7	Different environments in peptides.....	143
6.4	CONCLUSION	148
6.5	ACKNOWLEDGEMENT	148
6.6	BIBLOGRAPHY	149
7.0	SDS MONOMERS INDUCE XAO PEPTIDE POLYPROLINE II TO α -HELIX TRANSITION	151
7.1	INTRODUCTION	152
7.2	EXPERIMENTAL SECTION.....	154
7.2.1	Materials.....	154
7.2.2	Circular dichroism measurements.....	155
7.2.3	204 nm excited UVRR spectra.....	155
7.2.4	NMR diffusion measurements.....	156
7.3	RESULTS AND DISCUSSIONS.....	157
7.3.1	CD spectra of XAO in different SDS concentration solutions.....	157
7.3.2	XAO $^1\text{H}_\alpha$ NMR spectra	160
7.3.3	UVRR spectra of XAO dependence on SDS concentration.....	160
7.3.4	α -Helical fractions calculated from UVRR, CD and $^1\text{H}_\alpha$ NMR	161
7.3.5	The effects of long alkyl chains and negatively charged surfactant head groups.....	163
7.3.6	Diffusion coefficients measured by 1H-NMR DOSY experiment.....	166
7.3.7	Modeling of XAO and SDS diffusion coefficients.....	170

7.3.8	Mechanism of SDS-induced XAO α -helix formation.....	172
7.3.9	Ramachandran ψ angle distributions of XAO in SDS solution	174
7.4	CONCLUSIONS	179
7.5	ACKNOWLEDGEMENT	180
7.6	SUPPORTING INFORMATION	181
7.6.1	Modeling of XAO and SDS interactions.....	181
7.6.2	Modeling of SDS Micellation.....	182
7.6.3	Supplementary figures	183
7.7	BIBLOGRAPHY	185
8.0	DEPENDENCE OF RAMAN AND RESONANCE RAMAN INTENSITIES ON SAMPLE SELF ABSORPTION	189
8.1	INTRODUCTION	190
8.2	RESULTS AND DISCUSSION	192
8.2.1	Dependence of Raman Intensities on Self Absorption Where the Raman Intensities Derive from Preresonance and Resonance Raman Enhancement from a Single Resonant Absorption Band.....	192
8.2.2	Dependence of Raman Intensities on Self Absorption Where the Sample Contains Absorption Bands That Do Not Enhance The Analyte Raman Bands	197
8.2.3	Conditions for Sample Self Absorption That Causes Resonance Raman Intensity Minima	200
8.2.4	Dependence of Preresonance Raman Intensities on Self Absorption	206
8.3	CONCLUSIONS	209
8.4	ACKNOWLEDGEMENT	209

8.5	BIBLIOGRAPHY	210
9.0	CONCLUDING SUMMARY	211
10.0	FUTURE WORK	213
10.1	SEARCHING FOR NEW SPECTRAL MARKER(S) FOR PEPTIDE BACKBONE CONFORMATION	213
10.2	INVESTIGATING THE ROLE OF MEMBRANE ON THE AGGREGATION OF INTRINSICALLY DISORDERED PEPTIDES	216
10.3	BIBLIOGRAPHY	219

LIST OF TABLES

Table 1. 1. Three types of helices	5
Table 3. 1. NMA Bond lengths of ground state and second electronic excited state	52
Table 3. 2. Potential energy distribution of NMA amide vibrations	52
Table 3. 3. Bond length in NMA for different H-bonded waters	54
Table 3. 4. NBO bond order in NMA with different H-bonded waters.....	54
Table 3. 5. NBO atomic charges in NMA with different H-bonded waters	56
Table 3. 6. NMA vibrational frequency dependence on H-bonding state in vacuum	57
Table 4. 1. Peptides studied in this work.....	75
Table 4. 2. Purchased chemical used in this work	76
Table 5. 1. Thermodynamic parameters obtained by Zimm-Bragg model fitting of the CD melting curves.....	87
Table 5. 2 Equilibrium constants obtained by fitting the CD and Raman pH dependent helical fractions with eq. (7).....	95
Table 6. 1. Gdn and gdn-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	120
Table 6. 2. EG and EG-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	121

Table 6. 3. Arg and Arg-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	123
Table 6. 4. AAO and AAO-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	125
Table 6. 5. Calculated bond lengths of gdn and EG in the ground and first excited states *	127
Table 6. 6. Calculated atomic charges of gdn and EG in the ground and first excited states * ...	127
Table 6. 7. First electronic excited state gdn and gdn-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	136
Table 6. 8. First electronic excited state EG and EG-d vibrational normal modes involving at least 5% contribution from CN ₃ stretching.....	137
Table 6. 9. Fractions of different Arg hydration states, equilibrium constants and corresponding Gibbs free energies in AP and AEP peptides.....	146
Table 7. 1. Diffusion coefficients and hydrodynamic radius of species in XAO-SDS solution	169
Table 7. 2. Hydrogen bonding states of peptide bonds in aqueous solution.....	176
Table 7. 3. Conformational states resolved in aggregated in the aggregated XAO.....	179

LIST OF FIGURES

Figure 1. 1. Peptide backbone dihedral angles	2
Figure 1. 2. Ramachandran plot.....	3
Figure 1. 3. CPK models of α -helix, 3_{10} -helix and π -helix	4
Figure 1. 4. Mixed β -sheet including antiparallel and parallel sheets	6
Figure 1. 5. Polyproline II conformation	7
Figure 1. 6. Comparison of enthalpy (ΔH), entropy (ΔS) and Gibbs free energy (ΔG).....	12
Figure 1. 7. Diagram of free energy surface along folding coordinate of three models.....	15
Figure 1. 8. Cartoons showing three folding scenarios.....	18
Figure 1. 9. Folding funnels.....	19
Figure 2. 1. Classical view of light scattering by a molecule	26
Figure 2. 2. Diagram showing two-photon process from time-dependent perspective	31
Figure 2. 3. Three types of Raman scattering	34
Figure 2. 4. Flow chart of the home-developed PED calculation program	38
Figure 2. 5. Ground and excited state potential surfaces and vertical transition	40
Figure 2. 6. Duschinsky rotation matrix of formaldehyde between electronic ground and first excited states	41
Figure 3. 1. Selectively excitation of myoglobin resonance Raman spectra	46
Figure 3. 2. Temporal profile of IR pulse, water heating and T-jump resulted from the IR pulse.....	48

Figure 3. 3. Frontier molecular orbitals and electronic transitions in an amide unit	50
Figure 3. 4. Resonance enhanced amide backbone vibrations.....	51
Figure 3. 5. Definition of atomic labels and H-bonding sites in NMA	53
Figure 3. 6. Resonance structures of NMA.....	53
Figure 3. 7. 204 nm excited UVRR spectra of oligoglycines	59
Figure 3. 8. UVRR spectra of horse myoglobin dependent on pH.....	60
Figure 3. 9. UVRR spectra of PGA dependent on temperature.....	61
Figure 3. 10. Relative positions of C _α -H and N-H bonds in different configurations	62
Figure 3. 11. The correlation between Am III band frequency and ψ angle	63
Figure 3. 12. Aromatic side chains	66
Figure 4. 1. Diagram of UVRR spectroscopy instrumentation.....	72
Figure 5. 1. a) Temperature dependence of CD spectra of AEP at pH 7.0. The solid lines indicate the heating process, the dashed lines indicate the spectra measured after cooling process. Inset compares CD spectra of AP and AEP. b) Temperature dependent α -helical fraction and melting curves of AEP at different pH. The α -helical fractions are calculated from the CD spectra by using eq. (1). The melting curve is the fit of the temperature dependent α -helical fraction by using eq. (3). The inset shows the AEP pH 7.0 melting curves, calculated from CD and Raman spectra, and the AP pH 7.0 melting curve calculated from the CD spectra.....	84
Figure 5. 2. a) Temperature dependence of UVRR spectra of AEP at pH 7.0. b) AmIII ₃ and C _α -H band frequencies and integrated intensities from deconvolution of UVRR spectra. The error bars shown are standard deviations from the deconvolution.....	88
Figure 5. 3. UV Resonance Raman spectra of AEP and AP and their difference spectrum at 4 °C, pH 7.0.....	91

Figure 5. 4. a) pH dependence of helical fraction, f_H of AEP at 4 °C in pure water calculated from the CD and UVRR spectra via eq. (1). The curves show the fit to eq. (7). b) Calculated AEP species fractions from CD spectra (see Scheme 5. 1). Red: state A; Blue: state B, Green: state C, Magenta: state D. Dashed lines: α -helix conformation; Dotted lines: unfolded conformation; Solid lines: total AEP fraction in each state including helix and unfolded conformations. Thick solid black line: total helical fraction..... 97

Figure 5. 5. a) pH dependence of the α -helical fraction, f_H of AEP at 4 °C in pure water and 1.0 M NaCl. b) pH dependence of average enthalpy change (Δh) and entropy change (Δs) per residue for forming an α -helical peptide bond conformation at 4 °C in pure water and in 1.0 M NaCl. See Table 5.1..... 100

Figure 5. 6. a) pH dependence of AEP UVRR spectra at 4 °C. b) pH dependence of AmIII₃ and C α -H band frequencies and integrated intensities from deconvolution of UVRR spectra. The error bars shown are standard deviations from the deconvolution. 101

Figure 5. 7. Temperature dependence of CD spectra of AEP at pH 2.0. 107

Figure 5. 8. Temperature dependence of CD spectra of AEP at pH 4.0. 107

Figure 5. 9. Temperature dependence of CD spectra of AEP at pH 10.0. 108

Figure 5. 10. pH dependence of CD spectra of AEP at 4 °C. 108

Figure 6. 1. The calculated electronic ground state equilibrium geometries of gdn, EG, Arg and AAO..... 115

Figure 6. 2. The 204 nm excited UVRR spectra of gdn, EG, Arg, AAO in both H₂O and D₂O. 119

Figure 6. 3. The calculated EG potential energy surface along the C13-C10-N3-C2 torsion angle. 129

Figure 6. 4. Definitions of CN₃ stretching vibrations in EG, Arg and AAO. 130

Figure 6. 5. The calculated structure and molecular orbitals of EG in the first electronic excited state.	132
Figure 6. 6. 204 nm excited UVRR spectra of EG and EG-d (top panel) and calculated resonance Raman spectra of EG and EG-d (bottom panel).	135
Figure 6. 7. a) 204 nm excited UVRR spectra and b) UV absorbance spectra of EG in water and in acetonitrile.	139
Figure 6. 8. The integrated Raman scattering cross sections and frequencies of the $\sim 1170\text{ cm}^{-1}$ band measured from the 204 nm excited UVRR spectra of EG and AAO in different acetonitrile-water solutions.	143
Figure 6. 9. The fitting of the UVRR bands of the Arg side chains in AP and AEP using two Gaussian bands. The dashed bands are tentatively assigned to the backbone AmIII ₃ bands of turn structures.	145
Figure 7. 1. (A) CD spectra of 1.0 mM XAO in 0.1 M aqueous NaClO ₄ solution at different SDS concentrations. The inset shows the dependence on the SDS concentration of the CD spectra 222 nm ellipticity. (B) ¹ H _α NMR spectra of XAO in 0.1 M NaClO ₄ at different SDS concentrations. (C) 204 nm excited UVRR spectra of XAO in 0.1 M NaClO ₄ at different SDS concentrations. The difference spectrum calculated emphasizes the Ala methyl group umbrella bending band at 1382 cm ⁻¹ . (D) SDS concentration dependence of the α-helical fraction of 1.0 mM XAO in 0.1 M NaClO ₄ calculated from the CD, ¹ H _α NMR and UVRR spectra.	158
Figure 7. 2. The 222 nm ellipticities of 1.0 mM XAO in the presence of different surfactants.	164
Figure 7. 3. ¹ H NMR DOSY spectra of XAO.	167
Figure 7. 4. (A) Diffusion coefficients of 1.0 mM XAO and SDS in XAO-SDS solution and the diffusion coefficients of pure SDS in aqueous solution. All solutions contain 0.1 M NaClO ₄ . The	

data were experimentally measured by Pulsed Field Gradient NMR; the curves are calculated by using parameters obtained from experimental data. (B) Calculated concentrations of different species in XAO-SDS solutions by using parameters obtained from experimental data. 168

Figure 7. 5. Comparison between fractions of aggregated XAO and α -helical XAO as determined by CD, NMR and UVRR. 172

Figure 7. 6. Mechanism of the XAO₄-SDS₁₆ aggregation and the formation of XAO α -helix-like conformations. 173

Figure 7. 7. Bands resolved Am III region of the room temperature pure α -helix-like UVRR spectrum of XAO-SDS aggregates. This pure α -helix-like XAO-SDS aggregate UVRR spectrum is calculated by averaging the essentially identical pure α -helix-like XAO UVRR spectra at different SDS concentrations as weighted by their respective α -helical fractions. The α -helix-like UVRR spectrum at each SDS concentration is calculated by subtracting the PPII spectrum contribution. The Am III₃ deconvoluted bands are shown in yellow. The two presumed identical $\sim 1270\text{ cm}^{-1}$ Am III₃ Gaussian bands derive from the C-terminal and N-terminal peptide bonds in the α -helix segment. 175

Figure 7. 8. Ramachandran ψ angle distributions calculated for XAO-SDS aggregates from room temperature 204 nm excited UVRR spectra. 178

Figure 7. 9. 204 nm excited UVRR spectra of XAO in the presence and in the absence of 0.1 M NaClO₄ and the difference spectrum. 183

Figure 7. 10. CD spectra of 1.0 mM XAO in the absence of NaClO₄ at different SDS concentration. 184

Figure 7. 11. Pure α -helix-like conformation UVRR spectra of XAO with 2 mM, 4 mM and 20 mM SDS and their difference spectra. 185

Figure 8. 1. Raman backscattering geometry.....	192
Figure 8. 2. Impact of impurity absorption on observed Raman intensity - case 1	198
Figure 8. 3. Observed Raman intensity excitation frequency local minimum and maximum....	204
Figure 8. 4. Impact of impurity absorption on observed Raman intensity - case 2	205
Figure 8. 5. Impact of impurity absorption on observed Raman intensity - case 3	208
Figure 10. 1. 204 nm excited UVRR spectrum of Ac-AP-NH ₂ at 4 °C.....	214
Figure 10. 2. NMA second (π - π^*) excited state equilibrium geometry.....	215
Figure 10. 3. φ angle dependence of Ac-Ala-NHMe amide out-of-plane bending frequency ...	216
Figure 10. 4. Photograph of the 1 mM XAO in 0 , 6.0 mM and 20 mM SDS, incubated in room temperature for five days.	218

LIST OF SCHEMES

Scheme 5. 1. Species in AEP solution	93
Scheme 6. 1. Thermodynamic cycle that describes the transition among four Arg states in the peptides AP and AEP.....	147
Scheme 7. 1. Molecular structures of different charged head group surfactants.....	165

PREFACE

I would like to thank my advisor, Prof. Sanford Asher. It is a great pleasure to have the chance to work in the research group. He offered me excellent opportunities to carry out these challenging projects and patiently guided and encouraged me during my whole PhD study.

I would like to express my gratitude to my committee members, Prof. Lillian Chong, Prof. Michael Trakselis and Prof. Ron Wetzel.

I am thankful to all former and current Asher group members, especially to Dr. Zeeshan Ahmed, Dr. Bhavya Sharma, Dr. Lu Ma, Dr. Sergei Bykov, Dr. Edward Gooding, Jonathan Wert, David Punihaole and Tomas Brizer for helpful discussions and continuous encouragement in my research.

I am also grateful to Dr. Damodaran Achary for his help in PFG-NMR experiments, and to Tom Gasmire and Jeff Sicher in the Department of Chemistry machine shop.

In addition, I would like to thank National Institute of Health and Office of Naval Research for financial support and in part by the University of Pittsburgh, Center for Simulation and Modeling through the supercomputing resources provided.

1.0 PROTEIN FOLDING

Proteins consist of one or more chains of amino acid residues. The length of a single chain ranges from several amino acids to thousands of amino acids. Proteins are one of the most abundant and important biological molecules in a living organism. They exist in every cell and actively participate in all biological processes. Proteins serve many essential biological functions, including, but not limit to, catalysis of biochemical reactions, cell signaling, immune responses, and mechanical functions. The study of proteins and their functions is vital toward understanding of various biological processes, disease diagnoses and drug discovery.

1.1 PROTEIN STRUCTURES

Unlike random coiled polymers, proteins adopt certain secondary and tertiary structures to maintain specific biological functions. Protein misfolding causes many diseases including Alzheimer's disease, Huntington's disease, Parkinson's disease and Creutzfeldt-Jakob disease, etc.¹ The structures of proteins are determined by the amino acids sequence. Pauling first described two types of protein secondary structures, α -helix and β -sheet.² Other secondary structures, like β -turns and polyproline II, were discovered later.

1.1.1 Dihedral angle and Ramachandran plot

The bond between the carbonyl C atom and the N atom shows partial double bond character due to the delocalization of π orbital electrons of the carbonyl. This partial double bond character restrains the amide from free rotation. The two σ bonds, between the N atom and the C_α atom, the C_α atom and the C atom, are allowed to rotate freely. To quantitatively describe the rotation, a pair of dihedral angles (φ , ψ) has been defined. The φ angle is defined as the dihedral angle between the amide plane and the C_α -carbonyl plane, and the ψ angle is defined as the dihedral angle between of C_α -C-N plane and amide plane (Figure 1. 1). Collectively, the (φ , ψ) dihedral pairs of all residues dictate the protein secondary and tertiary structures.

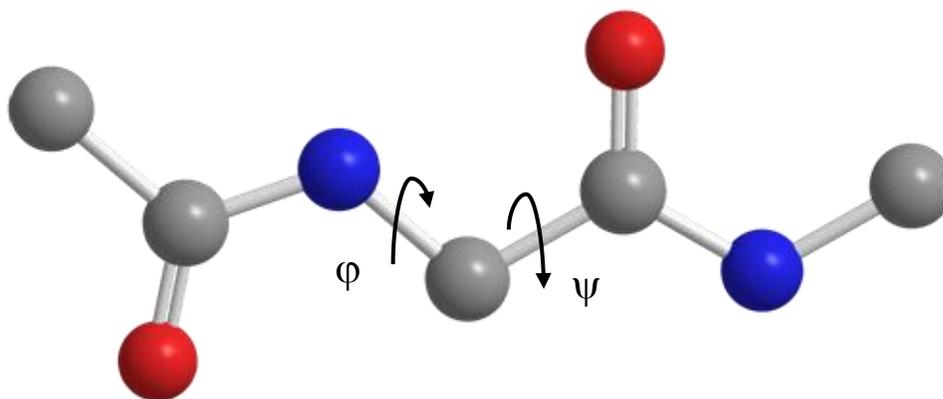


Figure 1. 1. Peptide backbone dihedral angles

A Ramachandran plot for an amino acid is generated by plotting the probabilities that the peptide bond adopts a certain dihedral angle pair. Due to the steric hindrance of side chains, the amino acids are unable to access all (φ , ψ) space (Figure 1. 2D). The probability is much higher for negative φ angle than positive one, due to the chirality of amino acids. The smaller the side chain is, the less constraint the amino acid has. The smallest amino acid, glycine (Gly), which lacks side chain and is not chiral, can access most of (φ , ψ) space and shows a symmetric

Ramachandran plot (Figure 1. 2A). Proline (Pro), which is a confined tertiary amide, can only access a small (φ , ψ) region of the Ramachandran plot (Figure 1. 2B).

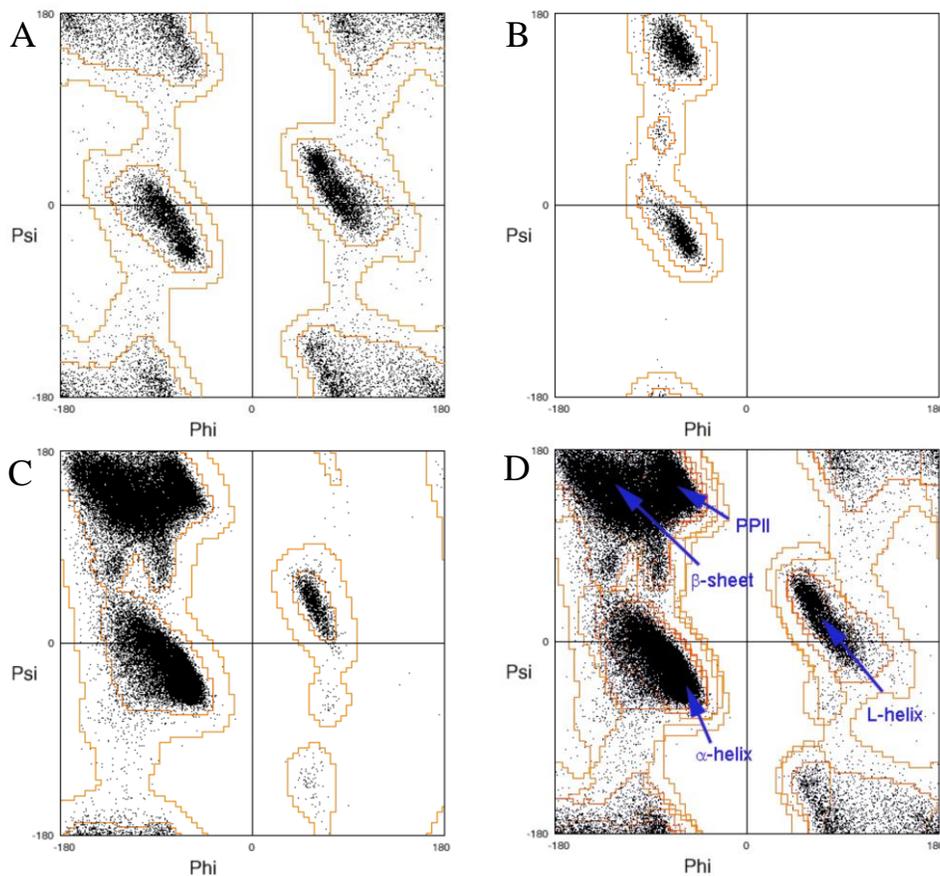


Figure 1. 2. Ramachandran plot

(A) glycine, (B) proline, (C) other amino acids except glycine and proline, (D) all amino acids. These plots are reproduced from <http://kinemage.biochem.duke.edu/validation/model.html>

The Ramachandran plot is classified as "core", "allowed", "generously allowed" and "forbidden" regions.³ This classification is a qualitative energetic scale. The "core" regions involve the most energetically preferred conformations, whereas the "forbidden" regions involve the most unfavorable conformations. Different secondary structures have different dihedral angles and locate in different regions of the Ramachandran plot (Figure 1. 2D). Most of the amino acids fall in the "core" or "allowed" regions.

1.1.2 Secondary structure

Flory predicted that a completely flexible homogeneous polymer chain in solution without any specific interactions is in a random coil state based on the three dimensional random walk model.⁴ However, a peptide is neither completely flexible nor homogeneous. The partial double bond character of the amide and the steric hindrance of side chains restrict polypeptide backbone from free rotation. In addition, the carbonyl oxygen (C=O) can serve as a hydrogen bond (H-bond) acceptor to form two H-bonds; the amide hydrogen (N-H), as a donor, can form one H-bond. Therefore, peptides can form intrapeptide H-bonds that also restrict free rotation. The side chains can also be involved in electrostatic interactions, hydrophobic interactions, cation- π interactions, etc. All these interactions together contribute to the determination of peptide secondary structures.

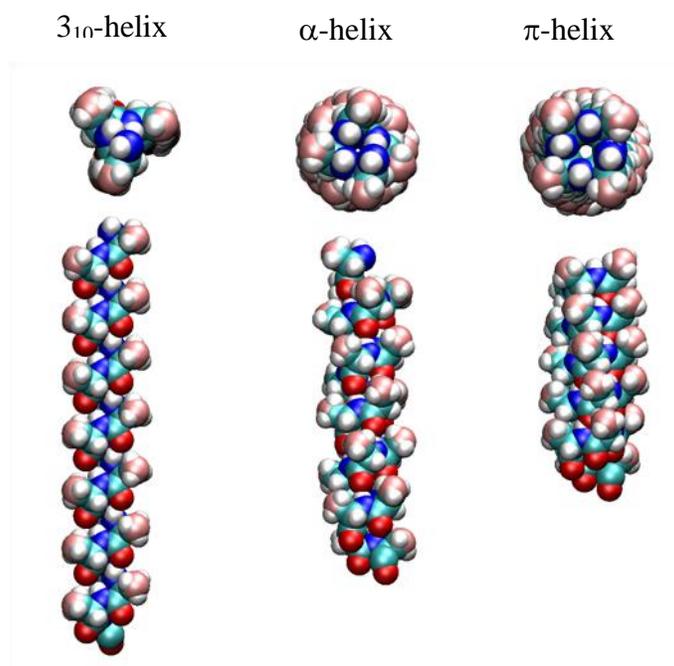


Figure 1. 3. CPK models of α -helix, 3_{10} -helix and π -helix

The most common secondary structure is the α -helix. The α -helix is a right handed helix with 3.6 residues per turn and the pitch is 5.4 Å. The N-H of the i th residue forms an intrapeptide H-bond with the C=O of the $i+4$ th residue. The N-H groups of the first four N-terminus residues and the C=O groups of the first four C-terminus residues in an α -helix cannot form intrapeptide H-bonds effectively.

In a long α -helix, each amide carries a dipole. All these dipoles from different amides form a macrodipole which is parallel to the helix axis with a direction pointing from C-terminus toward N-terminus. The perpendicular components of these dipoles are cancelled out. This macrodipole plays an important role in helix stabilization and melting.⁵⁻⁶

The 3_{10} -helix and π -helix are right handed helices. The 3_{10} -helix is a tightly twisted α -helix with 3 residues per turn. In contrast, the loosely twisted π -helix has 4.4 residues in each turn. 3_{10} -helix occasionally can be found in proteins, while the π -helix is rare. Table 1. 1 and Figure 1. 3 show the difference among three types of helices.

Table 1. 1. Three types of helices

Helix type	Abundance	Residues/turn	Height/turn	H-bond type	(φ, ψ)
α -helix	Abundant	3.6	5.4 Å	i,i+4	(-60,-45)
3_{10} -helix	Infrequent	3.0	6.0 Å	i,i+3	(-50,-25)
π -helix	Rare	4.4	4.8 Å	i,i+5	(-60,-70)

The second class of the protein secondary structure is β -sheet, which are also abundant in proteins. β -sheets are formed by β -strands that are H-bonded with the adjacent strands. The two flanking β -strands in a β -sheet are only partially involved in inter-strand H-bonding.

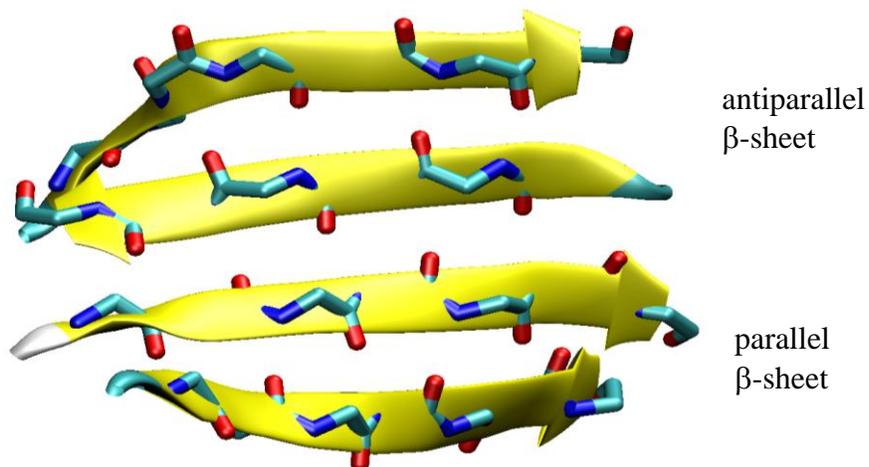


Figure 1. 4. Mixed β -sheet including antiparallel and parallel sheets

The consecutive strands in a β -sheet can be either parallel with dihedral angles of $(-120, 115)$ or antiparallel with dihedral angles of $(-140, 135)$. In the antiparallel arrangement, the consecutive strand reverses the propagation direction so that the N-terminus of one strand is adjacent to the C-terminus of the next strand. This arrangement produces stronger inter-strand H-bonds between the C=O and N-H of adjacent strands, due to the optimized orientation of H-bonds. Each amide bond of a β -strand also carries a dipole like that in an α -helix. All the dipoles sum up into a macrodipole pointing from C-terminus toward N-terminus. The interaction of macrodipoles of adjacent antiparallel strands is favorable, which stabilizes the antiparallel arrangement. In the parallel arrangement, the successive strands are arranged in the same direction. Therefore, the N (or C)-terminus of one strand is adjacent to the N (or C)-terminus of the next strand. The macrodipole interaction between adjacent strands is unfavorable. In addition, the configuration of the carbonyl and amine in adjacent strands is slightly off from the optimal orientation. As a result, the parallel β -sheet is slightly less stable than the antiparallel β -sheet.

Turns are the third class of secondary structure.⁷⁻¹⁰ A turn links two secondary structures together into a supersecondary structure. Based on the number of residues involved in a turn, turns are divided into different types. The most common type is the β -turn, which involves four residues. The β -turn includes various subclasses, namely, type I and I', type II and II', type III, type V, type VIa, type VIb and type VIII. All these β -turns involve a H-bond between CO of the i th residue and NH of the $(i+3)$ th residue. The smaller γ -turn involves three residues.

Proline is a special amino acid. It can only access to limited regions of the Ramachandran plot (Figure 1. 2B), due to its tertiary amide structure. Polyproline peptides assume two different conformations, polyproline I (PPI) and polyproline II (PPII). Recently, the PPII conformation with dihedral angles $(-75, 150)$ has received attentions since it has been proposed that unfolded non-proline peptides adopt the PPII conformation¹¹⁻¹² instead of the random coil. The PPII conformation is a left handed helix, where all the amides are H-bonded to water.

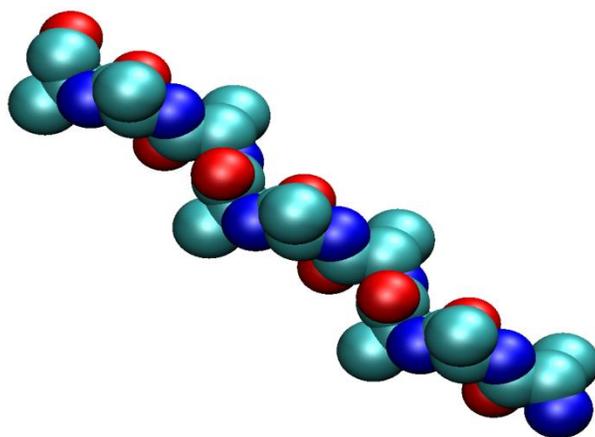


Figure 1. 5. Polyproline II conformation

1.1.3 Tertiary structure

Proteins are composed of multiple secondary structures that assemble into specific configurations that are known as the tertiary structure of proteins. Tertiary structure is the overall three dimensional arrangement of all atoms of a protein. Amino acids that are far apart in the primary sequence of a protein can be very close to each other in the tertiary structure and interact with each other. The interaction may further help to maintain the tertiary structure.

Based on the tertiary structure, proteins can be classified into three groups: fibrous proteins, globular proteins and intrinsically disordered proteins (IDPs). Fibrous proteins usually contain β -sheets or β -strands and a few β -turns, which are usually assembled into fibers. Globular proteins contain several types of secondary structures and are usually folded into a spherical or globular shape. In contrast, intrinsically disordered proteins are flexible, lacking persistent secondary / tertiary structure.¹³⁻¹⁵ Fibrous proteins are generally insoluble in water and rigid, providing functions such as structural support. Most enzymes and regulatory proteins are globular proteins with hydrophobic cores buried inside and hydrophilic surfaces, which maintain the solubility of proteins. The IDPs remain unstructured so to facilitate the folding upon binding to their biological targets.¹⁶⁻¹⁷ The fibrillation of globular proteins involves the aggregation of proteins, which may cause many serious diseases.^{1,18-26}

1.2 INTERACTIONS IN PROTEIN

The stability of protein relies on various interactions.²⁷ These interactions include H-bonding,²⁸ van der Waals interaction, hydrophobic interaction, electrostatic interaction (ion pair or salt

bridge)²⁹⁻⁴⁹ and cation- π interaction^{45,50-53}, etc. Proteins are only marginally stable in aqueous solution at room temperature. A small perturbation may shift protein structure equilibrium and denature a protein.

1.2.1 Hydrogen bonding

H-bonding is an interaction that a hydrogen atom is shared by two electronegative atoms. Usually the H-bond donor is an amine or hydroxyl, while the acceptor is a carbonyl or other group with large electronegative atoms having lone pairs. The H-bond usually is described as an electrostatic interaction, though it shows some covalent bonding features. The H-bond acceptor can form a limited number of H-bonds depending on its number of lone pairs, while the donor can only form one H-bond.

The peptide backbone involves amides, each of which has an N-H group as a potential H-bonding donor and a carbonyl as H-bonding acceptor. It is therefore expected that a peptide in aqueous solution would form not only intra-peptide H-bonds between N-H and C=O but also intermolecular H-bonds with water molecules. Side chains with H-bonding donor or acceptor can also form H-bonds.

The thermodynamics of H-bonding has also been extensively studied. It is adequate to describe the system by classical thermodynamics treatment. The enthalpy and entropy are temperature independent and the Gibbs free energy is thus simply a linear function of temperature. Figure 1. 6A shows the diagram of thermodynamic parameters for H-bonding.

H-bonding is the dominant driving force for secondary structure formation. The intra-peptide H-bonding is usually slightly stronger than the peptide-water inter-molecular H-bonding. If the number of intra-peptide H-bonds is sufficiently large, the intra-peptide H-bonding drive the

peptide to form appropriate compact secondary structures, usually α -helix or β -sheet from PPII or other extended structures where peptide are fully inter-molecularly H-bonded to water.

The importance of H-bonding in the peptide secondary structure formation is proved by experiments. Water has a very strong H-bonding donation ability and a moderate H-bonding acceptance ability.⁵⁴ Transferring peptides from aqueous solution into solvents that have weaker H-bonding donation ability and/or weaker H-bonding acceptance ability stabilizes the secondary structures. Indeed, trifluoroethanol (TFE)⁵⁵⁻⁵⁶ that has extremely strong H-bonding donation ability and no H-bonding acceptance ability strongly promotes the formation of α -helix in peptides. Acetonitrile⁵⁷ that has both weak H-bonding donation and acceptance abilities also increases the α -helix stability.

For an α -helix in aqueous solution, polar side chains usually pack along the backbone, shielding the backbone intra-peptide H-bonding from water molecules and thus increasing the stability of the α -helix.⁴⁶⁻⁴⁷ Also, peptides with side chains having appropriate H-bonding acceptors, such as Asn, Asp, Glu, Ser, Gln, etc, close to the α -helix N-terminus can stabilize the α -helix by forming H-bonds between these side chains and the N-terminal amide.⁵⁸⁻⁵⁹ Side chains possessing H-bonding donors, such as Lys, His, Arg, Gln, etc, located nearby the α -helix C-terminus would stabilize the α -helix by forming H-bonds with the C-terminal carbonyl.⁶⁰⁻⁶¹

1.2.2 Hydrophobic effect

The hydrophobic effect is a phenomenon that causes nonpolar solutes to aggregate in aqueous solution to minimize their exposure to water.⁶² The hydrophobic effect is an important driving force for micelle formation and protein folding. The thermodynamics of hydrophobic effect is

well understood. For simple solutions, the Gibbs free energy shows a linear dependence on temperature, as the enthalpy and entropy are temperature independent. For solutions involving hydrophobic interaction, this transfer of a solute from nonpolar solvent into an aqueous solution gives rise to a large heat capacity change. Therefore, the enthalpy and entropy are functions of temperature. As a consequence, the Gibbs free energy is non-linearly dependent on temperature. At the first order approximation, the heat capacity C_p is taken to be constant, therefore the enthalpy ΔH and entropy ΔS are^{27,63-64}

$$\Delta H(T) = \Delta H(T_1) + \int_{T_1}^T C_p dT = \Delta H(T_1) + C_p(T - T_1) \quad (1.1)$$

$$\Delta S(T) = \Delta S(T_2) + \int_{T_2}^T \frac{C_p}{T} dT = \Delta S(T_2) + C_p \ln \frac{T}{T_2} \quad (1.2)$$

If we let $T_1 = T_h$ and $T_2 = T_s$ where $\Delta H(T_h) = 0$ and $\Delta S(T_s) = 0$, the Gibbs free energy ΔG is

$$\Delta G(T) = C_p \left[(T - T_h) - T \ln \frac{T}{T_s} \right] \quad (1.3)$$

Figure 1. 6B shows the diagram of the temperature dependent enthalpy, entropy and Gibbs free energy for hydrophobic effect. The minimal ΔG locates at exactly $T = T_s$.

Hydrophobic effect has been widely accepted as the most important driving force in protein folding based on the following evidence. (1) In protein crystals, most nonpolar residues are buried inside the core, while most polar residues are exposed to water. On average, 86% of the carbon and sulfur atoms are buried inside, while 40% of the neutral oxygen and nitrogen atoms and only 32% of the charged oxygen and nitrogen atoms are found inside the core.⁶⁵ (2) Nonpolar solvents tend to denature proteins. Nonpolar solvents reduce the free energy of unfolded state of proteins by solvating the exposed nonpolar residues. (3) The native state

conformation stability decreases at both low temperatures and high temperatures, indicating the existence of a free energy minimum, which is predicted by eq. (1.3).

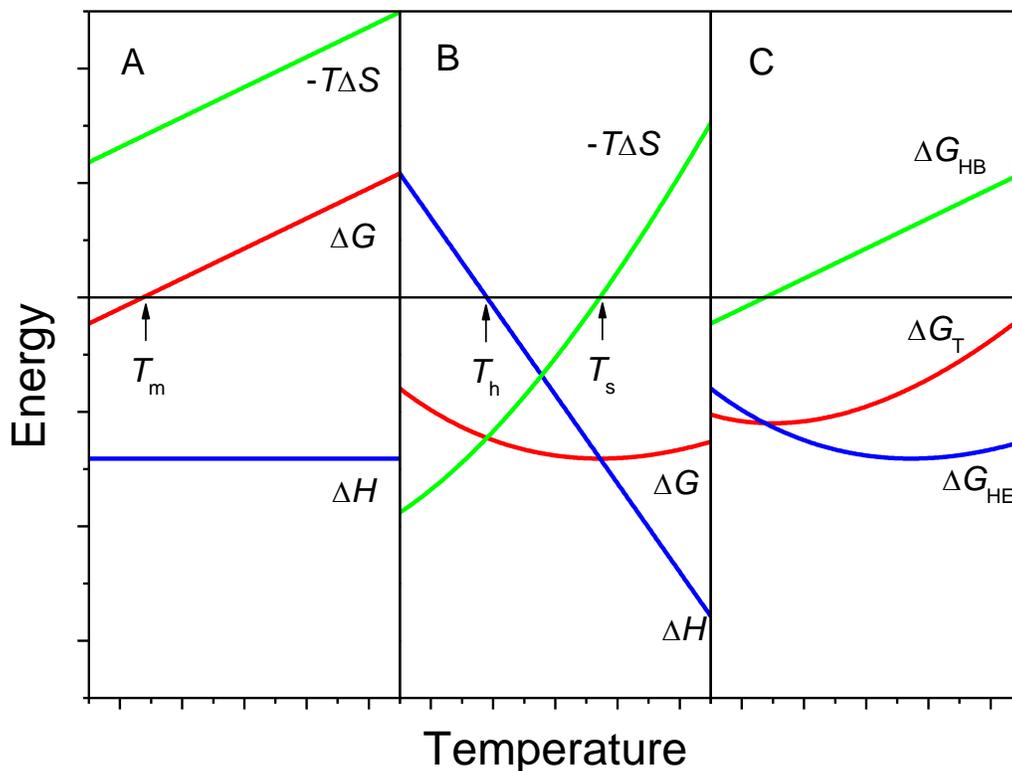


Figure 1. 6. Comparison of enthalpy (ΔH), entropy (ΔS) and Gibbs free energy (ΔG)

(A) H-bonding and (B) hydrophobic effect (C) comparison of ΔG for H-bonding and hydrophobic effect as well as their summation. Simulated using eqs. (1.1), (1.2) and (1.3)

Figure 1. 6C shows the contributions of H-bonding and hydrophobic effect to the total Gibbs free energy of protein folding. The minimum of Gibbs free energy curve with hydrophobic effect is shifted by H-bonding effect from T_s to a lower temperature, T_N . Cold denaturation of proteins will occur at temperature lower than T_N . As the temperature increases from T_N to T_s , the hydrophobic effect tends to stabilize the native state of proteins, while the weakened H-bonding decreases its stability. As the temperature increases above T_s , both H-bonding and the hydrophobic effect drive the protein toward the denatured state.

1.2.3 Electrostatic interactions

Electrostatic interactions in proteins involve a long range interaction among charged residues and macrodipoles. Although they are not the dominant driving forces in protein folding, electrostatic interactions play a role in perturbing the native state conformation stability and the folding dynamics.

Electrostatic interactions include ion pair and charge-dipole interaction. The interaction energy, U , follows

$$U \propto \sum \frac{q_A q_B}{\epsilon r} \quad (1.4)$$

where q is the charge, ϵ is the dielectric constant and r is the distance between charges. Since the interaction energy is proportional to the reciprocal of dielectric constant and distance, increasing ϵ and r will decrease the electrostatic interaction. High dielectric constant solvents and high salt concentrations will decrease the interaction energy of ion pairs separated by solvent. The interaction energy also depends on the product of charges. Changing the pH of protein solution can change the net charge of proteins and the energy. Therefore acid and base can destabilize the native state of proteins by introducing more net charge on the protein.

The macrodipoles in α -helices and β -sheets can interact with side chain charges⁵⁻⁶ and terminal charges.⁶⁶ The interaction between a charge and a pole that have the same sign is energetically unfavorable, while the interaction between a charge and a pole that have opposite sign are favorable.⁶⁷⁻⁷¹ For example, Glu and Asp at the α -helix N-terminus are most commonly found in proteins, in contrast, Lys and Arg usually locate at an α -helix C-terminus.⁶⁸ The configuration of Glu and Asp at the N-terminus, as well as Lys and Arg at the C-terminus, are expected to significantly increase α -helix stability.⁷²⁻⁷³

1.3 PROTEIN FOLDING MECHANISM

Anfinsen's dogma postulates that the native structure of protein that is determined by its primary sequence is a thermodynamically stable conformation and occurs at a minimum on the free energy surface.⁷⁴⁻⁷⁵ Kinetically, for a midsize protein with 100 residues, the free energy surface contains a huge number (10^{70}) of configurations. If the protein searches the free energy minimum randomly, it will take about 10^{57} seconds or 10^{50} years to complete the search and discover the native structure if it searches at an upper limit speed of 10^{-13} seconds per trial. This time scale is contradicted to the experimental results showing that most proteins fold in microseconds to seconds depending on size. The contradiction was first proposed by Cyrus Levinthal, and now has been named the "Levinthal Paradox".⁷⁶⁻⁷⁷ Many protein folding mechanisms have been proposed to resolve this paradox.

1.3.1 Folding pathway

All configurations of a protein are not energetically equivalent. The energy surface possesses a global minimum and a few local minima, at which protein has a better chance to adopt configurations compared with other locations. To resolve the paradox, Levinthal first introduced the concept of folding pathway, which is defined as a well-defined sequence of conformational transitions that leads towards the native state.⁷⁸ There are three models: (1) two-state model; (2) off-pathway model; (3) sequential model. Figure 1. 7 shows the diagram of free energy surface along the folding coordinate of these three models.

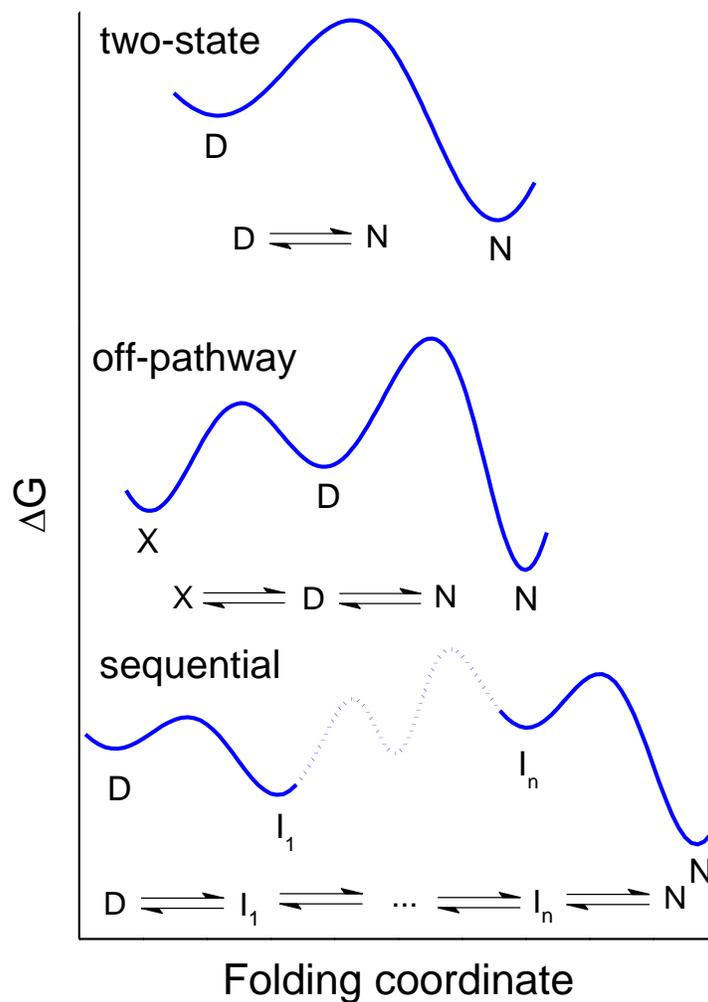


Figure 1. 7. Diagram of free energy surface along folding coordinate of three models

The study of protein folding intermediates has lasted for decades.⁷⁹⁻⁸² The first approach is to study the protein folding kinetics after a perturbation of the solution condition, which in most cases involves temperature change. By fitting the raw data to single- or multi-exponential decay, information about on-pathway intermediates and off-pathway intermediates can be obtained. Another method is to identify the intermediate by ultrafast experimental techniques. It provides more direct and substantial evidence about the intermediate.

1.3.2 Hierarchical folding

The idea of folding pathway solved half of the "Levinthal Paradox". It only answered the question "why proteins fold so fast" but not the question "how do proteins fold so fast". The hierarchical folding theory has been proposed to interpret the kinetics and the molecular detail of protein folding.⁸³⁻⁸⁴ This theory consists of a series of folding steps, the configuration space of each step is dramatically reduced. Figure 1. 8 shows three widely accepted scenarios.⁸⁵⁻⁸⁷

(1) Diffusion-collision model

The diffusion-collision model⁸⁸⁻⁸⁹ is based on the existence of fluctuating quasiparticles, which are locally folded microdomains. Microdomains are only marginally stable, fluctuating between fully unfolded and folded states. These microdomains, which may or may not be the incipient element of secondary structure, move diffusively and may collide with each other. When two matching microdomains collide, they may either coalesce into a certain secondary structure or assemble into a part of the tertiary structure. Those two microdomains involved in the collision may or may not be adjacent in the primary sequence.

The configuration space for a single microdomain is sufficiently small for the residues involved to traverse all possible configurations and determine the relative stable configuration in a short time period. All microdomains search through the corresponding spaces simultaneously, which greatly reduces the searching time. The collision step is the assembly of the microdomains, which has larger entropy cost than that of microdomain formation. Therefore, it is usually the rate-limiting step. Once the coalescence or assembly occurs, the enthalpy contribution from H-bonding or hydrophobic effect plus the specific electrostatic interactions (if any) would stabilize the structure.

(2) Hydrophobic collapse model

The hydrophobic collapse model²⁷ is based on the assumption that hydrophobic effect is the dominant driving force of protein folding. During protein folding, to avoid the unfavorable exposure of nonpolar residues to water, the unfolded protein undergoes a quick collapse into a significantly condensed volume, burying the nonpolar residue inside to form a hydrophobic core, and leaving the hydrophilic residue exposed outside toward the water. It is followed by arrangement of the buried residues.⁹⁰

The hydrophobic collapse is driven by the large entropy gain of the excluded water with a small energy barrier. Thermodynamically, it is a downhill transition, and kinetically, it is a fast and irreversible step. The condensed volume constrains the movement of the residues and hence significantly reduces the configuration space to a reasonable size that protein could search through the native conformation with insignificant enthalpic driven force. The arrangement and packing step should be the rate-limiting step of folding in this scenario.

(3) Nucleation-condensation model

The diffusion-collision model⁹¹⁻⁹³ and hydrophobic collapse model describe two extreme cases in protein folding. The diffusion-collision model emphasizes that the secondary structure formation is prior to the tertiary structure. On the contrary, the hydrophobic collapse model emphasizes the formation of tertiary structure prior to the secondary structure. The nucleation-condensation model merges the above two models. Figure 1. 8 compares the three scenarios.

The nucleation-condensation model postulates the existence of a nucleus. The nucleus possesses features of both secondary structure and the tertiary structure. The nonlocal interactions occur among the collapsed unfolded residues. Initialized by the nucleus, the unfolded residues condense and grow into a fully folded state with no energy barrier.

The formation of nucleus greatly reduces the configuration space. The relative ordered condensation and growth of the unfolded residues into folded state further decrease the number of configurations that protein needs to search.

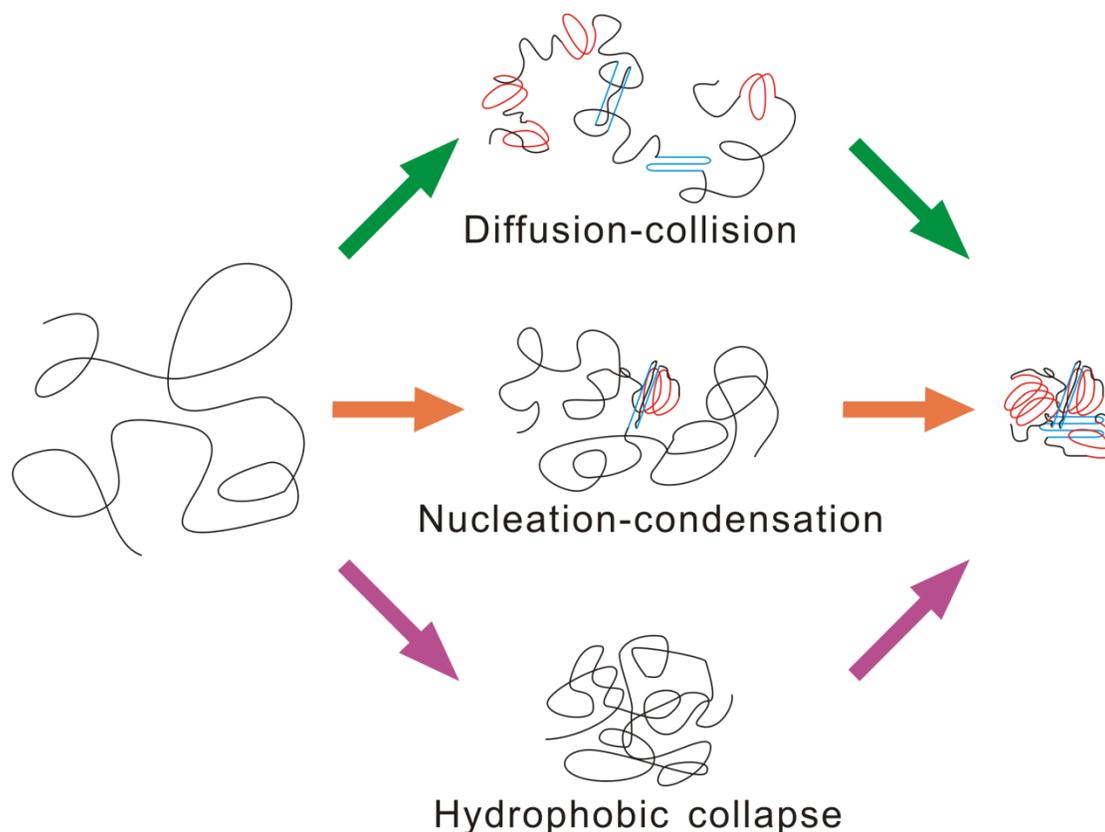


Figure 1. 8. Cartoons showing three folding scenarios

1.3.3 Energy landscape theory

Though the native state of a protein is a well-defined conformation, the denatured state of a protein is actually an ensemble of all unfolded structures instead of a single specific structure. Different unfolded structures may be slightly energetically different and constantly change from one structure to another. Based on this view, energy landscape theory has been proposed to interpret protein folding.^{76,86,94-95}

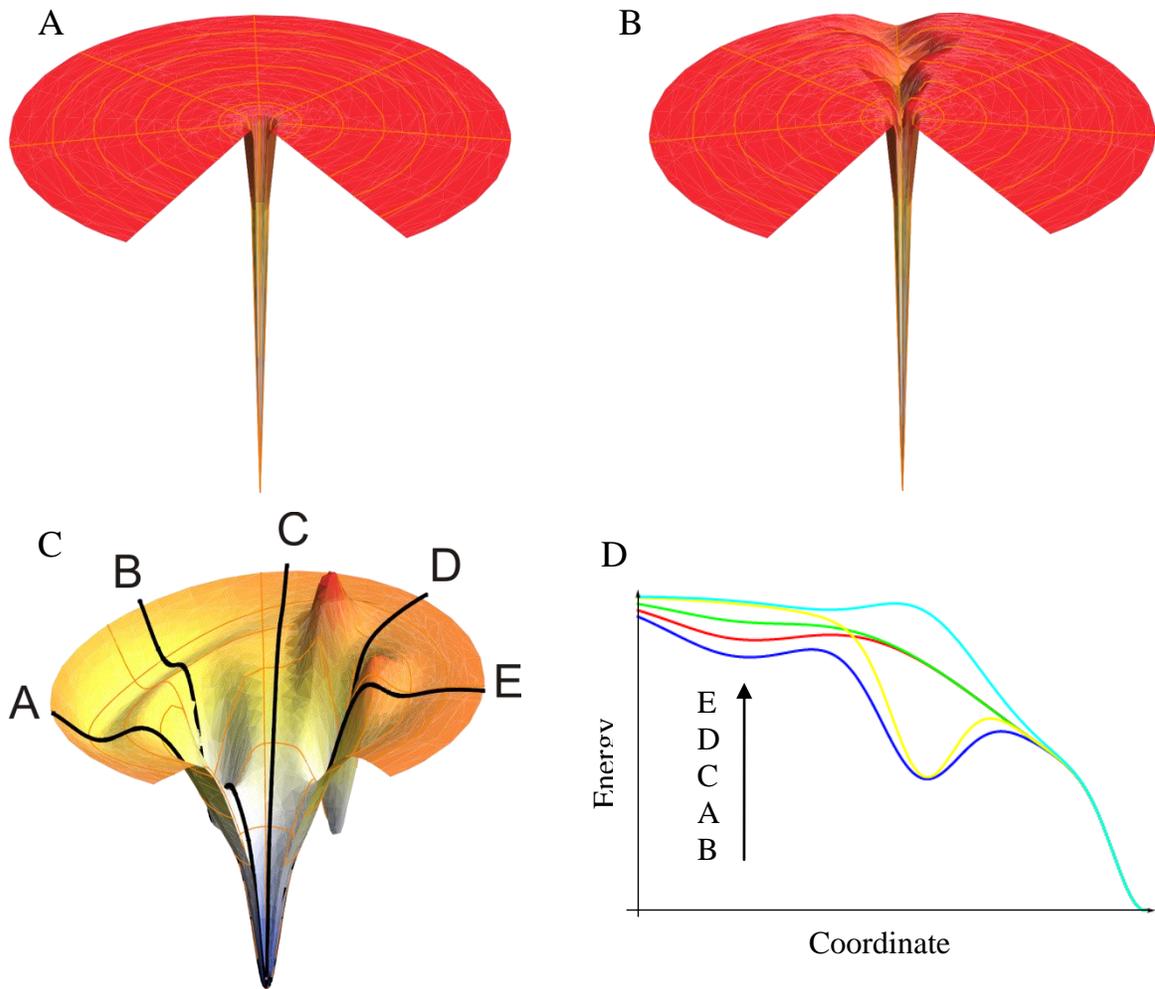


Figure 1. 9. Folding funnels

(A) Levinthal paradox, (B) single folding pathway, (C) rugged energy landscape and multiple folding pathways, (D) slices of folding pathways in (C)

The energy landscape theory constructs the energy surface in cylindrical coordinates instead of the normal Cartesian coordinates. The vertical axis represents the "internal free energy", which includes hydrophobic and solvation free energies, H-bonding free energy, ion pair free energy, etc. except conformational entropy. The conformational entropy, which is the conformational degeneracy with a certain free energy, is represented by the size of the surface at certain height. The global minimum is a deep well on the z-axis, representing the single well-defined native state conformation. As the free energy increases, the number of conformations

grows dramatically, so does the area of energy surface. The whole energy surface looks like a funnel.

The view of folding funnel has two significant impacts on understanding the protein folding. (1) Unfolded conformations at the funnel edge spontaneously proceeds towards the bottom of the deep well, which is the native state. (2) The unfolded conformations can undergo different routes towards the native state, in other words, a protein can fold in multiple pathways. The folding funnel is an ensemble of the unfolded state and folding pathways.

Figure 1. 9 shows the folding funnels in three cases. (1) The energy landscape of Levinthal paradox is shown in Figure 1. 9A. Except the native state conformation, all conformations, which are energetically degenerate, sit on a broad and flat surface. When the unfolded protein searches the native state, there are so many configurations it needs to try that it is kinetically impossible to fall in the hole. (2) The folding pathway view is shown in Figure 1. 9B. A trough connecting the unfolded conformations to the native state can guide the protein folding. It may have valleys (intermediates) and hills (transition states) on the trough or off the trough (off-pathway intermediates). (3) The energy landscape of protein is shown in Figure 1. 9C. The energy landscape is extremely rugged with hills and valleys. The local minima can be deep enough to kinetically trap proteins in misfolded conformations, as shown in pathway B and D in Figure 1. 9D.

1.4 BIBLIOGRAPHY

- (1) Selkoe, D. J. *Nature* **2003**, 426, 900.

- (2) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205.
- (3) Lovell, S. C.; Davis, I. W.; Adrendall, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 437.
- (4) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Interscience, 1969.
- (5) Hol, W. G. J. *Adv. Biophys.* **1985**, *19*, 133.
- (6) Hol, W. G. J. *Prog. Biophys. Mol. Bio.* **1985**, *45*, 149.
- (7) Chou, P. Y.; Fasman, G. D. *J. Mol. Biol.* **1977**, *115*, 135.
- (8) Rose, G. D.; Gierasch, L. M.; Smith, J. A. In *Advances in Protein Chemistry*; Anfinsen, C. B., Edsall, J. T., Richards, F. M., Eds.; Academic Press, 1985; Vol. 37.
- (9) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 2207.
- (10) Panasik, N.; Fleming, P. J.; Rose, G. D. *Protein Sci.* **2005**, *14*, 2910.
- (11) Shi, Z. S.; Chen, K.; Liu, Z. G.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877.
- (12) Shi, Z. S.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190.
- (13) Tompa, P. *Trends Biochem.Sci.* **2002**, *27*, 527.
- (14) Uversky, V. N. *Eur. J. Biochem.* **2002**, *269*, 2.
- (15) Uversky, V. N.; Dunker, A. K. *Biochim. Biophys. Acta. - Protein Proteomics* **2010**, *1804*, 1231.
- (16) Dyson, H. J.; Wright, P. E. *Curr. Opin. Struct. Biol.* **2002**, *12*, 54.
- (17) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.
- (18) Dobson, C. M. *Trends Biochem.Sci.* **1999**, *24*, 329.
- (19) Koo, E. H.; Lansbury, P. T.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9989.
- (20) Bucciardini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J. S.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. *Nature* **2002**, *416*, 507.
- (21) Lashuel, H. A.; Hartley, D.; Petre, B. M.; Walz, T.; Lansbury, P. T. *Nature* **2002**, *418*, 291.
- (22) Stefani, M.; Dobson, C. M. *J. Mol. Med.* **2003**, *81*, 678.
- (23) Baskakov, I. V.; Sipe, J. D. *Amyloid Proteins: The Beta Sheet Conformation and Disease*, 2005.
- (24) Glabe, C. G. *Neurobiol. Aging* **2006**, *27*, 570.
- (25) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333.
- (26) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215.
- (27) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.
- (28) Huyghuesdespointes, B. M. P.; Klingler, T. M.; Baldwin, R. L. *Biochemistry* **1995**, *34*, 13267.
- (29) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1983**, *168*, 867.
- (30) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8898.
- (31) Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1990**, *216*, 1031.
- (32) Lyu, P. C. C.; Gans, P. J.; Kallenbach, N. R. *J. Mol. Biol.* **1992**, *223*, 343.
- (33) Stellwagen, E.; Park, S. H.; Shalongo, W.; Jain, A. *Biopolymers* **1992**, *32*, 1193.
- (34) Huyghuesdespointes, B. M. P.; Scholtz, J. M.; Baldwin, R. L. *Protein Sci.* **1993**, *2*, 80.
- (35) Park, S. H.; Shalongo, W.; Stellwagen, E. *Biochemistry* **1993**, *32*, 12901.
- (36) Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 9668.
- (37) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211.

- (38) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol.* **1995**, *254*, 761.
- (39) HuyghuesDespointes, B. M. P.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 1965.
- (40) Mayne, L.; Englander, S. W.; Qiu, R.; Yang, J. X.; Gong, Y. X.; Spek, E. J.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1998**, *120*, 10643.
- (41) Sindelar, C. V.; Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1998**, *7*, 1898.
- (42) Smith, J. S.; Scholtz, J. M. *Biochemistry* **1998**, *37*, 33.
- (43) Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293*, 1241.
- (44) Olson, C. A.; Spek, E. J.; Shi, Z. S.; Vologodskii, A.; Kallenbach, N. R. *Proteins: Struct., Funct., Genet.* **2001**, *44*, 123.
- (45) Shi, Z. S.; Olson, C. A.; Bell, A. J.; Kallenbach, N. R. *Biopolymers* **2001**, *60*, 366.
- (46) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2782.
- (47) Ghosh, T.; Garde, S.; Garcia, A. E. *Biophys. J.* **2003**, *85*, 3187.
- (48) Errington, N.; Doig, A. J. *Biochemistry* **2005**, *44*, 7553.
- (49) Iqbalsyah, T. M.; Doig, A. J. *Biochemistry* **2005**, *44*, 10449.
- (50) FernandezRecio, J.; Vazquez, A.; Civera, C.; Sevilla, P.; Sancho, J. *J. Mol. Biol.* **1997**, *267*, 184.
- (51) Fernandez-Recio, J.; Romero, A.; Sancho, J. *J. Mol. Biol.* **1999**, *290*, 319.
- (52) Shi, Z. S.; Olson, C. A.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2002**, *124*, 3284.
- (53) Iqbalsyah, T. M.; Doig, A. J. *J. Am. Chem. Soc.* **2005**, *127*, 5002.
- (54) Marcus, Y. *Chem. Soc. Rev.* **1993**, *22*, 409.
- (55) Nelson, J. W.; Kallenbach, N. R. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 211.
- (56) Luo, P. Z.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 8413.
- (57) Bykov, S., personal communication.
- (58) Richardson, J.; Richardson, D. *Science* **1988**, *240*, 1648.
- (59) Doig, A. J.; Stapley, B. J.; Macarthur, M. W.; Thornton, J. M. *Protein Sci.* **1997**, *6*, 147.
- (60) Armstrong, K. M.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 11337.
- (61) Aurora, R.; Rose, G. D. *Protein Sci.* **1998**, *7*, 21.
- (62) Meyer, E. E.; Rosenberg, K. J.; Israelachvili, J. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15739.
- (63) Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 8069.
- (64) Becktel, W. J.; Schellman, J. A. *Biopolymers* **1987**, *26*, 1859.
- (65) Rose, G. D.; Wolfenden, R. *Annu. Rev. Biophys. Biomolec. Struct.* **1993**, *22*, 381.
- (66) Lockhart, D. J.; Kim, P. S. *Science* **1993**, *260*, 198.
- (67) Takahashi, S.; Kim, E.-H.; Hibino, T.; Ooi, T. *Biopolymers* **1989**, *28*, 995.
- (68) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987**, *326*, 563.
- (69) Sali, D.; Bycroft, M.; Fresht, A. R. *Nature* **1988**, *335*, 740.
- (70) Fairman, R.; Shoemaker, K. R.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Proteins: Struct., Funct., Genet.* **1989**, *5*, 1.
- (71) Lockhart, D. J.; Kim, P. S. *Science* **1992**, *257*, 947.
- (72) Forood, B.; Feliciano, E. J.; Nambiar, K. P. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 838.
- (73) Doig, A. J.; Baldwin, R. L. *Protein Sci.* **1995**, *4*, 1325.
- (74) Anfinsen, C. B. *Biochem. J.* **1972**, *128*, 737.
- (75) Anfinsen, C. B. *Science* **1973**, *181*, 223.
- (76) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10.
- (77) Karplus, M. *Fold. Des.* **1997**, *2*, S69.

- (78) Levinthal, C. *J. Chim. Phys. Phys.-Chim. Biol.* **1968**, 65, 44.
- (79) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1982**, 51, 459.
- (80) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1990**, 59, 631.
- (81) Creighton, T. E. *Biochemistry Journal* **1990**, 270, 1.
- (82) Baldwin, R. L. *Annu. Rev. Biophys.* **2008**, 37, 1.
- (83) Baldwin, R. L.; Rose, G. D. *Trends Biochem.Sci.* **1999**, 24, 26.
- (84) Baldwin, R. L.; Rose, G. D. *Trends Biochem.Sci.* **1999**, 24, 77.
- (85) Nolting, B.; Andert, K. *Proteins: Struct., Funct., Genet.* **2000**, 41, 288.
- (86) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. *Annu. Rev. Biophys.* **2008**, 37, 289.
- (87) Ivarsson, Y.; Travaglini-Allocatelli, C.; Brunori, M.; Gianni, S. *Eur. Biophys. J. Biophys. Lett.* **2008**, 37, 721.
- (88) Karplus, M.; Weaver, D. L. *Protein Sci.* **1994**, 3, 650.
- (89) Dobson, C. M.; Sali, A.; Karplus, M. *Angew. Chem. Int. Ed.* **1998**, 37, 868.
- (90) Murphy, K. P.; Privalov, P. L.; Gill, S. J. *Science* **1990**, 247, 559.
- (91) Fersht, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, 92, 10869.
- (92) Fersht, A. R. *Curr. Opin. Struct. Biol.* **1997**, 7, 3.
- (93) Fersht, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 1525.
- (94) Onuchic, J. N.; LutheySchulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, 48, 545.
- (95) Chan, H. S.; Dill, K. A. *Proteins: Struct., Funct., Genet.* **1998**, 30, 2.

2.0 RAMAN SPECTROSCOPY

Raman scattering is an inelastic light scattering phenomenon occurring when a photon impinges on a molecule. The incident electromagnetic field and the molecule exchange a quantum of energy. The difference of energy for incident and scattered photons thus provide information about the molecular quantum energy levels. This feature of Raman scattering has been widely exploited to study molecules and materials of interest.

2.1 RAMAN SCATTERING POLARIZABILITY

2.1.1 Classical theory

Classical theory¹ views a molecule as a collection of charges. An incident electric field would interact with the electron cloud and hence induce a dipole moment. Since the size of the molecules is usually much smaller than the wavelength of incident light, electric field that the molecule experiences is constant. Therefore, the dipole moment could be expanded as a Taylor series:

$$\mathbf{P} = \mathbf{P}_0 + \frac{\partial \mathbf{P}}{\partial \mathbf{E}} \mathbf{E} + \frac{1}{2} \frac{\partial^2 \mathbf{P}}{\partial \mathbf{E}^2} \mathbf{E}^2 + \dots \quad (2.1)$$

where \mathbf{P} is the dipole moment, and \mathbf{E} is the electric field. The first term is the permanent dipole moment. The second term is the induced dipole moment that governs the scattering. The induced dipole moment can be expressed as:

$$\mathbf{P}_{\text{ind}} = \alpha \mathbf{E} \quad (2.2)$$

where $\alpha = \frac{\partial \mathbf{P}}{\partial \mathbf{E}}$ is the polarizability tensor of the molecule, describing the electric field-induced displacement of the charges. The polarizability tensor α is a function of vibrational coordinate associated with every vibrational normal mode. We expand the expression of α in a Taylor series with respect to vibrational normal coordinate as

$$\alpha = \alpha_0 + \sum_k \frac{\partial \alpha}{\partial Q_k} Q_k + \frac{1}{2} \sum_{k,l} \frac{\partial^2 \alpha}{\partial Q_k \partial Q_l} Q_k Q_l \quad (2.3)$$

where Q_k is displacement from equilibrium position along the k th vibration normal mode coordinate. If we truncate after the first order term and refer to the derivative as α'_k , we may write eq. (2.3) as

$$\alpha = \alpha_0 + \sum_k \alpha'_k Q_k \quad (2.4)$$

We now consider Q_k . Assuming simple harmonic vibration, the time dependence of Q_k is given by

$$Q_k = Q_{k0} \cos(\omega_k t + \xi_k) \quad (2.5)$$

where Q_{k0} is the maximum displacement, $\omega_k = \sqrt{\frac{k_k}{\mu_k}}$ is the angular frequency of the vibration and ξ_k is a phase factor. Substituting eq. (2.5) into eq. (2.4) yields

$$\alpha = \alpha_0 + \sum_k \alpha'_k Q_{k0} \cos(\omega_k t + \xi_k) \quad (2.6)$$

The incident electric field is oscillating with a frequency ω_0 , given by eq. (2.6)

$$\mathbf{E} = \mathbf{E}_0 \cos \omega_0 t \quad (2.7)$$

Substituting eq. (2.6) and eq. (2.7) into eq. (2.2), and using the trigonometric identity, we then obtain:

$$\begin{aligned} \mathbf{P}_{\text{ind}} = & \alpha_0 : \mathbf{E}_0 \cos \omega_0 t + \frac{1}{2} \sum_k Q_{k0} \alpha'_k : \mathbf{E}_0 \cos[(\omega_0 - \omega_k)t + \varrho_k] \\ & + \frac{1}{2} \sum_k Q_{k0} \alpha'_k : \mathbf{E}_0 \cos[(\omega_0 + \omega_k)t + \varrho_k] \end{aligned} \quad (2.8)$$

It is clearly seen from eq. (2.8) that the induced moment resulted from the interaction of light with molecule is composed of three terms. The first term is an oscillating dipole at the same frequency as the incident electromagnetic field, therefore the scattered photon radiated from this dipole has the same frequency as the incident photon. This is known as Rayleigh scattering. The second term is the summation of oscillating dipoles over all vibrational normal modes. The frequencies of these dipoles are downshifted by vibrational normal mode frequencies. As a consequence, the scattered photons radiated from these dipoles are a series of frequencies that are downshifted from the incident frequency. This is known as Stokes Raman scattering. The third term is similar to the second term, except it results from a frequency upshift. This term describes the so called anti-Stokes Raman scattering.

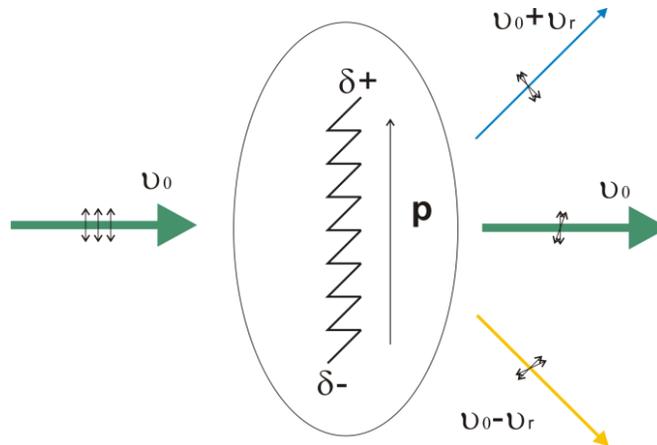


Figure 2. 1. Classical view of light scattering by a molecule

2.1.2 Vibronic coupling theory

While the classical theory provides a framework to treat the Raman scattering, it cannot provide insight into the molecular origin of Raman scattering. To get insight into the factors that dictate the Raman scattering, quantum mechanics is applied to treat the molecule, while the electromagnetic field is treated classically for the sake of simplification¹.

Analogous to the classical treatment, we expand the dipole moment

$$\mathbf{P}_{fi} = \mathbf{P}_{fi}^{(0)} + \mathbf{P}_{fi}^{(1)} + \mathbf{P}_{fi}^{(2)} + \dots \quad (2.9)$$

where subscript fi refers to the transition from initial state $|i\rangle$ to final state $|f\rangle$; $\mathbf{P}_{fi}^{(0)}$ is the permanent dipole moment; $\mathbf{P}_{fi}^{(1)}$, which is linear with \mathbf{E} , is the induced dipole moment. The total transition electric dipole is given by

$$\mathbf{P}_{fi} = \langle \Psi_f | \hat{p} | \Psi_i \rangle \quad (2.10)$$

where Ψ_i and Ψ_f are the time-dependent wave functions of the initial and final states of the molecules. Perturbation theory allows Ψ_i and Ψ_f to be given by:

$$\begin{aligned} \Psi_i &= \Psi_i^{(0)} + \Psi_i^{(1)} + \Psi_i^{(2)} + \dots \\ \Psi_f &= \Psi_f^{(0)} + \Psi_f^{(1)} + \Psi_f^{(2)} + \dots \end{aligned} \quad (2.11)$$

where the superscript is the order of the perturbation. Therefore, plugging eq. (2.11) into eq. (2.10) and collecting terms, we then find the expression of induced dipole moment as

$$\mathbf{P}_{fi}^{(1)} = \langle \Psi_f^{(0)} | \hat{p} | \Psi_i^{(1)} \rangle + \langle \Psi_f^{(1)} | \hat{p} | \Psi_i^{(0)} \rangle \quad (2.12)$$

We can further express $\Psi_i^{(1)}$ and $\Psi_f^{(1)}$ as linear combinations of a complete set of unperturbed time-independent wave functions $|r\rangle$

$$\begin{aligned}
|\Psi_i^{(1)}\rangle &= \sum_r \beta_{ir}^{(1)} e^{-i(\omega_r - i\Gamma_r)t} |r\rangle \\
\langle \Psi_f^{(1)}| &= \sum_r \beta_{fr}^{(1)*} e^{i(\omega_r + i\Gamma_r)t} \langle r|
\end{aligned} \tag{2.13}$$

$\omega_r = \frac{E_r}{\hbar}$ is the angular frequency and Γ_r is a phenomenological damping factor of state $|r\rangle$ and equals the reciprocal of state $|r\rangle$ lifetime. Plugging eq. (2.13) in time-dependent Schrödinger equation, the coefficients are found as²⁻³

$$\begin{aligned}
\beta_{ir}^{(1)} &= \frac{-i}{2\hbar} \sum_n \int_0^t \langle r | -\hat{p}\mathbf{E}_0 (e^{-i\omega_0 t'} + e^{i\omega_0 t'}) |n\rangle \beta_n(0) e^{i(\omega_{rn} - i\Gamma_{rn})t'} dt' \\
\beta_{fr}^{(1)*} &= \frac{i}{2\hbar} \sum_n \int_0^t \langle n | -\hat{p}\mathbf{E}_0 (e^{-i\omega_0 t'} + e^{i\omega_0 t'}) |r\rangle \beta_n^*(0) e^{-i(\omega_{rn} + i\Gamma_{rn})t'} dt'
\end{aligned} \tag{2.14}$$

where $-\hat{p}\mathbf{E}_0 (e^{-i\omega_0 t} + e^{i\omega_0 t})$ is the time-dependent perturbation Hamiltonian, $\omega_{rn} = \omega_r - \omega_n$ and $\Gamma_{rn} = \Gamma_r - \Gamma_n$. At $t=0$, the system is at state $|i\rangle$, therefore $\beta_n(0) = \delta_{ni}$. Also, the states $|i\rangle$ and $|f\rangle$ have infinite lifetime, $\Gamma_{i,f} = 0$. We can remove the summation and integrate the right-hand side of eq. (2.14),

$$\begin{aligned}
\beta_{ir}^{(1)} &= \frac{i}{2\hbar} \langle r | \hat{p}\mathbf{E}_0 | i \rangle \left\{ \frac{e^{i(\omega_{ri} - \omega_0 - i\Gamma_r)t} - 1}{i(\omega_{ri} - \omega_0 - i\Gamma_r)} + \frac{e^{i(\omega_{ri} + \omega_0 - i\Gamma_r)t} - 1}{i(\omega_{ri} + \omega_0 - i\Gamma_r)} \right\} \\
\beta_{fr}^{(1)*} &= \frac{-i}{2\hbar} \langle f | \hat{p}\mathbf{E}_0 | r \rangle \left\{ \frac{e^{-i(\omega_{rf} + \omega_0 + i\Gamma_r)t} - 1}{-i(\omega_{rf} + \omega_0 + i\Gamma_r)} + \frac{e^{-i(\omega_{rf} - \omega_0 + i\Gamma_r)t} - 1}{-i(\omega_{rf} - \omega_0 + i\Gamma_r)} \right\}
\end{aligned} \tag{2.15}$$

Substituting eq. (2.15) into eq. (2.13) and evaluating each term in eq. (2.12), we get

$$\begin{aligned}
\langle \Psi_f^{(0)} | \hat{p} | \Psi_i^{(1)} \rangle &= \frac{1}{2\hbar} \sum_r \langle f | \hat{p} | r \rangle \langle r | \hat{p} | i \rangle \mathbf{E}_0 \left\{ \frac{e^{-i(\omega_0 - \omega_{fi})t} - e^{-i(\omega_{rf} - i\Gamma_r)t}}{\omega_{ri} - \omega_0 - i\Gamma_r} \right. \\
&\quad \left. + \frac{e^{i(\omega_0 + \omega_{fi})t} - e^{-i(\omega_{rf} - i\Gamma_r)t}}{\omega_{ri} + \omega_0 - i\Gamma_r} \right\}
\end{aligned} \tag{2.16}$$

$$\begin{aligned} \langle \Psi_f^{(1)} | \hat{p} | \Psi_i^{(0)} \rangle = & \frac{1}{2\hbar} \sum_r \langle f | \hat{p} | r \rangle \langle r | \hat{p} | i \rangle \mathbf{E}_0 \left\{ \frac{e^{-i(\omega_0 - \omega_{fi})t} - e^{i(\omega_{ri} + i\Gamma_r)t}}{\omega_{rf} + \omega_0 + i\Gamma_r} \right. \\ & \left. + \frac{e^{i(\omega_0 + \omega_{fi})t} - e^{i(\omega_{ri} - i\Gamma_r)t}}{\omega_{ri} - \omega_0 + i\Gamma_r} \right\} \end{aligned}$$

It is noticed that the electric field is modulated into four frequencies, $\omega_0 - \omega_{fi}$, $\omega_0 + \omega_{fi}$, $\omega_{ri} + i\Gamma_r$ and $\omega_{rf} - i\Gamma_r$. The four terms with frequencies involving ω_r rapidly vanish due to the summation over all $|r\rangle$ states. The two terms with the frequency of $\omega_0 - \omega_{fi}$ account for Raman scattering, of which we are most concerned.

Therefore, for Raman scattering, we have

$$\left(\mathbf{P}_\rho^{(1)} \right)_{fi} = \frac{1}{2\hbar} \sum_r \left\{ \frac{\langle f | \hat{p}_\rho | r \rangle \langle r | \hat{p}_\sigma | i \rangle}{\omega_{ri} - \omega_0 - i\Gamma_r} \mathbf{E}_{\sigma 0} e^{-i\omega_s t} + \frac{\langle f | \hat{p}_\sigma | r \rangle \langle r | \hat{p}_\rho | i \rangle}{\omega_{rf} + \omega_0 + i\Gamma_r} \mathbf{E}_{\sigma 0} e^{i\omega_s t} \right\} \quad (2.17)$$

where $\omega_s = \omega_0 - \omega_{fi}$ is the frequency of the scattered photon, subscripts ρ and σ represent the coordinates. Taking the complex conjugate into account, it is straightforward to write the Raman scattering polarizability tensor

$$\left(\alpha_{\rho\sigma} \right)_{fi} = \frac{1}{\hbar} \sum_r \left\{ \frac{\langle f | \hat{p}_\rho | r \rangle \langle r | \hat{p}_\sigma | i \rangle}{\omega_{ri} - \omega_0 - i\Gamma_r} + \frac{\langle f | \hat{p}_\sigma | r \rangle \langle r | \hat{p}_\rho | i \rangle}{\omega_{rf} + \omega_0 + i\Gamma_r} \right\} \quad (2.18)$$

This Kramers-Heisenberg-Dirac (KHD) expression reveals the microscopic picture of the light scattering process. As the incident electromagnetic field interacts with the molecule, the molecule undergoes a transition from initial state to a virtual state, and immediately undergoes another transition to final state from the virtual state, accompanying with the "emission" of a photon. The virtual state is actually a linear combination of all eigenstates of the system. The Rayleigh scattering is a process in which the final and initial states involved are identical. The Stokes Raman scattering is a process in which the final state is in higher energy level than initial state, while in the anti-Stokes Raman scattering, the energy level of final state is lower.

Γ_r is usually small compared with ω_{ri} and ω_0 . As the excitation frequency ω_0 approaches electronic excited state energy, the denominator of the first term in eq. (2.18) approaches Γ_r . This term thus dominates and is called the “resonance term”. The second term is called the “non-resonance term”.

2.1.3 Time-dependent theory

Time-dependent theory views Raman scattering process dynamically as the wave packet propagating in phase space (Figure 2. 2). The wave packet $\chi_{1i}^{(0)}$, which is a superposition of wave functions, initially propagates on an electronic state i (usually ground state) potential energy surface. After some time, at t_1 , the wave packet is excited to an intermediate potential surface r and propagates along this surface. At t_2 , another photon arrives and scatters the wave packet to the final potential surface f .⁴ Therefore, assuming the wave packet at the initial or final potential surface has an infinite lifetime, the second order amplitude of the wave packet at time t is

$$\begin{aligned}
\chi_{1i}^{(2)}(t) = & \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_2 \int_0^{t_2} dt_1 e^{-i\hat{H}_1(t-t_2)/\hbar} [-\hat{p}_{fr} \mathbf{E}_2 e^{i\omega_s t_2}] \\
& \times e^{-i(\hat{H}_r/\hbar - i\Gamma_r)(t_2-t_1)} [-\hat{p}_{ri} \mathbf{E}_1 e^{-i\omega_0 t_1}] e^{-i\hat{H}_1 t_1/\hbar} \chi_{1i}^{(0)} \\
& + \sum_r \int_0^t dt_2 \int_0^{t_2} dt_1 e^{-i\hat{H}_1(t-t_2)/\hbar} [-\hat{p}_{fr} \mathbf{E}_1 e^{-i\omega_0 t_2}] \\
& \times e^{-i(\hat{H}_r/\hbar - i\Gamma_r)(t_2-t_1)} [-\hat{p}_{ri} \mathbf{E}_2 e^{i\omega_s t_1}] e^{-i\hat{H}_1 t_1/\hbar} \chi_{1i}^{(0)}
\end{aligned} \tag{2.19}$$

The second term in eq. (2.19) which describes the process that a scattered photon emits before the incident photon arrives is the non-resonance term and will be discarded hereafter.

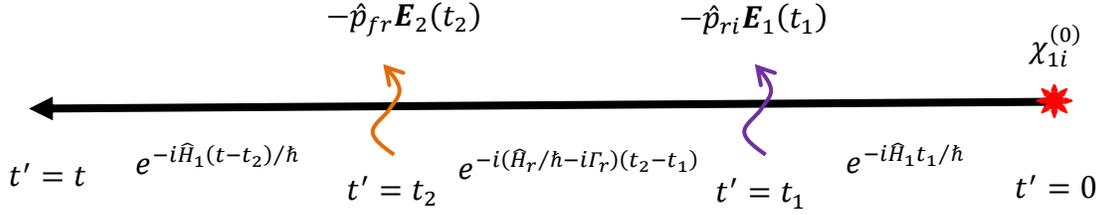


Figure 2. 2. Diagram showing two-photon process from time-dependent perspective

The amplitude for finding the molecule in state $\chi_{1f}^{(0)}$ through a Raman scattering process at time t is thus

$$\begin{aligned}
C_{1f \leftarrow 1i}^{(2)} &= \langle \chi_{1f}^{(0)}(t) | \chi_{1i}^{(2)}(t) \rangle \\
&= -\left(\frac{1}{i\hbar}\right)^2 \int_0^t dt_2 e^{i(\omega_{fi} + \omega_s - \omega_0)t_2} \mathbf{E}_1 \mathbf{E}_2 \\
&\quad \times \sum_r \int_0^{t_2} dt_1 e^{i(\omega_{1i} + \omega_0)t_1} \langle \chi_{1f}^{(0)} | \hat{p}_{fr} e^{-i(\hat{H}_r/\hbar - i\Gamma_r)(t_2-t_1)} \hat{p}_{ri} | \chi_{1i}^{(0)} \rangle
\end{aligned} \tag{2.20}$$

Defining

$$f_{fi}(\omega_0) = -\frac{1}{i\hbar} \sum_r \int_0^\infty dt_1 e^{i(\omega_{1i} + \omega_0)t_1} \langle \chi_{1f}^{(0)} | \hat{p}_{fr} e^{-i(\hat{H}_r/\hbar - i\Gamma_r)t_1} \hat{p}_{ri} | \chi_{1i}^{(0)} \rangle \tag{2.21}$$

Inserting $\sum |r\rangle\langle r|$ into the right-hand side of eq. (2.21),

$$f_{fi}(\omega_0) = -\frac{1}{i\hbar} \sum_r \int_0^\infty dt_1 e^{-i(\omega_{ri} - \omega_0 - i\Gamma_r)t_1} \langle \chi_{1f}^{(0)} | \hat{p}_{fr} | r \rangle \langle r | \hat{p}_{ri} | \chi_{1i}^{(0)} \rangle \tag{2.22}$$

It is immediately recognized that by performing the integration, eq. (2.22) is equivalent to the resonance term of the KHD expression of Raman scattering polarizability tensor in eq. (2.18) which is derived from vibronic coupling theory.

2.2 VIBRATIONAL RESONANCE RAMAN SCATTERING

2.2.1 Resonance Raman scattering

In normal Raman scattering (Figure 2. 3A), the energy of the incident photon is much lower than electronic excited state, the virtual state involves large contributions from the vibrational states on the electronic ground state and is lower than the first electronic excited state. The non-resonance term in eq. (2.18) has a magnitude comparable to the resonance term.

Time-dependent theory also describes normal Raman scattering. The time t_2-t_1 in which the scattering process takes place is so short that the energy is large enough for the wave packet to reach at least the first electronic excited state due to Heisenberg's uncertainty principle.⁵ The time that the wave packet propagates on the electronic excited state potential surface is extremely short and eq. (2.22) converges very rapidly.

As the energy of incident photon approaches an electronic excited state, the virtual state adopts a large contribution from that vibronic state and minor contributions from other states. This is the pre-resonance situation shown in Figure 2. 3B.

Figure 2. 3C shows the resonance Raman scattering. In this situation, the incident photon has sufficient energy to excite a molecule from the electronic ground state to an electronic excited state. Consequently, the virtual state is contributed mainly from the closely lying vibronic states of this electronic excited state; and the summation in eq. (2.18) can be approximately confined to these vibronic states. Moreover, the denominator of the resonance term is much smaller than the non-resonance term. It is appropriate to ignore the non-resonance term. The resultant equation is written as:

$$(\alpha_{\rho\sigma})_{fi} = \frac{1}{\hbar} \sum_r \frac{\langle f|\hat{p}_\rho|r\rangle\langle r|\hat{p}_\sigma|i\rangle}{\omega_{ri} - \omega_0 - i\Gamma_r} \quad (2.23)$$

Applying the Born-Oppenheimer approximation, we can separate the vibrational wave functions from the electronic wave functions. Assuming the initial state is state $|g0\rangle$, the final state is state $|g1\rangle$, and the states $|r\rangle$ are states $|ev\rangle$, eq. (2.23) is rewritten as:

$$(\alpha_{\rho\sigma})_{fi} = \frac{1}{\hbar} \frac{\langle g|\hat{p}_\rho|e\rangle\langle e|\hat{p}_\sigma|g\rangle}{\omega_{eg} - \omega_0 - i\Gamma_e} \sum_v \langle 1|v\rangle\langle v|0\rangle \quad (2.24)$$

The terms $\langle g|\hat{p}_\rho|e\rangle$ and $\langle e|\hat{p}_\sigma|g\rangle$ are the electronic transition dipole moment, therefore, eq. (2.24) predicts that the resonance Raman polarizability is proportional to the square of electronic transition dipole moment. The terms $\langle 1|v\rangle$ and $\langle v|0\rangle$ are known as Franck-Condon integrals, depending on the overlap between the initial vibrational state and the vibrational state in the electronic excited state and the overlap of the final vibrational state and the vibrational state in the electronic excited state.

Application of Born-Oppenheimer approximation on eq. (2.22) leads to⁶

$$(f_{\rho\sigma})_{fi} = -\frac{1}{i\hbar} \sum_v \langle 1|v\rangle\langle v|0\rangle \int_0^\infty dt e^{-i(\omega_{eg} - \omega_0 - i\Gamma_e)t} \langle g|\hat{p}_\rho|e\rangle\langle e|\hat{p}_\sigma|g\rangle \quad (2.25)$$

The Franck-Condon integral in eq. (2.24) and eq. (2.25) is the overlap of vibrational wave functions in electronic ground and excited states. For non-zero Franck-Condon integral, the electronic excited potential surface has to shift or differ from the ground state potential surface. Otherwise the vibrational wave functions belong to the same complete set of eigenfunctions and the integral vanishes as dictated by the orthogonality of the vibrational wave functions. The estimation of the Franck-Condon integral is not trivial. It requires the knowledge of the vibrational wave functions on the electronic excited state, which are usually difficult to obtain,

especially for polyatomic molecules. In the limit of no Duschinsky rotation, the multidimensional Franck-Condon integral can be separated and evaluated in one-dimension.⁶⁻⁷

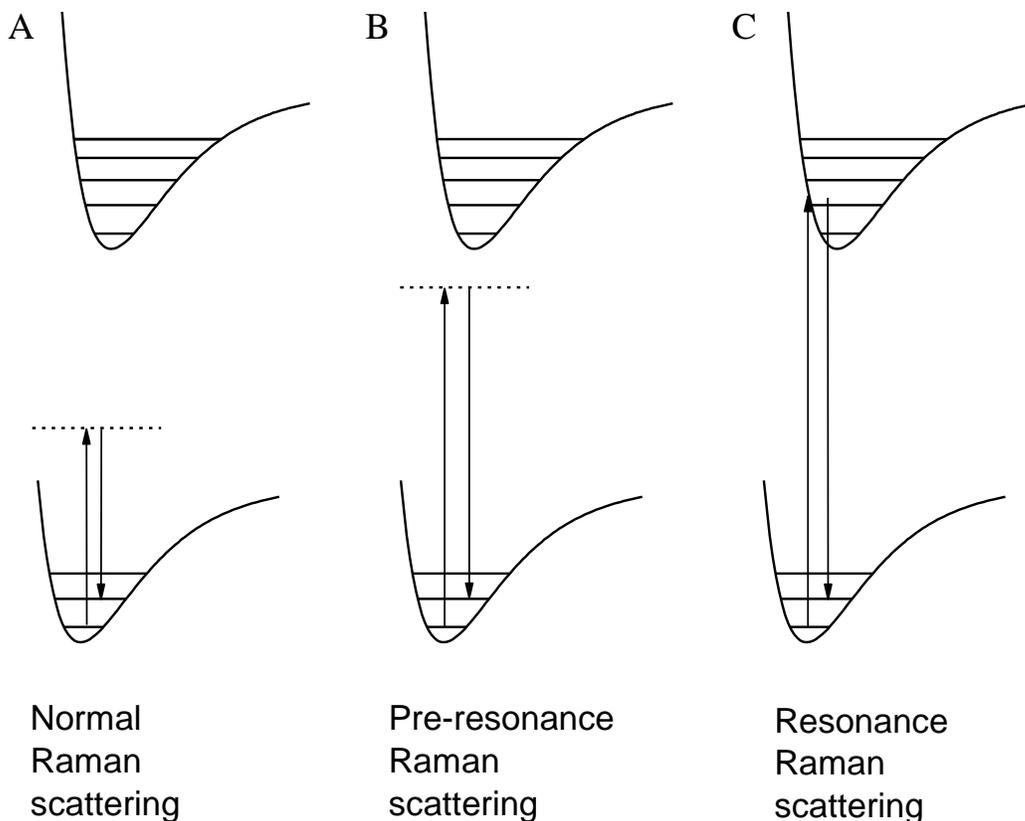


Figure 2. 3. Three types of Raman scattering

In time dependent theory, if it is assumed that a Gaussian wave packet propagates on a harmonic potential surface, under short time dynamics approximation,⁴ semiclassical treatment⁸⁻⁹ gives an expression of Raman scattering cross section¹⁰⁻¹¹

$$\sigma_k^R \sim \omega_0 (\omega_0 - \omega_k)^3 \frac{V_k^2}{\omega_k} \quad (2.26)$$

where ω_0 and ω_k are the angular frequencies of excitation and k th vibration, $V_k \equiv -\left. \frac{\partial E_e}{\partial Q_k} \right|_{Q_k=Q_{k0}}$ is the force of the first electronic excited state along k th mass weighted vibrational normal coordinate at the electronic ground state equilibrium geometry. Eq. (2.26) requires determining

the force of electronic excited potential surface that is easier to evaluate than the vibrational frequencies of electronic excited potential surface. Eq. (2.26) avoids the difficulty of estimating the Franck-Condon integrals in eq. (2.24) and eq. (2.25).

If the condition of no Duschinsky rotation or frequency changes on electronic excited state is imposed, eq. (2.26) can be rewritten as

$$\sigma_k^R \sim \omega_0 (\omega_0 - \omega_k)^3 \Delta_k^2 \omega_k^2 \quad (2.27)$$

where $\Delta_k = (Q_k - Q_{k0})/\sqrt{\omega_k}$ is the displacement of the electronic excited state equilibrium away from electronic ground state equilibrium along k th vibrational dimensionless normal coordinate. Eq. (2.27) provides a way to investigate the excited state equilibrium geometry.^{7,12}

2.2.2 Normal modes analysis

The geometry of a molecule is described by $3N$ Cartesian coordinates \mathbf{X} . The kinetic energy T and potential energy V of the molecule are respectively¹³⁻¹⁴

$$T = \frac{1}{2} \dot{\mathbf{X}}^T \mathbf{M} \dot{\mathbf{X}} \quad (2.28)$$

$$V = \frac{1}{2} (\mathbf{X} - \mathbf{X}_0)^T \mathbf{F} (\mathbf{X} - \mathbf{X}_0)$$

where \mathbf{M} is a $3N \times 3N$ matrix containing masses of atoms and \mathbf{F} is the Cartesian Hessian matrix, \mathbf{X}_0 is the equilibrium position.

Excluding translational and rotational motions, a molecule has $3N-6$ ($3N-5$ for linear molecule) vibrations. Intuitively, the vibrations are described by the changes of bond length, bond angle and dihedral angle, which constitute internal coordinates \mathbf{R} . Then eq. (2.28) becomes

$$T = \frac{1}{2} \dot{\mathbf{R}}^T \mathbf{g} \dot{\mathbf{R}} \quad (2.29)$$

$$V = \frac{1}{2} (\mathbf{R} - \mathbf{R}_0)^T \mathbf{f} (\mathbf{R} - \mathbf{R}_0)$$

where $\mathbf{g} = \mathbf{B}^T \mathbf{M} \mathbf{B}$ is a $(3N-6) \times (3N-6)$ matrix containing reduced masses for each internal coordinate and matrix $\mathbf{f} = \mathbf{B}^T \mathbf{F} \mathbf{B}$ is the internal coordinate Hessian. The transformation matrix \mathbf{B} is a $(3N-6) \times 3N$ matrix and can be obtained from taking the derivative of internal coordinates \mathbf{R} with respect to Cartesian coordinates \mathbf{X} .¹⁴⁻¹⁵ Therefore,

$$\mathbf{R} = \mathbf{B} \mathbf{X} \quad (2.30)$$

The $3N-6$ vibrations of a molecule form a vector in a $3N-6$ dimensional Euclidean space. Internal coordinates \mathbf{R} is the projection of this vector onto a basis, which is usually not orthogonal. However, it is possible to find an orthogonal basis, each basis vector of which is a normal vibration where all atoms move in phase with a particular frequency. The projection of the vibration vector on this orthogonal basis is called normal coordinates \mathbf{Q} . Eq. (2.29) becomes

$$T = \frac{1}{2} \dot{\mathbf{Q}}^T \mathbf{G} \dot{\mathbf{Q}} \quad (2.31)$$

$$V = \frac{1}{2} (\mathbf{Q} - \mathbf{Q}_0)^T \mathbf{\Lambda} (\mathbf{Q} - \mathbf{Q}_0)$$

where $\mathbf{G} = \mathbf{L}^T \mathbf{g} \mathbf{L}$ is a $(3N-6) \times (3N-6)$ matrix containing reduced masses for each normal coordinate. The normal coordinates Hessian matrix $\mathbf{\Lambda} = \mathbf{L}^T \mathbf{f} \mathbf{L}$ is a diagonal matrix with corresponding force constants of the normal vibrations as diagonal elements. The off-diagonal elements vanish by taking second derivative with respect to the orthogonal normal coordinates.

The transformation matrix \mathbf{L} connects internal coordinates \mathbf{R} and normal coordinates \mathbf{Q} .

$$\mathbf{R} = \mathbf{L} \mathbf{Q} \quad (2.32)$$

However, \mathbf{L} cannot be straightforwardly obtained by taking the derivatives of \mathbf{R} with respect to \mathbf{Q} , because \mathbf{Q} is usually unknown. \mathbf{L} can be obtained by solving the generalized eigenvalue problem¹⁴

$$\mathbf{fL} = \mathbf{gLA} \quad (2.33)$$

The \mathbf{L} matrix that satisfies eq. (2.33) transforms \mathbf{g} into a unit matrix. The diagonal force constants \mathbf{A} is obtained through the solving of eq. (2.33).

The potential energy distributions (PED) are defined as¹⁴

$$c_{jk} = \frac{\mathbf{f}_{jj}\mathbf{L}_{jk}^2}{\sum_i \mathbf{f}_{ii}\mathbf{L}_{ik}^2} \quad (2.34)$$

Eq. (2.34) describes the relative potential energy contribution of j th internal coordinate to the k th normal coordinate. PED provides a quantitative way to understanding the nature of a normal vibrational mode. PED are also very useful in understanding and predicting how a particular interaction, for example, H-bonding, affects the band frequency that we observe in a vibrational spectrum. For instance, vibrations involving C=O stretching downshifts if the carbonyl group is H-bonded; if an N-H group is H-bonded, the vibrations involving N-H stretching will downshift, while vibrations involving N-H bending upshifts.

Based on the methods described above, We wrote a computer program¹⁶ in matlab to perform normal mode analysis and calculate PED with Gaussian frequency calculation results as input. Figure 2. 4 shows the flow chart of this program. This program helps to generate internal coordinates and combines internal coordinates into symmetry adapted coordinates.¹⁷ It can also calculate the PED of isotope substituted molecule. With a known excited state geometry or potential energy gradients obtained from Gaussian calculation, this program can also calculate resonance Raman spectra by using eq. (2.26) or eq. (2.27).

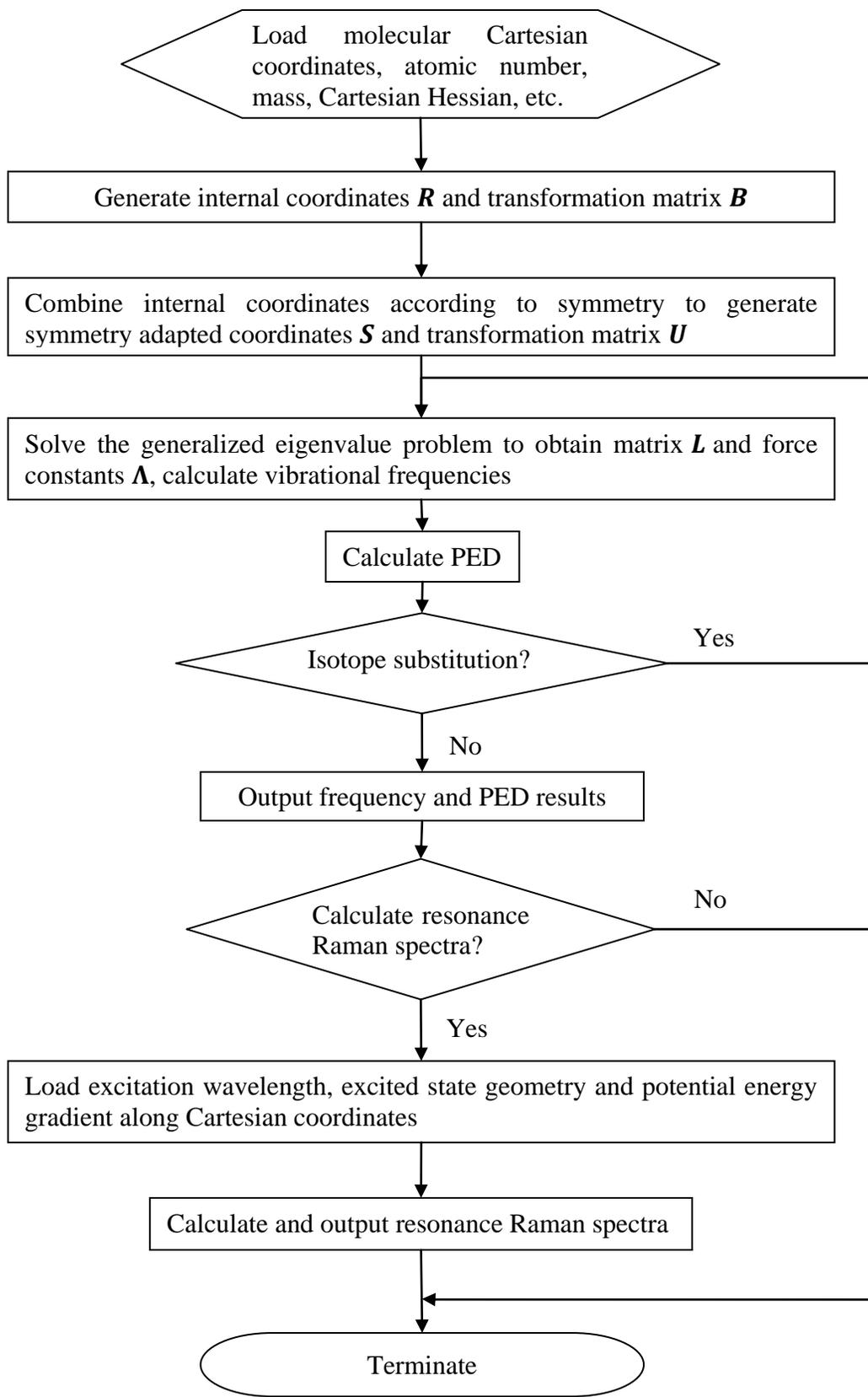


Figure 2. 4. Flow chart of the home-developed PED calculation program

2.2.3 Duschinsky rotation

In the limit of no Duschinsky rotation, the Franck-Condon integrals can be factored into one-dimension integrals and evaluated.^{6-7,18} Eq. (2.27) also assumes no Duschinsky rotation or frequency change occurs on electronic excited state. However, the electronic excited state potential surface is in general different from that of electronic ground state. The vibrational normal coordinates are thus different in general between ground state and electronic excited state. Duschinsky rotation describes a compositional change between the vibrational normal coordinates of the electronic excited and ground states. Figure 2. 5 shows two-dimensional ground and excited state potential surfaces. The excited state potential surface is shifted from the ground state. The excited state normal coordinates are rotated from those of the ground state, indicating a Duschinsky rotation. The frequencies differ as well.

With zero-order axis-switching approximation,¹⁹ The excited state normal coordinates \mathbf{Q}_{ex} thus can be connected to the ground state normal coordinates \mathbf{Q}_{gr} by the Duschinsky rotation matrix \mathbf{J} and a displacement vector \mathbf{K} .

$$\mathbf{Q}_{ex} = \mathbf{J}\mathbf{Q}_{gr} + \mathbf{K} \quad (2.35)$$

The calculation of \mathbf{K} is straightforward, however, the calculation of \mathbf{J} is much more complicated since it must involve rotational motions.²⁰⁻²² The translational and rotational axes differ between the electronic ground and excited states and are thus coupled with the vibrational coordinates. A method to calculate the Duschinsky rotation matrix \mathbf{J} has been proposed.^{19,23}

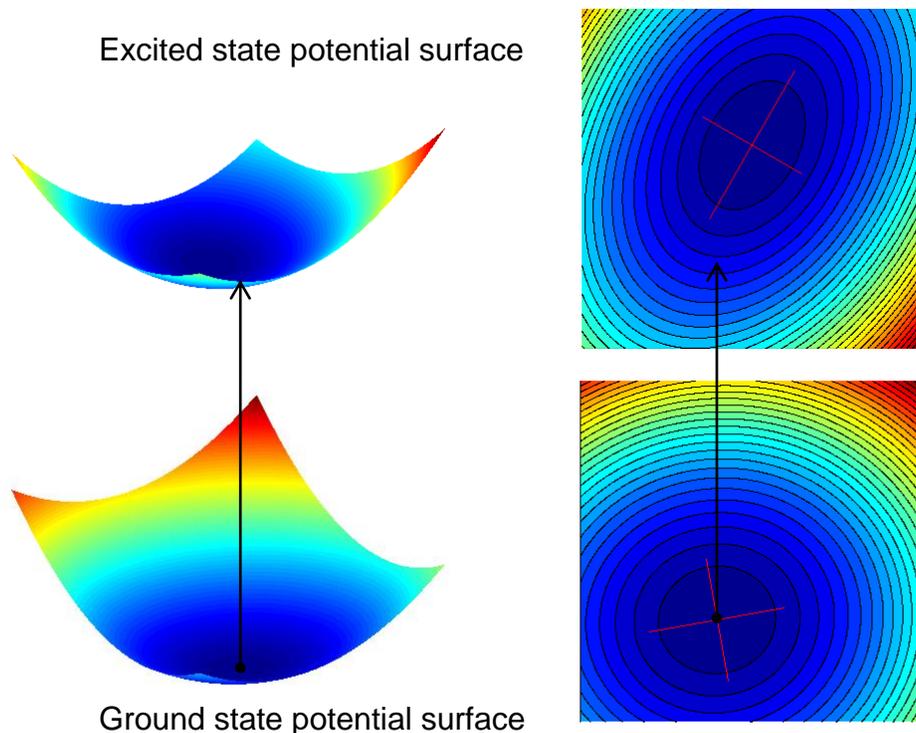


Figure 2. 5. Ground and excited state potential surfaces and vertical transition

The red lines indicate the directions of normal coordinates, the intersection points of the red lines are the equilibrium positions of the potentials. The black arrow indicates a vertical transition from the ground state equilibrium position to the excited state potential surface

Figure 2. 6 shows the Duschinsky rotation matrix of formaldehyde between the ground and the first excited state calculated by Gaussian 09.²⁴ The original Duschinsky rotation matrix is converted into angles between an electronic ground state normal mode and the electronic first excited state normal mode. It is clear that for formaldehyde, the corresponding normal modes in the ground and first electronic excited states are either parallel or antiparallel with an angle close to 0° or 180° , as the blocks along the diagonal are either in dark blue or dark red, substantially distinct from the off-diagonal blocks. For large molecules, in general, more significant Duschinsky rotations occur and the diagonal elements of the matrix are less distinct from off-diagonal elements.

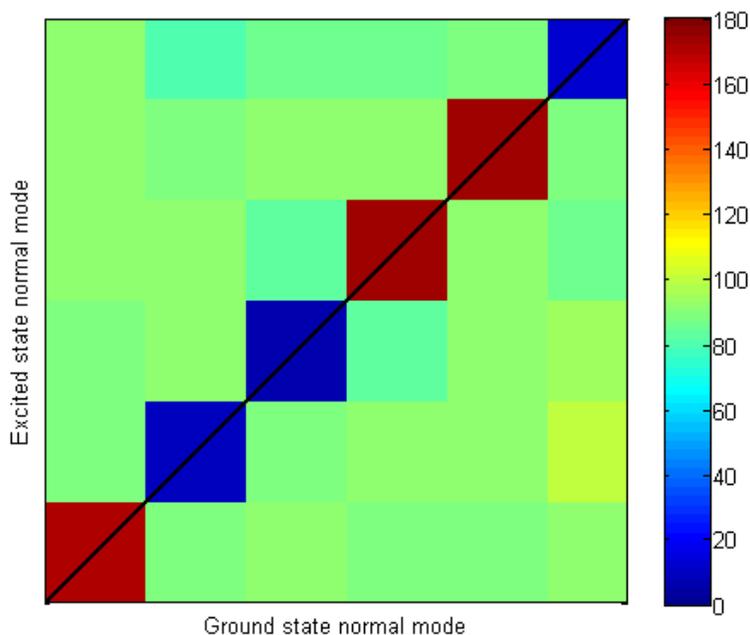


Figure 2. 6. Duschinsky rotation matrix of formaldehyde between electronic ground and first excited states

Along the horizontal axis and vertical axis are the electronic ground state and first excited state vibrational normal modes, respectively, ordered by their frequencies. Each block connects a vibrational normal coordinate of the electronic ground state normal mode and a vibrational normal coordinate of the first excited state normal mode. The Duschinsky rotation matrix elements are converted to the angle between these two normal coordinates and plotted in different color. The color of the block indicates the values of these angles. Dark blue (0°) indicates the two normal coordinates are parallel, while dark red (180°) indicates the two normal coordinates are antiparallel

2.3 BIBLIOGRAPHY

- (1) Long, D. A. *The Raman Effect: A unified treatment of the theory of raman scattering by molecules*; John Wiley & Sons Ltd.: West Sussex, England, 2002.
- (2) Schatz, G. C.; Ratner, M. A. *Quantum Mechanics in Chemistry*; Dover Publication, Inc: Mineola, New York, 2002.

- (3) Steinfeld, J. I. *Molecules and Radiation: An Introduction to Modern Molecular Spectroscopy*; 2nd ed.; Dover Publication, Inc: Mineola, New York, 2005.
- (4) Lee, S.-Y.; Heller, E. J. *J. Chem. Phys.* **1979**, *71*, 4777.
- (5) Craig, D. P.; Thirunamachandran, T. *Molecular Quantum Electrodynamics: An Introduction to Radiation Molecule Interactions*; Dover Publication, Inc.: Mineola, New York, 1998.
- (6) Myers, A. B.; Mathies, R. A. In *Biological Applications of Raman Spectroscopy*; Spiro, T. G., Ed.; John Wiley & Sons: New York, 1987; Vol. 2.
- (7) Myers, A. B.; Mathies, R. A.; Tannor, D. J.; Heller, E. J. *J. Chem. Phys.* **1982**, *77*, 3857.
- (8) Heller, E. J. *J. Chem. Phys.* **1975**, *62*, 1544.
- (9) Tannor, D. T. *Introduction to Quantum Mechanics: A Time-Dependent Perspective*; University Science Books: Sausalito, California, 2007.
- (10) Tannor, D. J.; Heller, E. J. *J. Chem. Phys.* **1982**, *77*, 202.
- (11) Heller, E. J.; Sundberg, R.; Tannor, D. *J. Phys. Chem.* **1982**, *86*, 1822.
- (12) Chen, X. G.; Asher, S. A.; Schweitzerstener, R.; Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1995**, *117*, 2884.
- (13) E. Bright Wilson, J.; Decius, J. C.; Cross, P. C. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; Dover Publication, Inc.: New York, 1980.
- (14) Groner, P. In *Handbook of Vibrational Spectroscopy*; John Wiley & Sons, Ltd, 2006; Vol. III.
- (15) Lee, S.-H.; Palmo, K.; Krimm, S. *J. Comput. Chem.* **1999**, *20*, 1067.
- (16) McIntosh, D. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **2009**, *125*, 177.
- (17) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.
- (18) Henderson, J. R.; Muramoto, M.; Willett, R. A. *J. Chem. Phys.* **1964**, *41*, 580.
- (19) Sando, G. M.; Spears, K. G. *J. Phys. Chem. A* **2001**, *105*, 5326.
- (20) Eckart, C. *Phys. Rev.* **1935**, *47*, 552.
- (21) Hougen, J. T.; Watson, J. K. G. *Can. J. Phys.* **1965**, *43*, 298.
- (22) Louck, J. D.; Galbraith, H. W. *Rev. Mod. Phys.* **1976**, *48*, 69.
- (23) Chen, P. In *Unimolecular and Bimolecular Ion-Molecule Reaction Dynamics*; Ng, C.-Y., Baer, T., Powis, I., Eds.; John Wiley & Sons: Chichester, 1994.
- (24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2010., 2010.

3.0 UV RESONANCE RAMAN STUDY OF PROTEIN SECONDARY STRUCTURES

Many theoretical and experimental approaches have been developed to study protein folding. Currently, limited by the computational power available, it is impossible to simulate protein folding from first principles. Monte Carlo (MC) and molecular dynamics (MD) with empirical or semiempirical force fields are two practical simulation methods currently available for protein folding.¹

Various experimental techniques, especially spectroscopy, have been applied to study protein folding. X-ray diffraction (XRD) with angstrom resolution is a superior technique to study the protein molecular structure. Unfortunately, it cannot provide information on solution protein dynamics. Nuclear magnetic resonance (NMR) spectroscopy is another powerful method for studying protein folding. It can be applied not only to static state structure, but also to fleeting dynamic structures at a temporal resolution of microseconds. However, the temporal resolution is not high enough for protein folding where fast events occur on the order of nanoseconds. Optical spectroscopy has excellent temporal resolution. The ultrafast laser spectroscopy and pump-probe techniques can increase the time resolution to femtoseconds. Circular dichroism (CD) spectroscopy is a useful and classical method to study protein secondary structures. Vibrational spectroscopy gives many details about the proteins. However, infrared (IR) spectroscopy can only be carried out in D₂O. Compared with the above techniques, UV resonance Raman (UVRR) spectroscopy offers many advantages to study protein conformations and folding.²⁻³

3.1 ADVANTAGES OF UV RESONANCE RAMAN SPECTROSCOPY

3.1.1 High sensitivity

The intensity of a Raman band for orientationally averaged molecules is given by:⁴⁻⁵

$$I = \sigma^R N I_0 W(\Omega) \quad (3.1)$$

where σ^R is the Raman scattering cross section associated with the vibrational band of interest with units of $\text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$. N is the number of molecules irradiated and I_0 is the intensity of the excitation. $W(\Omega)$ is a geometric parameter depending on the solid angle of collection. The spontaneous Raman cross section σ^R is proportional to the square of polarizability tensor and approximately to the fourth power of the excitation frequency. The expression for σ^R is:

$$\sigma^R = \frac{1}{32\pi^2 \epsilon_0^2 c^4} \omega_0 (\omega_0 - \omega_{fi})^3 |\alpha|^2 \quad (3.2)$$

In general, the Raman scattering cross section is very small ($\sim 10^{-30} \text{ cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$), compared to that of absorbance ($\sim 10^{-17} \text{ cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$) or fluorescence ($\sim 10^{-20} \text{ cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$). For normal Raman scattering, the intensity increases as the excitation frequency, ω_0 , increases, according to a fourth power relationship between σ^R and ω_0 as shown in eq. (3.2). In normal Raman spectroscopy, the sample is usually excited in the visible or near IR region with wavelength between 1200 and 400 nm. For a colorless material, which has little absorbance in the visible region, tuning the excitation frequency from 1200 nm to 400 nm will increase the Raman intensity by a factor of 81 (3^4). If the excitation wavelength is further lowered to 200 nm, an

additional factor of 16 can be achieved. Therefore, in general, moving the excitation to shorter wavelength helps to increase the sensitivity.

Tuning the excitation into resonance of an electronic transition gives an even larger enhancement of the Raman scattering cross section. The resonance Raman polarizability tensor estimated by eq. (2.24) is usually 2 or 3 orders of magnitude larger than the normal Raman polarizability tensor in eq. (2.18), which leads to an enhancement factor of 10^4 - 10^6 . This great enhancement allows studies on dilute samples and enables investigations of biological samples in living cell or tissue.

3.1.2 High selectivity

Resonance Raman scattering does not equally increase the intensities of all vibrational modes. For a particular resonant electronic excited state, only the vibrational modes with large Frank-Condon integrals are resonance enhanced. This requires the electronic excited state potential surface to significantly shift from the ground state minimum along those vibrational modes, which results in a large force on the excited state potential surface at the ground state equilibrium geometry. This fact enables resonance Raman spectroscopy to selectively enhance different vibrations by tuning the excitation wavelength to different electronic transition wavelengths.

The selectivity of resonance Raman spectroscopy is important for the study of macromolecules, especially proteins and DNA. For macromolecules, say, of 100 atoms, the number of vibrational modes would be too large (294) to separate one vibration from other vibrations in the spectrum. By tuning the excitation wavelength, we can selectively enhance the vibrational modes that we are interested in.

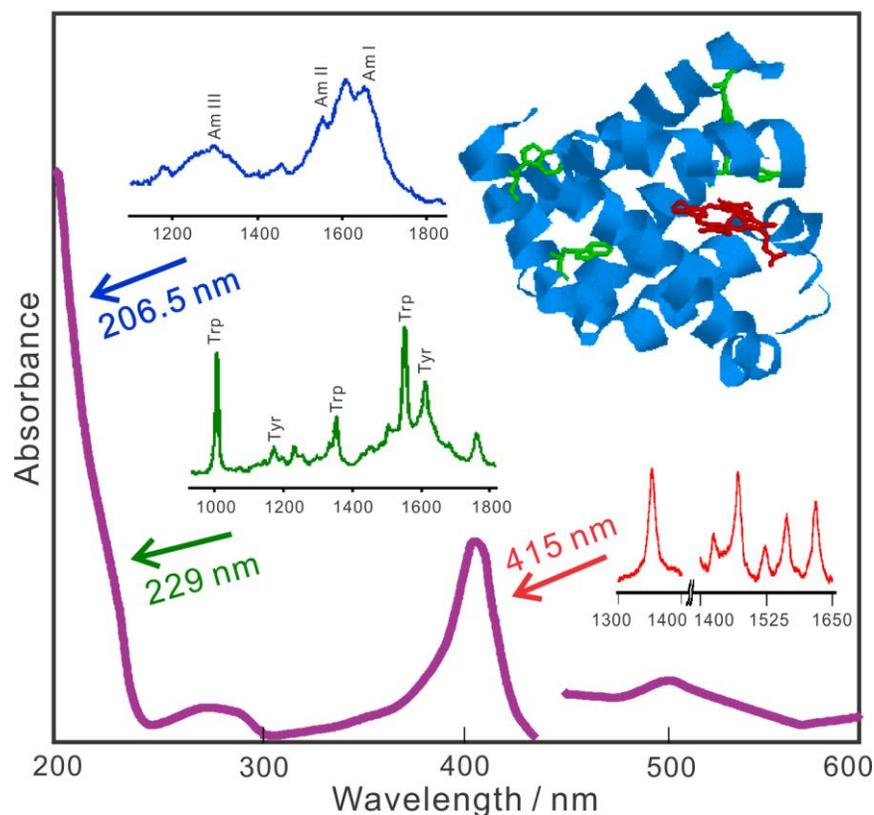


Figure 3. 1. Selectively excitation of myoglobin resonance Raman spectra

Figure 3. 1 shows an example of myoglobin Raman spectra excited at different wavelengths⁶. The weak absorption band at ~500 nm involves a charge transfer transitions between the iron and its ligands. The Raman spectrum excited at this wavelength therefore shows the vibrations involving the iron and its ligands. The Raman spectrum excited at the Soret band (~400 nm) of the myoglobin heme π - π^* transition is dominated by the in-plane vibrations of heme group. The absorption band between 300 and 270 nm is due to aromatic side chains, tryptophan (Trp) and tyrosine (Tyr). The large absorbance between 250 and 220 nm arises from Trp and Tyr as well. Therefore, excitation between 220 to 230 nm gives rise to a resonance Raman spectrum with predominately Trp and Tyr vibrations with little peptide backbone amide vibrations. The very intense absorbance below 220 nm derives from the peptide amide π - π^* transition. Raman spectra excited below 220 nm result in the great enhancement of the amide

backbone vibrations, which contain information on protein secondary structure because the frequencies and intensities of these vibrational bands are sensitive to the protein secondary structure and H-bonding states of the amides. The high selectivity of UVRR spectroscopy also allows the study of proteins in complex environments, such as membrane proteins,⁷⁻⁸ protein-ligand binding,⁹ and protein-DNA interaction.¹⁰⁻¹¹

3.1.3 Excellent temporal resolution

A complete protein (un)folding process occurs on a time regime of picoseconds to seconds; the (un)folding of peptide secondary structures usually occurs from the nanosecond to microsecond time regime; and fast events, such as proton transfer and H-bonding, may extend into the picosecond time regime. Traditional techniques such as NMR do not have the temporal resolution to study proteins (un)folding process on those fast time scales at the molecular level.

For vibrational spectroscopy, such as UVRR spectroscopy, the required temporal resolution can be easily achieved by using pump-probe techniques with short pulsed lasers.

The temperature-jump (T-jump) technique involving an IR pump pulse and a UV probe pulse has been developed and used to investigate the (un)folding of peptide secondary structures.¹²⁻¹⁸ Water has an absorbance band attributed to the combination mode of O-H stretching and H-O-H bending at $\sim 1.9 \mu\text{m}$ with an absorption coefficient of $\sim 130 \text{ cm}^{-1}$, and also the overtone of O-H stretching at $\sim 1.5 \mu\text{m}$ with an absorption coefficient of $\sim 32 \text{ cm}^{-1}$.¹⁹⁻²⁰ The vibrational relaxation of the water combination mode and overtone are fast ($<10 \text{ ps}$),²¹ giving rise to a heating temporal profile shown in Figure 3. 2 (blue curve). Therefore, both near IR light at $1.9 \mu\text{m}$ ²²⁻²³ and $1.5 \mu\text{m}$ ²⁴⁻²⁵ can serve as the pump wavelength to heat the sample.

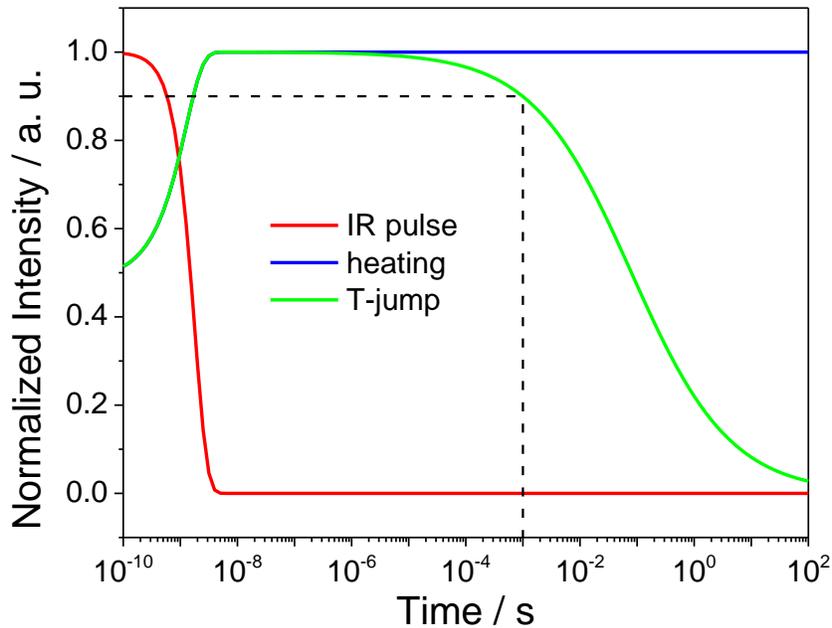


Figure 3. 2. Temporal profile of IR pulse, water heating and T-jump resulted from the IR pulse

The IR pulse is assumed to have a Gaussian temporal profile with a FWHM of 3 ns. The heating curve is the integral of the IR pulse intensity. The overall T-jump curve is calculated from the heating curve assuming $\Delta T = \frac{T_{\text{heat}}}{a+2\sqrt{\pi D_{\text{heat}} t}}$

$$\text{with } a = 1 \text{ and } D_{\text{heat}} = 1$$

The diffusion of heat from the hot spot to external cool regions is fairly slow compared with the heating step,²⁵ which results in an overall T-jump temporal profile (green curve) in Figure 3. 2. The T-jump stays almost constant until ~ 1 ms, when the temperature decreases by 10%. It can be seen that the laser induced T-jump can last for a large time window from 10 ns to 1 ms time regime, where most secondary structure (un)folding processes occur.

The T-jump temporal profile shown in Figure 3. 2 is for static samples. The slow cooling time limits maximum the laser repetition rate. However, for flowing samples, the hot spot can be moved away from the probing spot before it cools down. For example, for a typical size probing spot, say 100 μm , on a 1 $\text{m}\cdot\text{s}^{-1}$ stream, the heated sample stays in probing spot for <100 μs . This allows the use of kHz lasers, which significantly increases the signal.

3.2 CORRELATIONS BETWEEN PROTEIN SECONDARY STRUCTURES AND UV RESONANCE RAMAN SPECTRA

UVRR can selectively enhance vibrations of interest. The backbone amide vibrations contain information about H-bonding and Ramachandran dihedral angles that can be used to identify the secondary structures. The side chain vibrations report on side chain H-bonding and hydration that are relevant to tertiary structures.

3.2.1 Resonance enhanced amide vibrations

The amide frontier orbitals involve mixing of two oxygen sp^2 orbitals and three p orbitals from oxygen, carbon and nitrogen, respectively. The resultant frontier molecular orbitals have two non-bonding lone pair orbitals and three π orbitals as shown in Figure 3. 3. An n orbital is the highest occupied molecular orbital (HOMO) and the π^* orbital is the lowest unoccupied molecular orbital (LUMO).

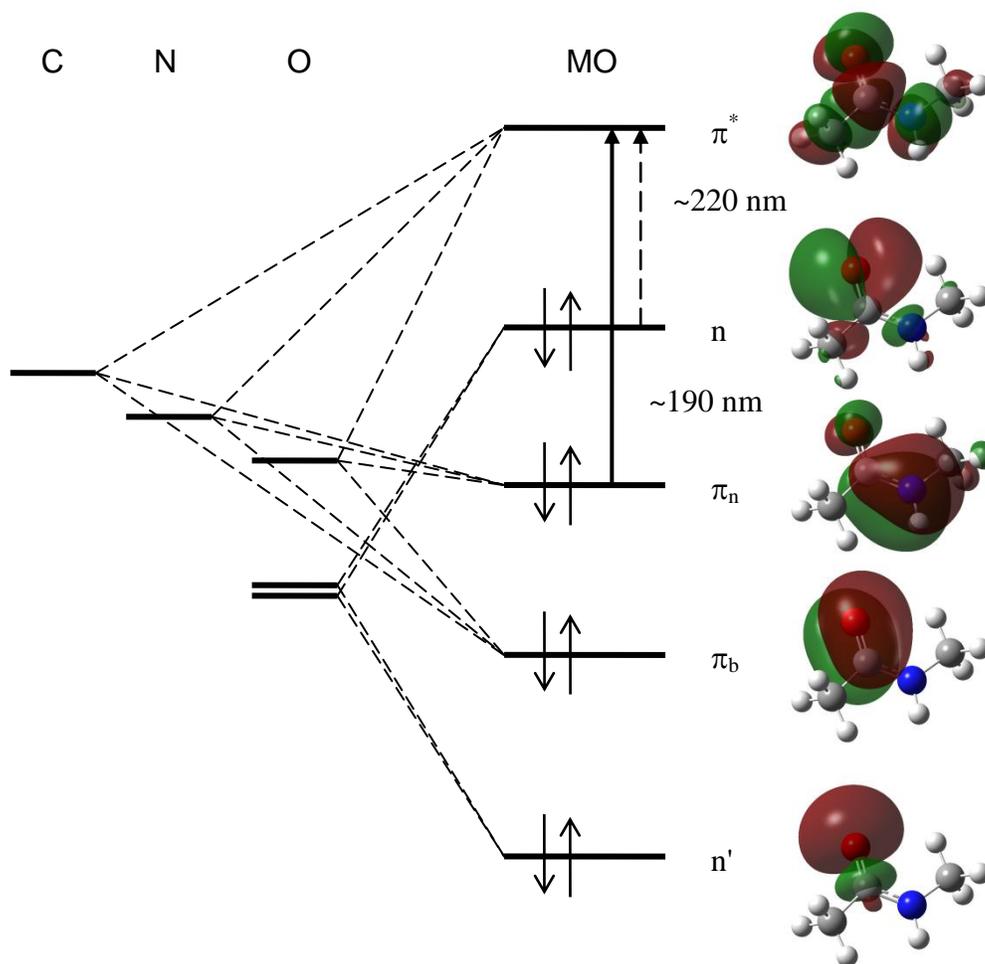


Figure 3. 3. Frontier molecular orbitals and electronic transitions in an amide unit

The lowest $n\text{-}\pi^*$ transition at ~ 220 nm is magnetically allowed but electronically forbidden. The first electronically allowed transition is therefore the $\pi_n\text{-}\pi^*$ transition at ~ 190 nm.²⁶ This $\pi_n\text{-}\pi^*$ transition results in significant displacements along carbonyl and C-N bond: an elongated carbonyl and an elongated C-N bond.²⁷ Thus any vibrations that involve CO and CN stretching are resonance enhanced if the amide bond is excited at ~ 190 nm. N-methylacetamide (NMA) is a classic model compound that mimics the peptide amide unit. Table 3. 1 shows the comparison of NMA calculated bond lengths between ground state equilibrium geometry and the second electronic excited state equilibrium geometry in vacuum. Upon excited to the second

electronic excited state, the C-N bond length increases most, followed by the C=O bond, compared with the ground state equilibrium geometry. The C $_{\alpha}$ -C bond contracts, while the N-H and N-C $_{\alpha}'$ bond length changes slightly. This result qualitatively agrees with previously determined NMA π - π^* excited state geometry by using UVRR spectroscopy.²⁷

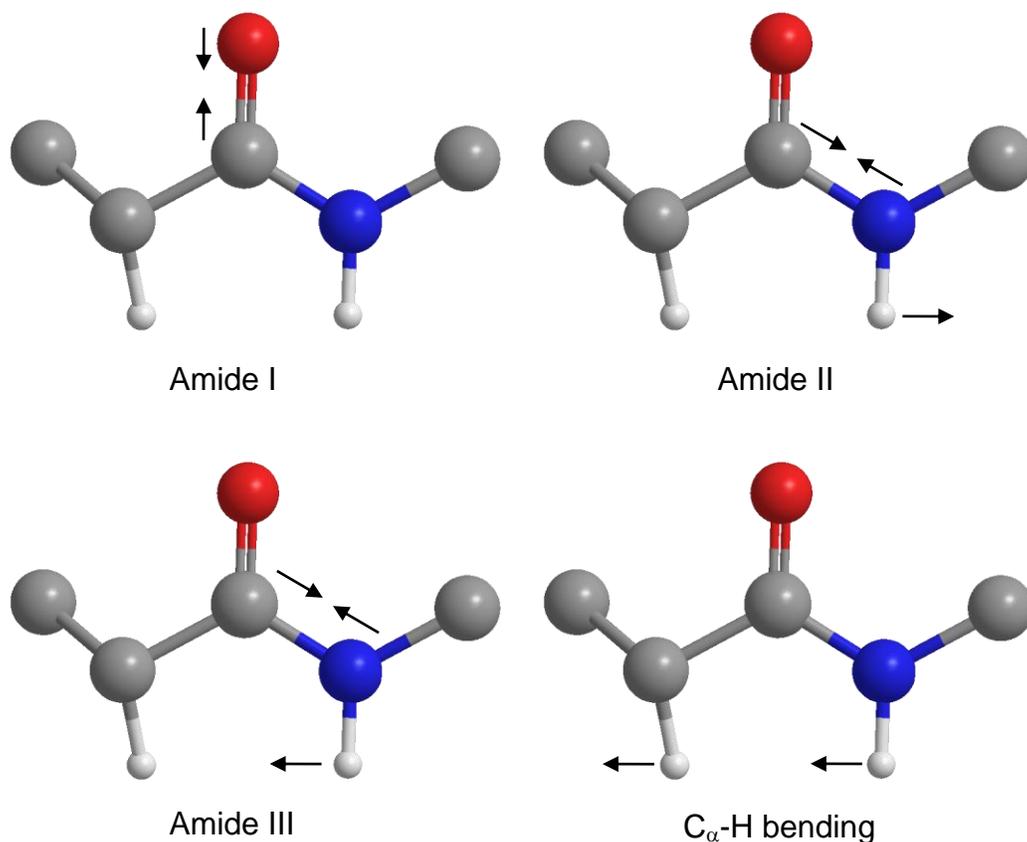


Figure 3. 4. Resonance enhanced amide backbone vibrations

In UVRR spectra, there are four amide vibrations within the region 1000-2000 cm^{-1} , namely the Amide I (Am I), Amide II (Am II), Amide III (Am III) and C $_{\alpha}$ -H bending (Figure 3. 4).²⁸⁻²⁹ The Am I vibration at $\sim 1660 \text{ cm}^{-1}$ is mainly contributed from the C-O stretching with a slight amount of C-N stretch. Am II vibration at $\sim 1550 \text{ cm}^{-1}$ and Am III vibration at $\sim 1250 \text{ cm}^{-1}$ involve the coupling of C-N stretching and N-H in-plane bending with different phases. Potential energy distributions (PED) of NMA vibrations are listed in Table 3. 2. The C $_{\alpha}$ -H bending at

$\sim 1380\text{ cm}^{-1}$ is resonance enhanced due to the coupling to N-H bending at some Ramachandran dihedral angles.³⁰ The assignment of these vibration bands in proteins has been well studied.³¹⁻³³

Table 3. 1. NMA Bond lengths of ground state and second electronic excited state

Bond	Bond length / Å		
	Ground state	Second excited state	Difference
C-O	1.2252	1.2625	0.0373
C-N	1.3691	1.4304	0.0613
N-H	1.0074	1.0079	0.0005
N-C $_{\alpha}'$	1.4520	1.4545	0.0025
C $_{\alpha}$ -C	1.5173	1.4999	-0.0175

*The ground state was optimized at MP2 / 6-311++G** level of theory and the second electronic excited state was optimized at CIS / 6-311++G** level of theory by using Gaussian 09.³⁴

Table 3. 2. Potential energy distribution of NMA amide vibrations

Vibrations	Frequency / cm^{-1}	Potential energy distribution (%)
Am I	1744	C=O str (76), C-C $_{\alpha}$ N ipb (7), C-N str (6)
Am II	1559	N-H ipb (41), C-N str (23), C $_{\alpha}'$ -H $_3$ asym def (12), C $_{\alpha}'$ -H $_3$ rock (10)
Am III	1265	C-N str (29), N-H ipb (20), C-O ipb (15), C $_{\alpha}'$ -H $_3$ asym def (11), N-C $_{\alpha}'$ str (5)

*Calculated at B3LYP / 6-311++G** level of theory with an optimized geometry at the same level of theory by using Gaussian 03.³⁵ The frequencies are not scaled. PED are cut off at 5%. str: stretching, ipb: in-plane bending, rock: rocking, asym def: asymmetric deformation.

3.2.2 Correlations between amide hydrogen bonding and UV resonance Raman spectrum

Amides can participate in H-bonding. The carbonyl oxygen with two lone pairs of electrons can be an H-bond acceptor, and the N-H hydrogen can be an H-bond donor. For example, NMA can H-bond with three water molecules as shown in Figure 3. 5. The formation of H-bonds can affect the vibration frequencies. The impact of H-bonding can be understood through the resonance model.³⁶ Figure 3. 6 shows two NMA resonance structures.

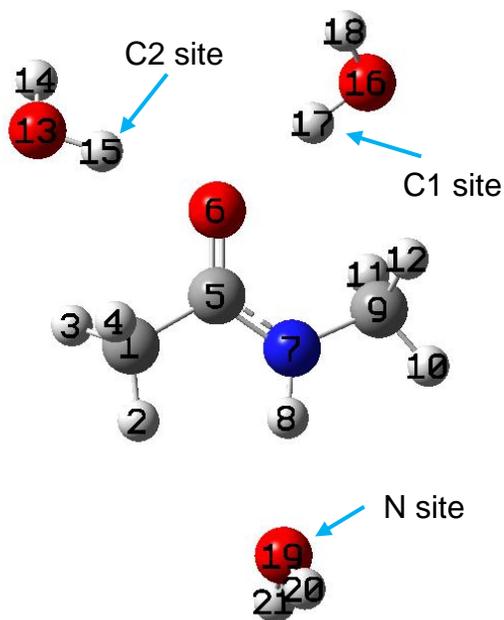


Figure 3. 5. Definition of atomic labels and H-bonding sites in NMA

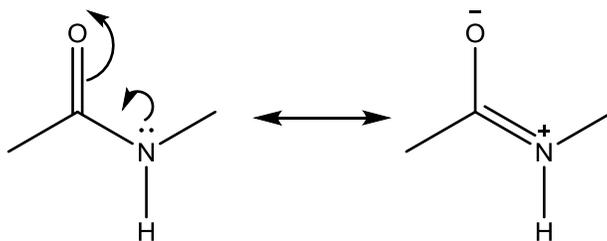


Figure 3. 6. Resonance structures of NMA

Table 3. 3. Bond length in NMA for different H-bonded waters

Bond	Bond length / Å					
	NMA	NMA-H ₂ O (N)	NMA-H ₂ O (C1)	NMA-H ₂ O (C2)	NMA-H ₂ O (C1, C2)	NMA-H ₂ O (N, C1, C2)
C-O	1.2211	1.2247	1.2291	1.2304	1.2387	1.2424
C-N	1.3676	1.3599	1.3552	1.3566	1.3482	1.3432
N-H	1.0069	1.0113	1.0065	1.0064	1.0068	1.0132
N- C _α '	1.4541	1.4538	1.4602	1.4570	1.4619	1.4597
C _α -C	1.5192	1.5200	1.5159	1.5139	1.5114	1.5120

*The geometry was optimized at B3LYP / 6-311++G** level of theory with PCM by using Gaussian 03³⁵ or Gaussian 09.³⁴

Table 3. 4. NBO bond order in NMA with different H-bonded waters

Bond	Bond order					
	NMA	NMA-H ₂ O (N)	NMA-H ₂ O (C1)	NMA-H ₂ O (C2)	NMA-H ₂ O (C1, C2)	NMA-H ₂ O (N, C1, C2)
C-O	1.66	1.65	1.60	1.60	1.54	1.52
C-N	1.19	1.21	1.23	1.22	1.26	1.29
N-H	0.83	0.78	0.82	0.82	0.82	0.77
N- C _α '	0.97	0.98	0.96	0.97	0.96	0.96
C _α -C	0.99	0.99	1.00	1.00	1.01	1.01

*The geometry was optimized at B3LYP / 6-311++G** level of theory with PCM by using Gaussian 03³⁵ or Gaussian 09.³⁴

H-bonding donors residing at C1 and C2 sites (Figure 3. 5) that can form H-bonds with the C=O stabilize the zwitterionic form of an amide (Figure 3. 6). Thus stronger H-bonding at these sites results in a longer C=O bond length and a lower C=O bond order. As a consequence, the Am I vibration, mainly C=O stretching, (Table 3. 2) downshifts its frequency. These H-bonds also decrease the C-N bond length and increase the C-N bond order, usually giving rise to upshifted Am II and Am III frequencies. However, the Am II vibration contains relatively little C-N stretching (Table 3. 2). Consequently, the upshift is insignificant.³⁷⁻³⁸

An H-bonding acceptor residing at the N site (Figure 3. 5), which is H-bonded to the N-H hydrogen, can also stabilize the zwitterionic form of the amide by donating some electron density to the N-H group. Therefore, a stronger H-bonding at the N site downshifts the Am I frequency and upshifts the Am II and Am III frequencies similar to what occurs when there is the H-bonding at C1 and C2 sites (Figure 3. 5). Additionally, a stronger H-bonding at the N site can increase the N-H bending frequency.³⁹ As a consequence, the Am II and Am III frequencies will further increase.

Table 3. 3 and Table 3. 4 list the calculated bond lengths and bond orders of relevant NMA amide bonds in different H-bonding states with water calculated by DFT and the natural bond orbital (NBO) method. The impact of H-bonding on NMA bond length change are very similar to a previously reported calculation performed at HF / 6-31++G** level of theory.⁴⁰ These calculated results are consistent with the previous qualitative analysis that H-bonding stabilizes the zwitterionic form of an amide unit. H-bonding at any site decreases the C=O bond order while increasing the C-N bond order, which is in agreement with the C=O and C-N bond length changes upon H-bonding. The impact of H-bonding on N-C α' and C α -C bond lengths are less significant but still noticeable. Table 3. 5 lists the NBO atomic charges in NMA upon H-

bonding with water. When NMA is H-bonded, the H-bonding donor and/or the acceptor pulls and/or pushes the electron density through O-C-N p orbitals to change the bond lengths and bond orders.

Table 3. 5. NBO atomic charges in NMA with different H-bonded waters

Atom	Charge					
	NMA	NMA-H ₂ O (N)	NMA-H ₂ O (C1)	NMA-H ₂ O (C2)	NMA-H ₂ O (C1, C2)	NMA-H ₂ O (N, C1, C2)
C1 (C _α)	-0.658	-0.661	-0.655	-0.664	-0.663	-0.663
H2	0.193	0.203	0.196	0.196	0.199	0.209
H3	0.228	0.222	0.230	0.245	0.252	0.242
H4	0.228	0.222	0.231	0.236	0.235	0.231
C5 (C)	0.672	0.669	0.685	0.685	0.698	0.697
O6	-0.642	-0.653	-0.681	-0.675	-0.720	-0.734
N7	-0.634	-0.651	-0.618	-0.621	-0.604	-0.621
H8	0.380	0.420	0.387	0.388	0.390	0.427
C9 (C _{α'})	-0.373	-0.365	-0.375	-0.366	-0.377	-0.376
H10	0.187	0.189	0.194	0.194	0.196	0.195
H11	0.187	0.200	0.204	0.209	0.205	0.202
H12	0.231	0.194	0.220	0.198	0.225	0.219
Total	0.000	-0.010	0.019	0.025	0.036	0.029

*The geometry was optimized at B3LYP / 6-311++G** level of theory with PCM by using Gaussian 03³⁵ or Gaussian 09.³⁴

Table 3. 6 lists some calculated NMA Am I, Am II and Am III frequencies in different H-bonding states with different partners. In general, stronger H-bonding downshifts the Am I frequency and upshifts Am II and Am III frequencies.

Table 3. 6. NMA vibrational frequency dependence on H-bonding state in vacuum

N site	C1 site	C2 site	Am I freq / cm ⁻¹	Am II freq / cm ⁻¹	Am III freq / cm ⁻¹
-	-	-	1744	1559	1265
H ₂ O	-	-	1732	1570	1296
EtOH	-	-	1731	1574	1300
NMA	-	-	1732	1600	1298
EtSH	-	-	1734	1560	1300
EtSMe	-	-	1732	1562	1294
(EtO) ₂ PO ₂ ⁻	-	-	1700	1638	1341
EtOSO ₃ ⁻	-	-	1704	1617	1330
-	H ₂ O-d ₂	-	1719	1557	1290
-	EtNH ₃ ⁺ -d ₈	-	1648	1589	1321
-	-	H ₂ O-d ₂	1712	1555	1291
-	-	NMA	1718	1570	1283
-	-	EtNH ₃ ⁺ -d ₈	1648	1589	1321
-	H ₂ O-d ₂	H ₂ O-d ₂	1687	1568	1304
-	H ₂ O-d ₂	NMA	1693	1568	1301

NMA	-	NMA	1698	1602	1326
H ₂ O-d ₆	H ₂ O-d ₆	H ₂ O-d ₆	1679	1597	1329
EtOH-d ₆	EtOH-d ₆	EtOH-d ₆	1677	1600	1332
TFE-d ₃	TFE-d ₃	TFE-d ₃	1673	1598	1332
H ₂ O-d ₂	H ₂ O-d ₂	NMA	1686	1598	1328
NMA	H ₂ O-d ₂	NMA	1679	1614	1340
(EtO) ₂ PO ₂ ⁻	H ₂ O-d ₂	H ₂ O-d ₂	1665	1671	1368
EtOSO ₃ ⁻	H ₂ O-d ₂	H ₂ O-d ₂	1671	1648	1631

*Calculated at B3LYP / 6-311++G** level of theory with optimized geometry at the same level of theory by using Gaussian 03³⁵ or Gaussian 09.³⁴ The frequencies are not scaled. Some ligands are deuterated to remove coupling between NMA amide vibrations and ligand vibrations.

3.2.3 Correlations between protein secondary structures and UV resonance Raman spectra

A protein or peptide contains multiple amide units and therefore multiple amide vibrations. The Am I vibration usually has a strong transition dipole moment and could couple with other in different amide units. The Am I vibration is widely used in IR spectroscopy to identify protein secondary structures. In UVRR spectroscopy, however, the Am I vibration is fairly insensitive to secondary structure changes.

Unlike the Am I, the Am II, Am III and C_α-H bending vibrations have much smaller transition dipole moments. The coupling of these vibrations between different amide units is negligible. Therefore, the vibrations are essentially localized in the amide unit and contain only

the information about that specific amide unit without interference from other units.⁴¹⁻⁴³ The UVRR spectrum of a whole protein or peptide can be modeled as a sum of spectra from all the individual amide units. For example, Figure 3. 7 shows a modeling of the 204 nm excited UVRR spectra of oligoglycines as a sum of the internal peptide bond spectra and the flanking peptide bond spectra.⁴⁴ If different secondary structures occur at different locations within a peptide, the overall peptide spectrum can be modeled as a sum of the basis spectra of different secondary structures. The fractions of occurring secondary structures can be readily calculated if the basis spectra are known.

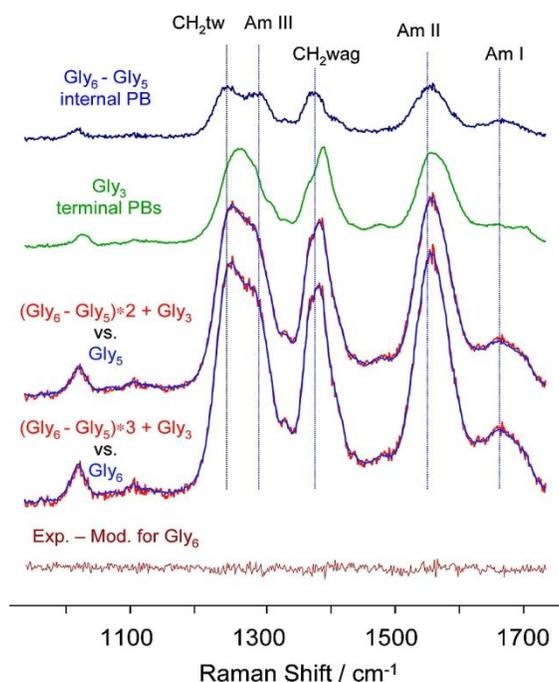


Figure 3. 7. 204 nm excited UVRR spectra of oligoglycines

Reproduced with permission from Ref⁴⁴

A quantitative methodology based on the linear modeling of measured spectra with basis spectra to determine the protein secondary structure has been developed.⁴⁵ The basis spectra of α -helix, β -sheet and unordered structure has been obtained from the UVRR spectra of a set of proteins with known fractions of secondary structures by applying a chemometrics method. By

fitting the unknown protein spectra with these basis spectra, the fractions of α -helix, β -sheet and unordered structure can be determined.

Figure 3. 8 shows 206.5 nm excited Raman spectra of horse myoglobin at various pH values.⁴⁶⁻⁴⁷ The pH dependent secondary structure fractions were calculated using the above quantitative method. As the pH decreases, the Am III band intensity increases and downshifts, the C_{α} -H band intensity increases significantly, and the Am I band shows a slight upshift.

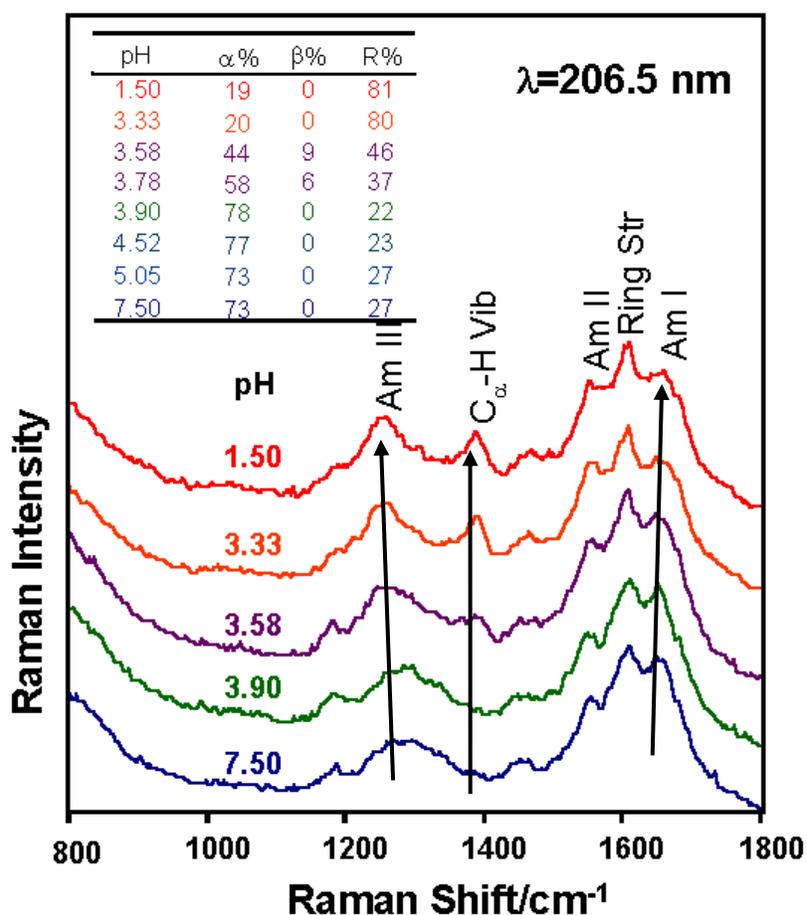


Figure 3. 8. UVRR spectra of horse myoglobin dependent on pH

Reproduced with permission from Ref⁴⁶

Figure 3. 9 shows the 204 nm excited Raman spectra of polyglutamic acid (PGA) at different temperatures. The temperature dependent secondary structure fractions were calculated

using the previous described quantitative method. As the temperature increases, α -helix melts to PPII conformation.⁴⁸ The spectral change includes a slight upshift of the Am I band, a significant C_{α} -H band intensity increase, an Am III band intensity increase and a frequency downshift.

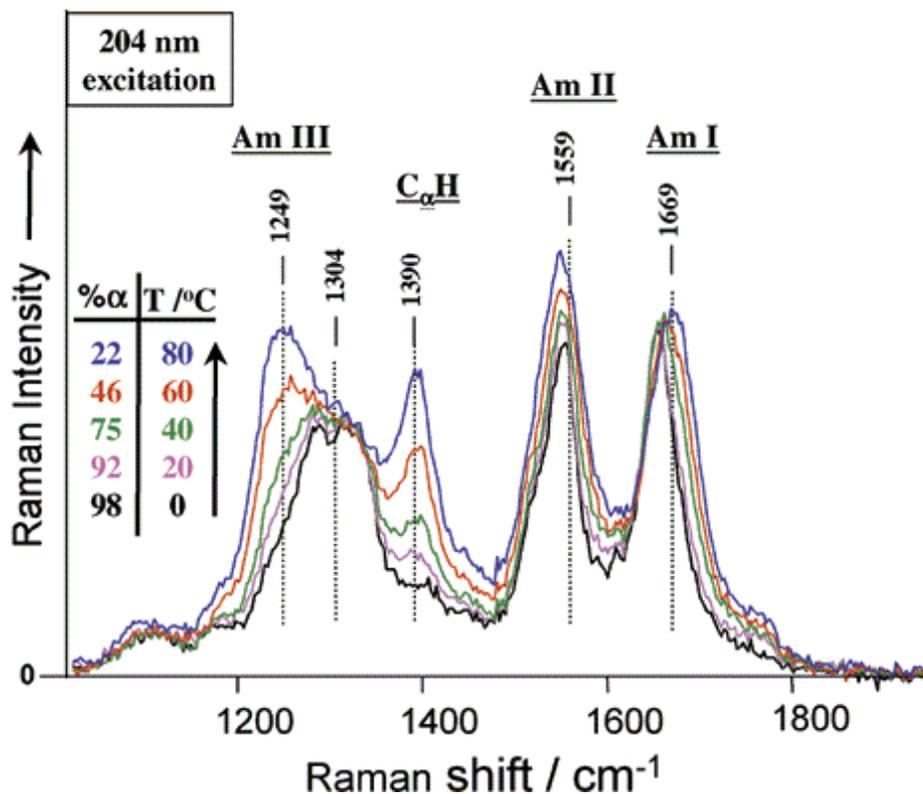


Figure 3. 9. UVRR spectra of PGA dependent on temperature

Reproduced with permission from Ref³⁰

These spectral changes arise from two origins: the change in H-bonding state and the change in conformation. The Am I band change is mainly due to the change of H-bonding as discussed in Section 3.2.2. As the conformation changes from α -helix to unordered structure, which is now proven to be PPII-like conformation⁴⁸⁻⁵¹, the intra-molecular H-bonds are broken and form weaker amide-water inter-molecular H-bonds, which upshifts the Am I band frequency.^{38,52-54}

The C_{α} -H band is also a marker to distinguish different secondary structures. If the C_{α} -H is "cis" to the N-H bond ($\varphi = \sim -75^{\circ}$, $\psi = \sim 150^{\circ}$, β -sheet or PPII conformation, Figure 3. 10A), the C_{α} -H bending couples with the N-H bending and hence is resonance enhanced. In contrary, when it is in "trans" position ($\varphi = \sim -60^{\circ}$, $\psi = \sim 45^{\circ}$, α -helix-like conformation, Figure 3. 10B), the spectrum will show little C_{α} -H band intensity.^{30,55}

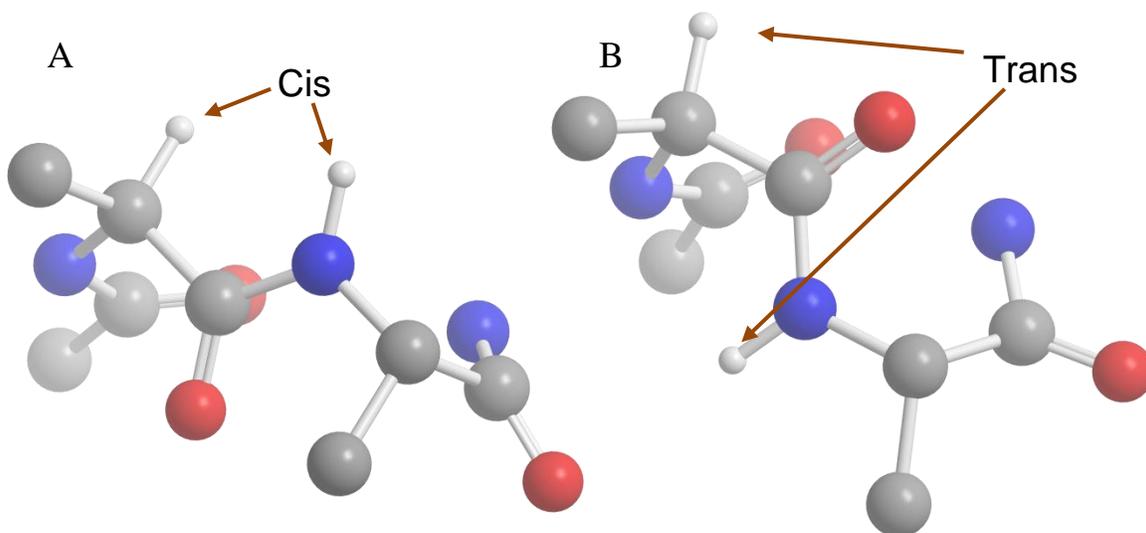


Figure 3. 10. Relative positions of C_{α} -H and N-H bonds in different configurations

(A) cis-configuration, (B) trans-configuration

The Am III band is sensitive to peptide conformation despite a small dependence on H-bonding. N-H bending contributes substantially to the Am III band (Table 3. 2). The coupling between the C_{α} -H bending and the N-H bending not only gives rise to the resonance enhancement of the C_{α} -H bending mode, but also affects the Am III frequency.³⁰ A larger coupling will result in a bigger downshift of the Am III frequency. This coupling depends on the projection of the C_{α} -H bending motion on the N-H bending, and has a sine dependence on the Ramachandran ψ angle.

The Am III frequency can be correlated to the ψ angle⁵⁶⁻⁵⁹ as

$$\nu(\psi, T, \text{H-bonding}) = \nu_0(T, \text{H-bonding}) + \Delta\nu \sin(\psi + \psi_0) \quad (3.3)$$

The offset, ν_0 , depends on temperature, T , and H-bonding state of the peptide bond. The amplitude, $\Delta\nu$, and phase, ψ_0 , can be experimental determined from small peptide crystals with known ψ angles. Figure 3. 11 plots the correlation between Am III band frequency and ψ angles. If the impact of temperature and H-bonding can be specified, we can use this correlation to calculate the ψ angle from the Am III frequency for monitoring the secondary structures.^{55,60-63}

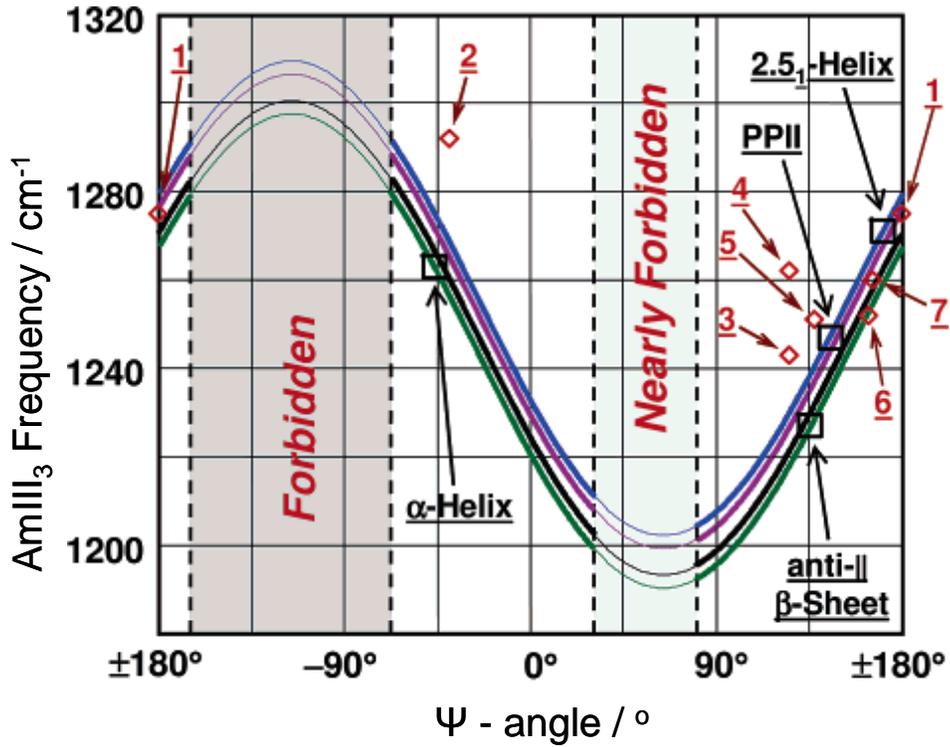


Figure 3. 11. The correlation between Am III band frequency and ψ angle

Reproduced with permission from Ref⁵⁵

3.2.4 Ramachandran ψ angle distribution and Gibbs free energy landscapes of peptide bond

Previous studies on small peptide crystals also found the homogeneous linewidth of Am III band to be 15 cm^{-1} (FWHM).⁴⁸ This value is significantly narrower than the common Am III bandwidth in solution spectrum, indicating that the Am III band in solution spectrum contains multiple ψ angle contributions that come either from a single fluctuating secondary structure or multiple different secondary structures. To distinguish these two situations, we could calculate the probability distribution along the ψ angle coordinates from the Am III band deconvolution.

An observed Am III band, \mathcal{S} , is an integral of continuously distributed Lorentzian bands

$$\mathcal{S} = \int \mathcal{S}_L(\nu_L, I_L(\nu_L), w_L(\nu_L)) d\nu_L \quad (3.4)$$

where $\mathcal{S}_L(\nu_L, I_L(\nu_L), w_L(\nu_L))$ is the Lorentzian band depending on center frequency, ν_L , intensity, $I_L(\nu_L)$, and band width, $w_L(\nu_L)$. In a discrete form, the integral is replaced with summation

$$\mathcal{S} = \sum_{\nu_L} \mathcal{S}_L(\nu_L, I_L(\nu_L), w_L(\nu_L)) \Delta\nu_L \quad (3.5)$$

Assuming the homogeneous line width does not change with different ψ angles, we can deconvolute the Am III band into a series of Lorentzian bands with a fixed homogeneous bandwidth w_L . The deconvoluted Lorentzian frequency interval, $\Delta\nu_L$, can be arbitrarily chosen.

By applying eq. (3.3), we could convert $I_L(\nu_L)$ to $I_L(\psi)$, which is related to normalized probability of an peptide bond occurring at ψ , through

$$P(\psi) = \frac{I(\psi)}{\Delta\psi \sum_{\psi} I(\psi)} \quad (3.6)$$

where $\Delta\psi$ is the resolution interval between ψ angles. $\Delta\psi$ can be shown to depend on $\Delta\nu_L$ from eq. (3.3)

$$\Delta\psi = \frac{\Delta\nu_L}{\sqrt{\nu_0^2 - (\nu_L - \nu_0)^2}} \quad (3.7)$$

Assuming the torsion potential of a peptide bond along its ψ angle coordinate is a parabolic around the optimal ψ angle, a secondary structure will give rise to Gaussian ψ angle distribution. Calculated ψ angle distributions deviating from a single Gaussian suggest the existence of multiple secondary structures. We could fit the calculated ψ angle distribution to multiple Gaussians to reveal the secondary structures. Moreover, the torsional force constants along ψ angle can also be calculated from the standard deviation σ of the Gaussian distribution that is obtained from the fitting

$$f = \frac{RT}{\sigma^2} \quad (3.8)$$

If multiple secondary structures resolved above are in equilibrium, we can calculate the Gibbs free energy landscapes (GFEL) by using the Boltzmann relationship:⁶⁰

$$\frac{P(\psi)}{P(\psi_0)} = e^{-\frac{G(\psi) - G(\psi_0)}{RT}} \quad (3.9)$$

where $P(\psi)$ and $G(\psi)$ are the probability density and Gibbs free energy of a peptide at a specific ψ angle; $P(\psi_0)$ and $G(\psi_0)$ are the probability and the Gibbs free energy of a peptide bond at the energy minimum ψ_0 .

3.2.5 Resonance enhanced protein side chain vibrations

Besides peptide bond vibrations, vibrations of aromatic side chains, including tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe), as well as histidine (His), arginine (Arg) (Figure 3. 12), asparagines (Asn) and glutamine (Gln) are also resonance enhanced with UV excitation.⁶⁴⁻⁶⁸

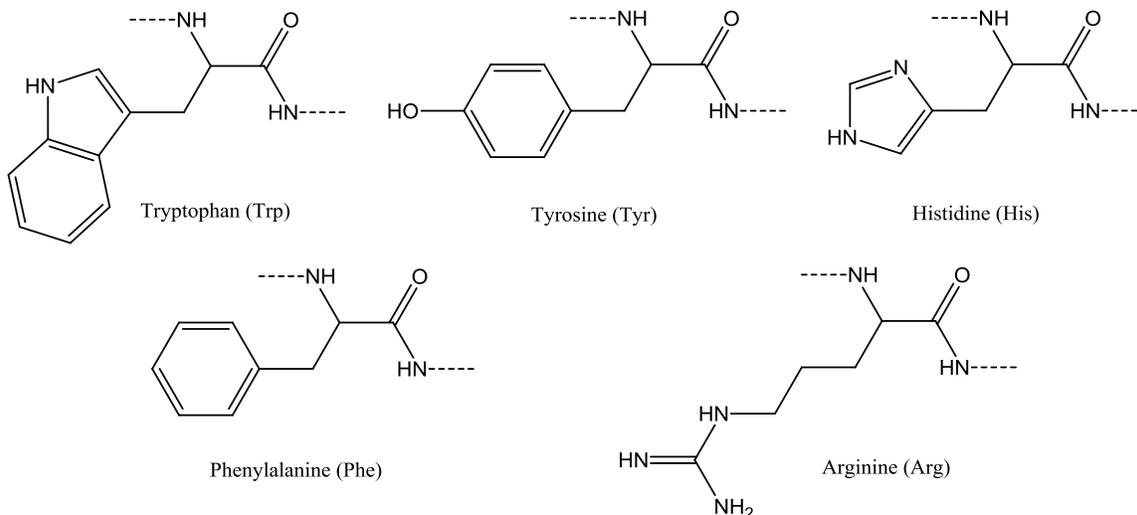


Figure 3. 12. Aromatic side chains

As demonstrated in Figure 3. 1, Trp, Tyr and Phe side chain vibrations are resonance enhanced with an excitation in their π - π^* transitions at ~ 230 nm. The π - π^* electronic transitions of these aromatic side chains can shift in different solvation environments due to the different (de)stabilization effect of solvation on ground and electronic excited states.⁶⁹ The resonance Raman scattering cross section is proportional to the square of the absorbance cross section (Section 2.2.1). Therefore, the change in absorptivity at a specific wavelength due to the electronic transition shift is manifested in Raman scattering cross section at that excitation wavelength. The UVRR intensities of these aromatic side chain vibrations can report on their environment. In folded state, these hydrophobic aromatic side chains are usually buried inside the protein; while in unfolded state, they are usually exposed to water. Therefore, the UVRR

spectra of these aromatic side chains can be correlated to their water exposure which depends upon the tertiary structure of a protein.

For example, the redshift of π - π^* transition in hydrophobic environment significantly increases the 229 nm excited Raman cross section for Trp and Tyr vibrations.⁷⁰ The Raman cross section that has been related with the solvent accessible surface area, was applied to the study of myoglobin acid denaturation⁴⁶ and Trp-cage thermal unfolding.⁷¹ A similar relationship between Phe vibration UVRR intensity and environment hydrophobicity has been reported⁷² and used to monitor lysozyme fibrillation.⁷³

The N-H on the indole ring of Trp and the O-H on the phenol group of Tyr can serve as H-bonding donors. Resonance enhanced vibrations with large contributions from the N-H bending or O-H bending can be used to report on the H-bonding in which these side chains participate.^{72,74-77}

The two nitrogen atoms on the imidazole ring in His side chains are often found to be involved in coordination. The C₄=C₅ stretching⁷⁷ has been used in the study of metal ion binding proteins including Cu, Zn superoxide dismutase,⁷⁸⁻⁷⁹ Zn finger peptide⁸⁰ and prion protein.⁸¹

Due to the electron delocalization on trigonal planar guanidinium group, Arg,⁸² is resonance enhanced at ~ 200 nm excitation. The primary amide side chain, Asn and Gln, are also resonance enhanced below ~ 200 nm excitation.

3.3 BIBLIOGRAPHY

- (1) Mirny, L.; Shakhnovich, E. *Annu. Rev. Biophys. Biomolec. Struct.* **2001**, *30*, 361.

- (2) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. *J. Phys. Chem. Lett.* **2011**, *2*, 334.
- (3) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. *Chem. Rev.* **2012**, *112*, 2604.
- (4) Asher, S. A. *Anal. Chem.* **1993**, *65*, A59.
- (5) Asher, S. A. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Griffiths, P. R., Eds.; John Wiley & Sons Ltd.: Chichester, 2002.
- (6) Asher, S. A. *Anal. Chem.* **1993**, *65*, A201.
- (7) Sanchez, K. M.; Neary, T. J.; Kim, J. E. *J. Phys. Chem. B* **2008**, *112*, 9507.
- (8) Halsey, C. M.; Xiong, J.; Oshokoya, O. O.; Johnson, J. A.; Shinde, S.; Beatty, J. T.; Ghirlanda, G.; JiJi, R. D.; Cooley, J. W. *Chembiochem* **2011**, *12*, 2125.
- (9) Okishio, N.; Tanaka, T.; Fukuda, R.; Nagai, M. *Biochemistry* **2002**, *42*, 208.
- (10) Wen, Z. Q.; Armstrong, A.; Thomas, G. J. *Biochemistry* **1999**, *38*, 3148.
- (11) Wojtuszewski, K.; Mukerji, I. *Protein Sci.* **2004**, *13*, 2416.
- (12) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 8074.
- (13) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *Biophys. J.* **1999**, *76*, A119.
- (14) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **2001**, *123*, 2388.
- (15) Balakrishnan, G.; Hu, Y.; Case, M. A.; Spiro, T. G. *J. Phys. Chem. B* **2006**, *110*, 19877.
- (16) JiJi, R. D.; Balakrishnan, G.; Hu, Y.; Spiro, T. G. *Biochemistry* **2006**, *45*, 34.
- (17) Mikhonin, A. V.; Asher, S. A.; Bykov, S. V.; Murza, A. *J. Phys. Chem. B* **2007**, *111*, 3280.
- (18) Balakrishnan, G.; Hu, Y.; Bender, G. M.; Getahun, Z.; DeGrado, W. F.; Spiro, T. G. *J. Am. Chem. Soc.* **2007**, *129*, 12801.
- (19) Palmer, K. F.; Williams, D. *J. Opt. Soc. Am.* **1974**, *64*, 1107.
- (20) Kou, L.; Labrie, D.; Chylek, P. *Appl. Opt.* **1993**, *32*, 3531.
- (21) Lindner, J.; Vöhringer, P.; Pshenichnikov, M. S.; Cringus, D.; Wiersma, D. A.; Mostovoy, M. *Chem. Phys. Lett.* **2006**, *421*, 329.
- (22) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. *Appl. Spectrosc.* **2005**, *59*, 1541.
- (23) Balakrishnan, G.; Hu, Y.; Spiro, T. G. *Appl. Spectrosc.* **2006**, *60*, 347.
- (24) Ballew, R. M.; Sabelko, J.; Reiner, C.; Gruebele, M. *Rev. Sci. Instrum.* **1996**, *67*, 3694.
- (25) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. *Biochemistry* **1997**, *36*, 9200.
- (26) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101–102*, 535.
- (27) Chen, X. G.; Asher, S. A.; Schweitzerstenner, R.; Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1995**, *117*, 2884.
- (28) Mayne, L. C.; Ziegler, L. D.; Hudson, B. *J. Phys. Chem.* **1985**, *89*, 3395.
- (29) Dudik, J. M.; Johnson, C. R.; Asher, S. A. *J. Phys. Chem.* **1985**, *89*, 3805.
- (30) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzerstenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775.
- (31) Chen, X. G.; Schweitzerstenner, R.; Asher, S. A.; Mirkin, N. G.; Krimm, S. *J. Phys. Chem.* **1995**, *99*, 3074.
- (32) Wang, Y.; Purrello, R.; Jordan, T.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359.
- (33) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. *J. Phys. Chem. B* **2004**, *108*, 19020.
- (34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.;

- Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2010., 2010.
- (35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; J. A. Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Lyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Gega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, G.; Pople, J. A. Gaussian, Inc., Wallingford CT, 2004., 2004.
- (36) Herrebout, W. A.; Clou, K.; Desseyn, H. O. *J. Phys. Chem. A* **2001**, *105*, 4865.
- (37) Triggs, N. E.; Valentini, J. J. *J. Phys. Chem.* **1992**, *96*, 6922.
- (38) Myshakina, N. S.; Ahmed, Z.; Asher, S. A. *J. Phys. Chem. B* **2008**, *112*, 11873.
- (39) Vanderkooi, J. M.; Dashnau, J. L.; Zelent, B. *Biochim. Biophys. Acta. - Protein Proteomics* **2005**, *1749*, 214.
- (40) Torii, H.; Tatsumi, T.; Tasumi, M. *J. Raman Spectrosc.* **1998**, *29*, 537.
- (41) Mix, G.; Schweitzer-Stenner, R.; Asher, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 9028.
- (42) Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2005**, *109*, 3047.
- (43) Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2007**, *111*, 4271.
- (44) Bykov, S. V.; Asher, S. A. *J. Phys. Chem. Lett.* **2010**, *1*, 269.
- (45) Chi, Z. H.; Chen, X. G.; Holtz, J. S. W.; Asher, S. A. *Biochemistry* **1998**, *37*, 2854.
- (46) Chi, Z. H.; Asher, S. A. *Biochemistry* **1998**, *37*, 2865.
- (47) Chi, Z. H.; Asher, S. A. *Biochemistry* **1999**, *38*, 8196.
- (48) Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433.
- (49) Shi, Z. S.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190.
- (50) Ding, L.; Chen, K.; Santini, P. A.; Shi, Z. S.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2003**, *125*, 8092.
- (51) Shi, Z. S.; Chen, K.; Liu, Z. G.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877.
- (52) Pimenov, K. V.; Bykov, S. V.; Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 2840.

- (53) Wang, Y.; Purrello, R.; Georgiou, S.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6368.
- (54) Amunson, K. E.; Kubelka, J. *J. Phys. Chem. B* **2007**, *111*, 9993.
- (55) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2006**, *110*, 1928.
- (56) Krimm, S.; Bandekar, J. *Biopolymers* **1980**, *19*, 1.
- (57) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2002**, *106*, 3391.
- (58) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2004**, *108*, 10923.
- (59) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2004**, *108*, 5438.
- (60) Mikhonin, A. V.; Myshakina, N. S.; Bykov, S. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 7712.
- (61) Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2006**, *128*, 13789.
- (62) Ma, L.; Ahmed, Z.; Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2007**, *111*, 7675.
- (63) Ma, L.; Hong, Z.; Sharma, B.; Asher, S. *J. Phys. Chem. B* **2011**, *116*, 1134.
- (64) Fodor, S. P. A.; Copeland, R. A.; Grygon, C. A.; Spiro, T. G. *J. Am. Chem. Soc.* **1989**, *111*, 5509.
- (65) Asher, S. A.; Ludwig, M.; Johnson, C. R. *J. Am. Chem. Soc.* **1986**, *108*, 3186.
- (66) Asher, S. A.; Murtaugh, J. L. *Appl. Spectrosc.* **1988**, *42*.
- (67) Sweeney, J. A.; Asher, S. A. *J. Phys. Chem.* **1990**, *94*, 4784.
- (68) Sension, R. J.; Hudson, B.; Callis, P. R. *J. Phys. Chem.* **1990**, *94*, 4015.
- (69) Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319.
- (70) Chi, Z. H.; Asher, S. A. *J. Phys. Chem. B* **1998**, *102*, 9595.
- (71) Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 10943.
- (72) Hildebrandt, P. G.; Copeland, R. A.; Spiro, T. G.; Otlewski, J.; Laskowski, M.; Prendergast, F. G. *Biochemistry* **1988**, *27*, 5426.
- (73) Shashilov, V. A.; Lednev, I. K. *J. Am. Chem. Soc.* **2008**, *130*, 309.
- (74) Miura, T.; Takeuchi, H.; Harada, I. *Biochemistry* **1988**, *27*, 88.
- (75) Miura, T.; Takeuchi, H.; Harada, I. *J. Raman Spectrosc.* **1989**, *20*, 667.
- (76) Rodgers, K. R.; Su, C.; Subramaniam, S.; Spiro, T. G. *J. Am. Chem. Soc.* **1992**, *114*, 3697.
- (77) Takeuchi, H. *Biopolymers* **2003**, *72*, 305.
- (78) Hashimoto, S.; Ohsaka, S.; Takeuchi, H.; Harada, I. *J. Am. Chem. Soc.* **1989**, *111*, 8926.
- (79) Wang, D.; Zhao, X.; Varghek, M.; Spiro, T. G. *J. Am. Chem. Soc.* **2000**, *122*, 2193.
- (80) Varghek, M.; Zhao, X.; Lai, Z.; McLendon, G. L.; Spiro, T. G. *Inorganic Chemistry* **1999**, *38*, 1372.
- (81) Miura, T.; Hori-i, A.; Mototani, H.; Takeuchi, H. *Biochemistry* **1999**, *38*, 11560.
- (82) Sension, R. J.; Hudson, B.; Callis, P. R. *J. Phys. Chem.* **1990**, *94*, 4015.

4.0 EXPERIMENTAL TECHNIQUES

In this work, UVRR spectroscopy is the main technique applied to study the peptide structure and folding dynamics. However, other techniques, such as CD spectroscopy and NMR spectroscopy have been used as complementary techniques.

4.1 UV RESONANCE RAMAN SPECTROSCOPY

4.1.1 UV resonance Raman spectroscopy instrumentation

The details about the UVRR spectroscopy instrument have been described in previous publication.¹ Figure 4. 1 shows a diagram of UVRR spectroscopy instrument consisting of a UV (and additional IR for T-jump experiment) laser light source, a sample holder, a spectrometer and a detector.

Two 204 nm UV laser light sources that are in resonance with the peptide bond first allowed $\pi \rightarrow \pi^*$ electronic transition and selectively enhance peptide backbone amides vibrations have been used in this work. (1) The 355 nm third harmonic of a Nd:YAG laser (Coherent Infinity) was anti-Stokes Raman shifted five harmonics to 204 nm in 40 psi hydrogen gas. This laser is operated at 100 Hz repetition rate with a 3 ns pulse duration. The average power is ~1 mW. (2) The third harmonic (272 nm) of a Nd:YLF (second harmonic at 527 nm) pumped

Ti:Sapphire laser (Photonic Industries, Inc.) was frequency mixed with the Ti:Sapphire fundamental (816 nm), giving rise to a 204 nm light source. This laser was operated at 1 kHz repetition rate with a 10 ns pulse duration. The average power was ~1 mW.

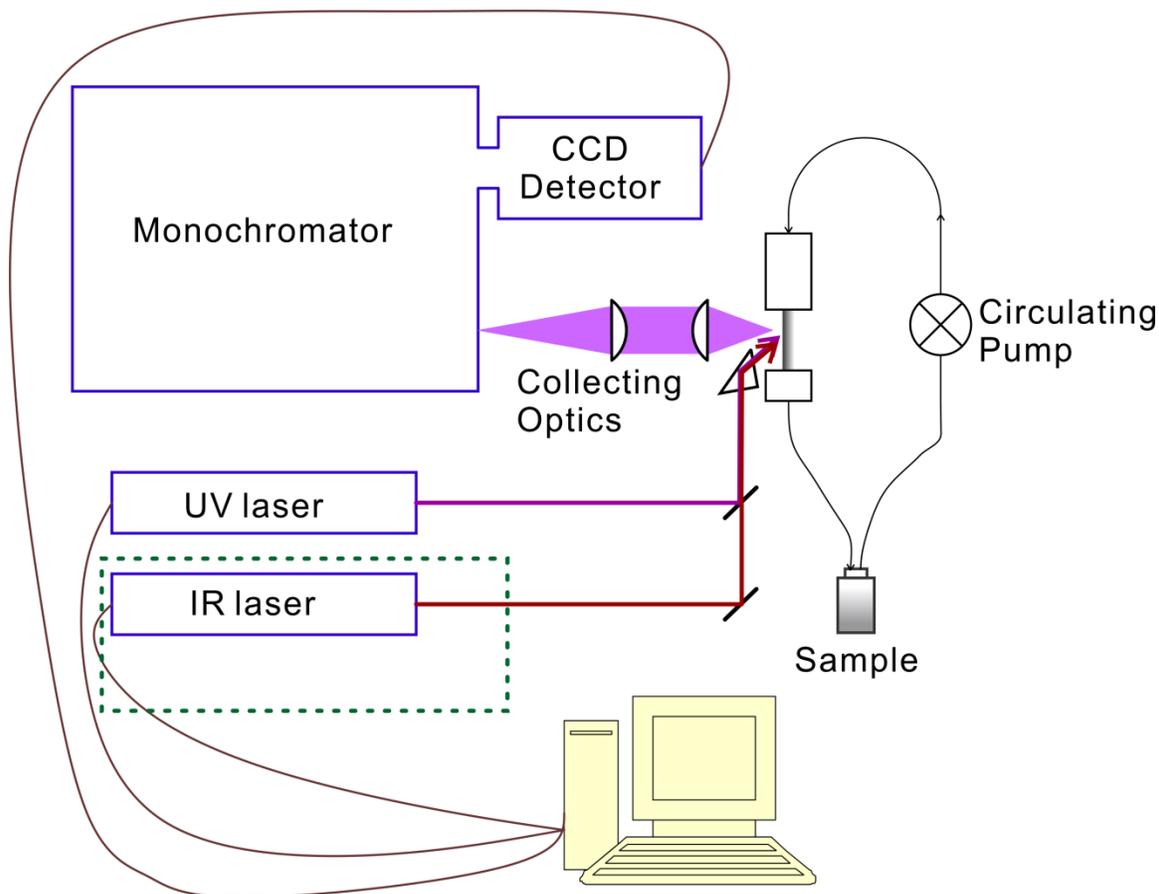


Figure 4. 1. Diagram of UVRR spectroscopy instrumentation

The excitation beam was directed to the sample. Depending on the purpose of the study and the properties of the sample, the sample can be held in different ways. (1) Solution samples can be circulated by a pump in an open flowing stream that is free of background (except air). The metal tubing is enclosed within a water-jacket to control the sample temperature. (2) For small volume samples, a spinning quartz NMR tube can be used. The quartz NMR tube would cause additional spectral background from quartz that can be subtracted. It is difficult to precisely control the sample temperature in this setup.

The back-scattered light was collected by optics and focused on a partially subtractive double monochromator that removed the Rayleigh scattering and disperses the Raman scattering signals. The dispersed signal was finally detected by a liquid nitrogen cooled, Lumogen coated back-thinned CCD camera (Princeton Instruments Spec-10:400B).

4.1.2 Calculation of Raman scattering cross section

The Raman scattering cross section of a specific band is calculated by:

$$\sigma_S^R = \frac{\sigma_{IS}^R c_{IS} I_S \eta_{IS}}{c_S I_S \eta_S} \quad (4.1)$$

where σ^R is the Raman scattering cross section, c is the concentration, I is the integrated intensity of a specific band, and η is the instrument efficiency at the frequency of a specific Raman band. The subscript S denotes the sample, while the subscript IS denotes the perchlorate internal standard. The instrument efficiencies at the sample band (η_S) and the internal standard band (η_{IS}) are assumed to be equal. The Raman scattering cross section of perchlorate has been estimated to be $0.12 \times 10^{-26} \text{ cm}^2 \cdot \text{molecule sr}^{-1}$ at 204 nm excitation by extrapolating the Raman cross section measurement of Dudik et al².

4.2 OTHER EXPERIMENTAL TECHNIQUES

4.2.1 Circular dichroism spectroscopy

A Jasco J-710 CD spectropolarimeter was used to measure the CD spectra of samples. Temperature controlled quartz cuvette with 0.2 mm, 1 mm or 10 mm path length was used

depending on the sample concentration. The CD spectra were collected at 0.2 nm data intervals and averaged over 5 or 10 scans. If necessary, the second order Savitzky-Golay method over 15 data points was applied to smooth the CD spectra. This smooth process does not affect the CD bandshapes because the band widths (> 10 nm) are usually much broader than the smooth window (3 nm).

4.2.2 Pulsed field gradient nuclear magnetic resonance spectroscopy

All diffusion coefficient measurements were performed using a Bruker Avance III 600 MHz spectrometer with a BBFO Plus probe. Temperatures were controlled to a ± 1 K accuracy using a Bruker BVT3000 temperature control system. The spectrometer was calibrated against the water self-diffusion coefficient at 25 °C. The samples were prepared in D₂O and the residual HOD resonance $\delta = 4.70$ was used as an internal chemical shift standard. Diffusion coefficients were determined using a stimulated echo pulsed field gradient pulse sequence with bipolar gradients.³⁻

⁴ To obtain the diffusion coefficient, the peak intensity vs. gradient strength data were fit to the Stejskal-Tanner equation⁵⁻⁶

$$I = I_0 e^{-4\pi^2 D_s \gamma^2 G^2 \xi^2 \left(\Xi - \frac{\xi}{3}\right)} \quad (4.2)$$

where I and I_0 are the intensities of a specific band (or the intensity at a specific chemical shift) with and without the magnetic field gradient; γ is the gyromagnetic ratio that equals 42.58 MHz·T⁻¹ for ¹H nuclei; ξ and Ξ are the pulse width and the pulse interval, D_s is the diffusion coefficient of the species in D₂O; and G is the gradient strength applied to the sample. For each sample, 16 spectra were taken with gradient strengths that varied from 0 to 50 G/cm, while the duration of the gradient ξ was held constant throughout the experiment.

To plot the results as two dimensional diffusion-ordered spectroscopy (DOSY) contours with respect to the chemical shift and diffusion coefficient, these 16 spectra were fitted to eq. (4.2) to obtain $D_s(\delta)$ and the intensity $I_0(\delta)$ at each chemical shift δ . A normal distribution of intensities was generated along the diffusion coefficient axis at each chemical shift. The peak height, center position and standard deviation of the normal distribution are labeled as $I_0(\delta)$, $D_s(\delta)$ and $\Delta D_s(\delta)$ (standard error of $D_s(\delta)$), respectively⁷

$$I(\delta, D) = I_0(\delta) e^{-\frac{(D - D_s(\delta))^2}{2\Delta D_s(\delta)^2}} \quad (4.3)$$

The integrated intensities of resonances from the same species were globally fit to obtain a single diffusion coefficient for qualitative analysis by using eq. (4.2).

4.3 MATERIALS

4.3.1 Peptides

Table 4. 1 lists the peptides studied in this work.

Table 4. 1. Peptides studied in this work

Peptide	Sequence	Manufacturer	purity
AP	A ₈ RA ₄ RA ₄ RA ₂	Anaspec, Inc	> 95%
AEP	A ₉ RA ₃ EA ₄ RA ₂	Anaspec, Inc	> 95 %
Ac-AP-NH ₂	Ac-A ₈ RA ₄ RA ₄ RA ₂ -NH ₂	Anaspec, Inc	> 95 %
XAO	Ac-X ₂ A ₇ O ₂ -NH ₂ *	Pittsburgh Peptide Facility	HPLC pure

*X: diaminobutyric acid; O: orthonine

The peptide concentrations for UVRR spectroscopy was typically $1.0 \text{ mg} \cdot \text{ml}^{-1}$. For CD spectroscopy, the peptide concentrations varied depending on the experimental requirement.

4.3.2 Purchased chemicals

Table 4. 1 lists purchased chemicals that were used in this work. All chemicals are used as received without further modification or purification.

Table 4. 2. Purchased chemical used in this work

Chemicals	Name	Manufacturer	purity
NaOH	Sodium hydroxide	Sigma-Aldrich	> 97%
HCl	Hydrochloric acid	J. T. Baker	36.5% - 38%
NaCl	Sodium chloride	Mallinckrodt	AR
NaClO ₄	Sodium perchlorate	Sigma-Aldrich	98%
Ba(ClO ₄) ₂	Barium perchlorate	Sigma-Aldrich	97%
AgClO ₄	Silver perchlorate	Sigma-Aldrich	Anhydrous, 97%
GC	Guanidinium chloride	Sigma-Aldrich	99%
EGS	Ethylguanidinium sulfate	Sigma-Aldrich	98%
Arg·HCl	L-Arginine hydrochloride	Sigma-Aldrich	98%
AAOC	N-acety-arginine methyl ester hydrochloride	Bachem	
SES	Sodium ethylsulfate	TCI	98%
SPS	Sodium pentylsulfate	Alfa Aesar	99%
SOS	Sodium octylsulfate	Alfa Aesar	99%

SDeS	Sodium decylsulfate	Alfa Aesar	99%
SDS	Sodium dodecylsulfate	Sigma-Aldrich	99%
DTAC	Dodecyltrimethylammonium chloride	Sigma-Aldrich	99%
DPC	Dodecylphosphocholine	Avanti polar lipids, Inc	99%
DDM	Dodecyl- β -D-maltoside	Avanti polar lipids, Inc	99%
CH ₃ CN	acetonitrile	Sigma-Aldrich	spectrophotometric grade
D ₂ O	Deuterium oxide	Cambridge Isotope Laboratories, Inc.	99.9%

4.3.3 Synthesized materials

Due to the high charge density of sulfate and chloride anions, EGS and AAOC are insoluble in acetonitrile. We exchanged the sulfate and chloride counter ions to perchlorate in order to make the compounds soluble.

To convert EGS and AAOC into ethylguanidinium perchlorate (EGPC) and Ac-Arg-OMe perchlorate (AAOPC), ~ 0.1 M Ba(ClO₄)₂ or ~ 0.1 M AgClO₄ was slowly added to the ~ 0.1 M EGS or ~ 0.1 M AAOC aqueous solutions to form insoluble BaSO₄ or AgCl, respectively. The suspension was centrifuged several times to remove the precipitate. The supernatant was collected and gently heated with stirring to evaporate the water. The resulting EGPC and AAOPC were very light yellow viscous ionic liquids at room temperature. The yellow color

derived from impurities. The calculated yields were 93% and 105%, respectively. The fact that the yield was greater than 100% indicated that a small amount of water remains in AAOPC.

4.4 BIBLIOGRAPHY

- (1) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. *Appl. Spectrosc.* **2005**, *59*, 1541.
- (2) Dudik, J. M.; Johnson, C. R.; Asher, S. A. *J. Chem. Phys.* **1985**, *82*, 1732.
- (3) Cotts, R. M.; Hoch, M. J. R.; Sun, T.; Markert, J. T. *J. Magn. Reson.* **1989**, *83*, 252.
- (4) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson. A* **1995**, *115*, 260.
- (5) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288.
- (6) Tanner, J. E. *J. Chem. Phys.* **1970**, *52*, 2523.
- (7) Nilsson, M.; Morris, G. A. *Anal. Chem.* **2008**, *80*, 3777.

5.0 CD AND UV RESONANCE RAMAN INDICATE LITTLE ARG-GLU SIDE CHAIN α -HELIX PEPTIDE STABILIZATION

This chapter has been published: *J. Phys. Chem. B* **2011**, *115*, 4234-4243

Authors: Zhenmin Hong, Zeeshan Ahmed⁺ and Sanford A. Asher*

Address: Department of Chemistry, University of Pittsburgh, PA 15260

⁺Current Address: Optical Technology Division, National Institute of Standards and Technology, Gaithersburgh, MD 20899

Electrostatic interactions between side chains can control the conformation and folding of peptides and proteins. We used CD and UV resonance Raman spectroscopy (UVRR) to examine the impact of side chain charge on the conformations of two 21 residue mainly polyala peptides with a few Arg and Glu residues. We expected that attractions between Arg-10 and Glu-14 side chains would stabilize the α -helix conformation compared to a peptide with an Arg-14. Surprisingly, CD suggests that the peptide with the Glu-14 is less helical. In contrast, the UVRR

show that these two peptides have similar α -helix content. We conclude that the peptide with Glu-14 has the same net α -helix content as the peptide with the Arg, but has two α -helices of shorter length. Thus, side chain interactions between Arg-10 and Glu-14 have a minor impact on α -helix stability. The thermal melting of these two peptides is similar. However the Glu-14 peptide pH induced melting forms type III turn structures that form α -helix-turn- α -helix conformations.

5.1 INTRODUCTION

For most proteins the information needed to fold into their native states is encoded in their primary sequences.¹⁻² Despite decades of studies of protein folding it is still impossible to predict the secondary structure of a protein from only its primary sequence, if an analogous sequence has not been previously characterized. The native structure is determined by amino acid side chain interactions such as hydrophobic interactions, hydrogen bonding and electrostatic interactions. The strength and long range nature of electrostatic interactions allow them to dominate side chain interactions. Side chain ion pairs are commonly found in protein crystal structures, especially within the hydrophobic cores of protein.³⁻⁷ However, the impact of these types of interactions on, for example, α -helix structures is still unclear.⁸⁻²⁴

We have been using UV resonance Raman spectroscopy (UVRR) to examine equilibrium protein and peptide solution conformations, as well as the kinetics of unfolding.²⁵⁻³⁷ We recently developed a method to determine the Gibbs free energy landscape along the Ramachandran Ψ angle folding coordinate.³⁷⁻³⁹ We also monitored the kinetics of folding of α -helical peptides and

were able to monitor both 3_{10} helix and π bulge conformations in equilibrium with the pure α -helix conformation.

In the work here we use UVRR and CD to examine the solution conformations and pH and ionic strength dependences of the conformations of two related mainly polyala peptides AP and AEP. The AP peptide sequence is $A_8\underline{R}A_3\underline{R}A_4RA_2$ while the AEP sequence is $A_8\underline{A}RA_3\underline{E}A_4RA_2$. We tested the hypothesis that the Arg-10 and Glu-14 salt bridge would form in AEP and would stabilize the α -helix conformation. This hypothesis was based on previous studies that showed the impact of salt bridges on the α -helix conformation.^{8-15,40-42} These studies indicated that oppositely charged side chains located at positions i and $i+4$ should stabilize α -helices by ~ 2 kJ mol⁻¹,⁴³ while an i and $i+3$ spacing should show a significantly smaller stabilization energy; while an i and $i+5$ spacing should show negligible stabilization. For titratable side chains, the stabilization will also depend upon pH, as well as, the solution ionic strength.

Surprisingly, we find that the Glu substitution in AEP does not enhance α -helix stability. Instead, it results in a similar α -helix content as in AP, but shifts the α -helix length distribution towards more numerous, but smaller length α -helix segments; AEP, thus, has an increased concentration of α -helix-turn- α -helix conformations.

5.2 EXPERIMENTAL SECTION

5.2.1 Materials

The AP peptide ($A_8RA_4RA_4RA_2$, > 95% purity) and the AEP peptide ($A_9RA_3EA_4RA_2$, > 95% purity) were synthesized by Anaspec Inc. Sodium hydroxide (97%+) was purchased from Sigma, hydrochloric acid (36.5%-38%) was purchased from J.T. Baker and sodium chloride (AR) was purchased from Mallinckrodt. All chemicals were used as received. The sample pH was adjusted by adding HCl or NaOH solution. The volume added was less than 1% of the sample solution.

5.2.2 Circular dichroism measurements

CD spectra were measured by using a Jasco J-710 spectrometer. We used temperature-controlled quartz cuvette of 0.2 mm (for 1.0 mg/ml samples) or 1 mm (for 0.2 mg/ml samples) path length. The CD spectra shown are the average of 10 scans.

5.2.3 UV resonance Raman spectra

The UVRR instrument was described in detail elsewhere.⁴⁴ The third harmonic of a Nd:YAG laser (Coherent Infinity) was anti-Stokes Raman shifted five harmonics in 40 psi hydrogen gas to 204 nm, which is in resonance with the peptide bond first allowed $\pi \rightarrow \pi^*$ transition. The backscattered Raman light was collected and dispersed with a partially subtractive double monochromator and detected by a Princeton Instruments Spec-10 CCD.

5.3 RESULTS AND DISCUSSION

5.3.1 Temperature dependence of conformational transitions

Figure 5.1a shows the temperature dependence of the CD spectra of AEP at pH 7.0. The 4 °C spectrum shows two minima at ~222 nm and ~205 nm, which derive from the α -helix conformation.⁴⁵ These spectra clearly indicate a dominating α -helix conformation at 4 °C. As the temperature increases the 222 nm trough becomes less evident, while the 205 nm minimum shifts to shorter wavelength due to α -helix melting. The temperature dependent spectra up to 50 °C show an isodichroic point at 202.4 nm, indicating that up to 50 °C the transition appears spectroscopically two-state. In contrast, the 60 – 70 °C spectra deviate from the isodichroic point, indicating formation of additional conformations. The high temperature AEP CD spectra look identical to those of AP, for which we have clearly demonstrated melting to a PPII-like conformation.³²

The CD spectra of samples cooled from 70 °C to 20 °C and 4 °C, show somewhat less negative ellipticities and troughs that are red shifted by ~1 nm compared to the original 20 °C and 4 °C CD spectra. This suggests irreversible sample alterations due to the high temperature exposure.

The magnitude of the 222 nm CD band that is characteristic of α -helix conformations can be directly related to the fraction of α -helix-like conformations (f_H).

$$f_H = \frac{\theta - \theta_{\text{PPII}}}{\theta_\alpha - \theta_{\text{PPII}}} \quad (1)$$

where θ is the mean residue ellipticity at 222 nm, and θ_α and θ_{PPII} are the mean residue ellipticities at 222 nm of the α -helix and PPII-like conformations. Our previous AP study²⁹

showed that θ_α and θ_{PPII} are $-26000 \pm 1000 \text{ deg cm}^2 \text{ dmol}^{-1}$ and $-3200 \pm 600 \text{ deg cm}^2 \text{ dmol}^{-1}$, respectively.

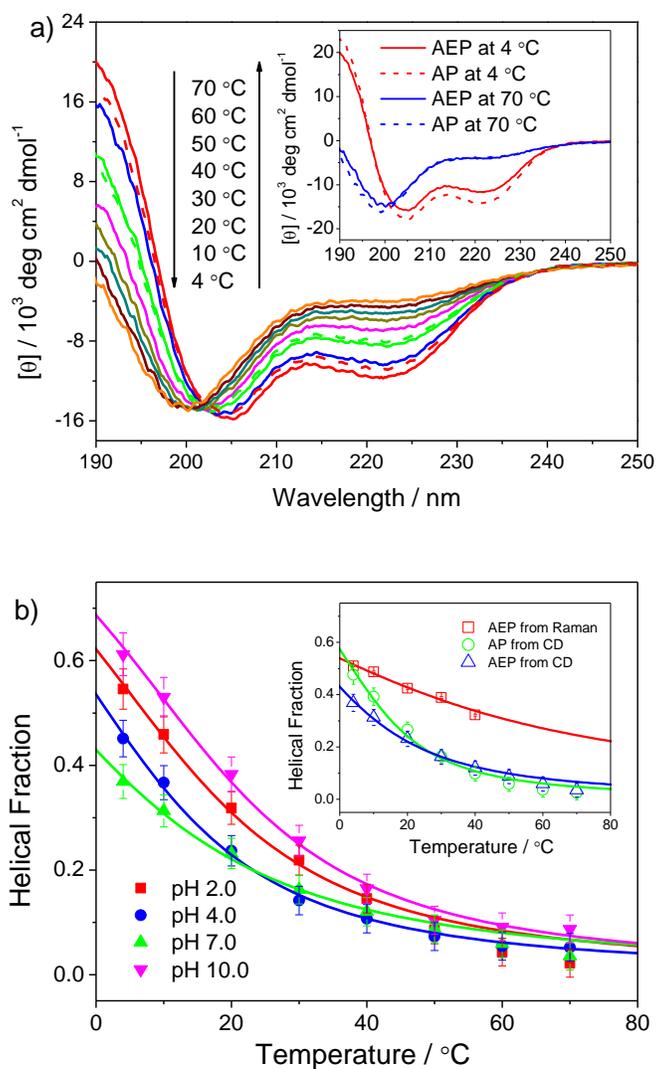


Figure 5. 1. a) Temperature dependence of CD spectra of AEP at pH 7.0. The solid lines indicate the heating process, the dashed lines indicate the spectra measured after cooling process. Inset compares CD spectra of AP and AEP. b) Temperature dependent α -helical fraction and melting curves of AEP at different pH. The α -helical fractions are calculated from the CD spectra by using eq. (1). The melting curve is the fit of the temperature dependent α -helical fraction by using eq. (3). The inset shows the AEP pH 7.0 melting curves, calculated from CD and Raman spectra, and the AP pH 7.0 melting curve calculated from the CD spectra.

We calculated the Zimm-Bragg model nucleation parameter σ and propagation parameter s from the Figure 5.1 melting data. σ is related to the likelihood of initiating an α -helical residue from a non- α -helix state, and in the typical model is considered temperature independent. s is the ratio of the α -helix partition function to the non- α -helix partition function, and is related to the enthalpy and entropy. If the partition function for the non- α -helix conformation is assigned a value of 1,

$$s = \exp\left(-\frac{\Delta h}{RT} + \frac{\Delta s}{R}\right) \quad (2)$$

where Δh and Δs are the enthalpy and entropy changes per residue, respectively, for the peptide bond to adopt an α -helix compared to a non- α -helix state. Statistical treatment with appropriate approximations gives the following quantitative melting curve:⁴⁶⁻⁴⁷

$$f_H = \frac{s}{(1+s) + \sqrt{(1-s)^2 + 4\sigma s}} \left(1 + \frac{(s-1) + 2\sigma}{\sqrt{(1-s)^2 + 4\sigma s}}\right) \quad (3)$$

Assuming that the α -helix is stabilized by interresidue hydrogen bonding, the thermodynamically least favored α -helix conformation would have a span of four α -helix residues. A four-residue α -helix span is too short to allow interresidue hydrogen bonding. Thus, this four α -helical residue conformation suffers an entropic penalty of $4\Delta s$ without any enthalpic hydrogen bonding compensation. Based on this argument, the temperature independent σ is calculated as

$$\sigma = \exp\left(\frac{4\Delta s}{R}\right) \quad (4)$$

The σ values obtained from the fitting our AEP data are comparable to those of previous studies⁴⁸⁻⁴⁹ but somewhat larger than values obtained from the host-guest methods.⁵⁰⁻⁵¹ It should

be noted that these host-guest method results have been questioned because poly(hydroxybutyl- or hydroxypropyl)-L-glutamine hosts strongly favor formation of long α -helices and thus distort the estimation of σ values of the guest residues.^{48-49,52}

When $s = 1$, T_m , the midpoint of the melting curve is:

$$T_m = \frac{\Delta h}{\Delta s} \quad (5)$$

Figure 5.1b shows the best fits of the temperature dependence of the calculated α -helix fractions between 4 °C to 50 °C. The 60 °C and 70 °C data were not included because of the irreversible changes found upon high temperatures incubation. The fitted thermodynamics parameters Δh , Δs and T_m are listed in Table 5. 1. These Δh and Δs values are close to those found in earlier studies,⁵³⁻⁵⁵ but are significantly smaller than those found recently.^{31,37,48} Also, the T_m values obtained from these fits are systematically smaller.^{31,37,48} This discrepancy might arise from the previous model which only permitted a single α -helix segment per peptide, which was thought to be reasonable for these short peptides. This approximation, as discussed below, appears inappropriate for our AEP peptide.

The hydrogen bonding enthalpy is expected to dominate Δh , but a relatively small hydrophobic contribution also occurs. A peptide chain of m α -helix segments of segment length L_m results in a total number of hydrogen bonds of $\Sigma L_m - 4m$, since each α -helix segment forms $L_m - 4$ hydrogen bonds. The zipper model assumes only one helix segment per peptide such that all helical peptide bonds are counted within one segment. This over counts the number of hydrogen bonds as $\Sigma L_m - 4$. As a consequence, Δh is overestimated by $4(m-1)$ hydrogen bonds enthalpies.⁵⁶ The zipper model may also overestimate Δs , since the entropy of α -helical terminal residues is somewhat greater than those of the central residues. The overestimation of Δs , however, is likely

to be smaller than the overestimation of Δh ; this may bias the zipper model calculated T_m towards higher temperature.

Table 5. 1. Thermodynamic parameters obtained by Zimm-Bragg model fitting of the CD melting curves

Peptide	Ionic strength	pH	$\Delta h / \text{kJ mol}^{-1}$	$\Delta s / \text{J mol}^{-1} \text{K}^{-1}$	$T_m / \text{°C}$	R^2
AP	0 M NaCl	7.0	-3.19 ± 0.05	-11.5 ± 0.2	4	0.985
AEP	0 M NaCl	2.0	-3.10 ± 0.03	-11.0 ± 0.1	9	0.994
		4.0	-3.12 ± 0.03	-11.3 ± 0.1	3	0.996
		7.0	-2.82 ± 0.04	-10.5 ± 0.1	-5	0.993
		10.0	-3.15 ± 0.03	-11.1 ± 0.1	11	0.994
	1.0 M NaCl	2.0	-3.18 ± 0.04	-11.0 ± 0.1	16	0.992
		4.0	-3.20 ± 0.03	-11.2 ± 0.1	13	0.995
		7.0	-3.08 ± 0.05	-10.9 ± 0.2	9	0.985
		10.0	-3.08 ± 0.08	-11.0 ± 0.3	7	0.969

Figure 5.2a shows the temperature dependence of the UVRR spectra of AEP between 4 °C to 40 °C at pH 7.0. These spectra are very similar to those of AP (Figure 5.3). At 4 °C, the broad band at $\sim 1650 \text{ cm}^{-1}$ is composed of at least two overlapping bands: the AmI band at $\sim 1661 \text{ cm}^{-1}$, and a 1645 cm^{-1} band which arises from the CN_3 asymmetric stretch of the guanidinium group of the Arg side chains. The AmII band occurs at $\sim 1556 \text{ cm}^{-1}$. The $\text{C}_\alpha\text{-H}$ b bands occur at $\sim 1367 \text{ cm}^{-1}$ and $\sim 1392 \text{ cm}^{-1}$. The intensities of the $\text{C}_\alpha\text{-H}$ b bands are inversely proportional to the α -helical fraction.^{30,36,57} The AmIII₁ band occurs at $\sim 1338 \text{ cm}^{-1}$, the AmIII₂ band occurs at $\sim 1307 \text{ cm}^{-1}$, and the AmIII₃ band occurs at $\sim 1256 \text{ cm}^{-1}$. The $\sim 1170 \text{ cm}^{-1}$ band consists of multiple contributions and involves NH_2 rocking of the Arg side chain guanidinium group, which makes

this band sensitive to the environment, especially hydrogen bonding. Thus, this band is potentially useful for identifying the environment of Arg side chains.

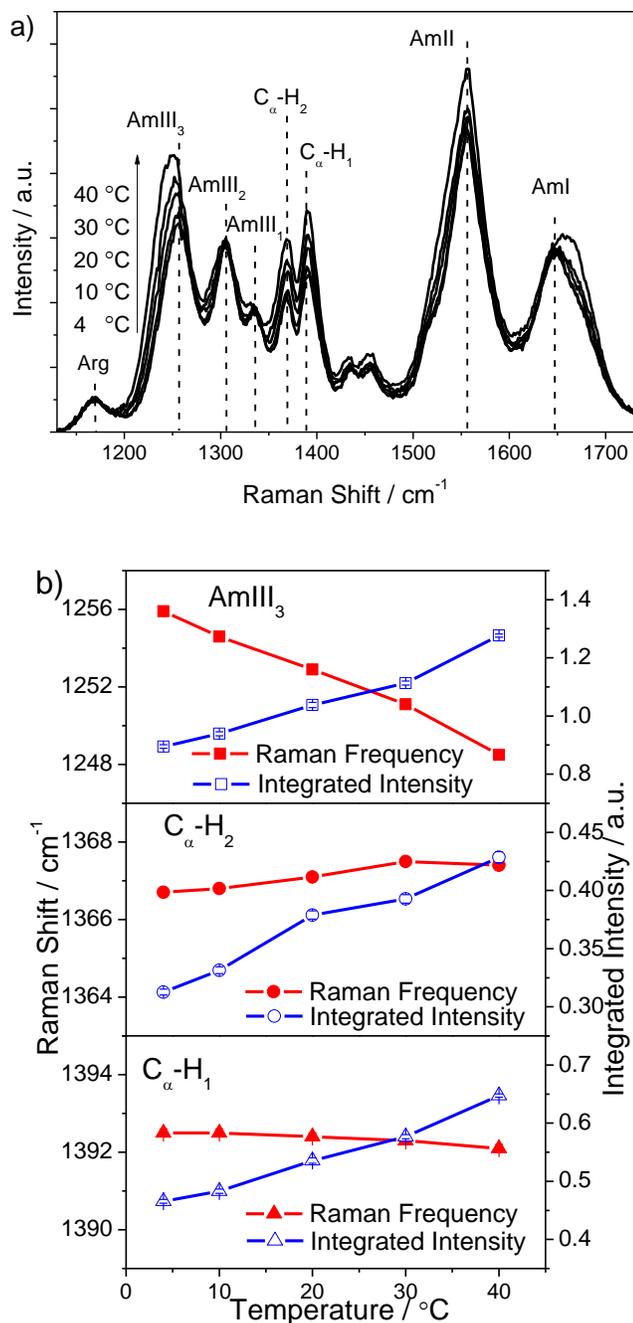


Figure 5. 2. a) Temperature dependence of UVRR spectra of AEP at pH 7.0. b) AmIII₃ and C_α-H band frequencies and integrated intensities from deconvolution of UVRR spectra. The error bars shown are standard deviations from the deconvolution.

The Figure 5.2a spectra show that as the temperature increases, the two C_{α} -H bands become more intense, due to melting of the α -helix and formation of PPII-like conformations (Figure 5.2b). The AmIII₃ band downshifts and its intensity increases. In contrast, the AmIII₁ band becomes weaker, whereas the AmIII₂ band shows little change. The similarity of the AEP and AP UVRR temperature dependence allows us to ascribe the observed AEP spectral changes to the melting of the α -helix to PPII-like conformations.

5.3.2 Comparison between AEP and AP

The CD spectra calculated α -helical fraction, f_H of the 4 °C pH 7.0 AEP sample (Figure 5.1) is smaller (0.37 ± 0.03) than that of AP (0.48 ± 0.04). In contrast, the Figure 5.3 UVRR spectra of AEP and AP in the Am III and C_{α} -H b regions are essentially identical, indicating identical Ramachandran ψ angle distributions of AP and AEP. This indicates identical conformational distributions and identical helical fractions, in contrast to the CD results.

This disagreement results from the different mechanisms of CD and Raman spectroscopy. The CD phenomenon results from the coupling of electric and magnetic dipole transition moments whose values increase supralinearly with the α -helix length.⁵⁸⁻⁶⁰ Thus, the CD signal per peptide bond is stronger for long α -helices, and weaker for shorter helices. We previously estimated that the mean residue ellipticity at 222 nm changed from $\sim -3000 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ for a fully unfolded peptide to $\sim -4000 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ when the peptide adopts only a single α -helix turn.⁶¹

In contrast to CD, UVRR of peptide bonds show a more linear intensity dependence since each peptide bond independently contributes to the AmIII₃ and the C_{α} -H b band intensities.^{35,62-63}

The only nonlinear dependence of the UVRR band peptide bond cross sections results from a bias against long α -helices due to the decrease in electronic transition oscillator strength because of the hypochromism which results from excitonic interactions between the transition dipoles of adjacent peptide bonds in the α -helix conformation.⁶⁴⁻⁶⁵ Since the Raman scattering cross section is roughly proportional to the square of the absorptivity, the α -helix Raman cross sections per α -helix peptide bonds will decrease for longer helices that have more excitonic interactions.

Considering the CD and Raman results together, we conclude that AEP has essentially the same α -helix fraction as AP, but differs in that it adopts more α -helix segments of shorter length than does AP; molecular dynamics simulation of polyala peptides found that α -helix-turn- α -helix conformations are common.^{56,66-67}

Recent MD simulation by Zhang et al. showed that the Fs peptides (which has the same sequence as AP, but is capped by acetyl and amide groups) forms a single α -helix and α -helix-turn- α -helix conformation.⁶⁷ Due to the lower intrinsic α -helical propensity of the AEP Glu-14 than that of Arg-14 in AP,⁶⁸ it is likely that AEP forms more α -helix-turn- α -helix conformations, with the turn structures occurring around Glu-14. The α -helix segment between Glu-14 and C-terminus is probably too short to contribute much of a CD signature, whereas this short segment contributes well to the UVRR.

AEP has two Arg compared to the three of AP. Thus, unless there is a conformational dependence of the Arg Raman intensities the $\sim 1170\text{ cm}^{-1}$ Arg band intensity should decrease by a third in AEP. However, Figure 5.3 shows that the AEP Arg band intensity decreases by only 20 percent and its frequency downshifts in AEP from that in AP demonstrating that the Arg Raman cross section and band frequencies depend upon the side chain environment and peptide conformation.

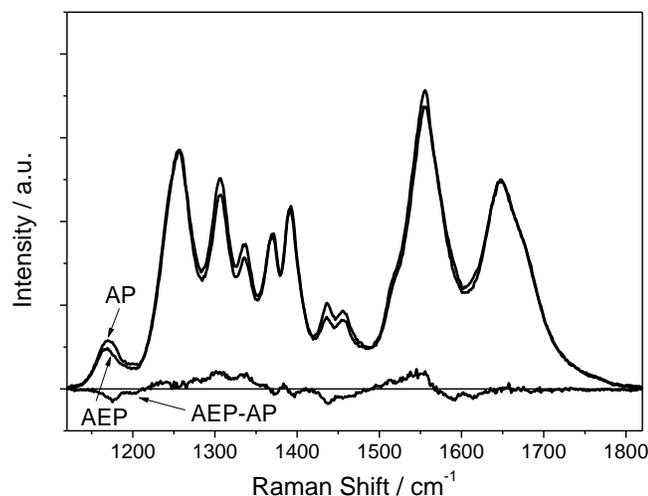


Figure 5.3. UV Resonance Raman spectra of AEP and AP and their difference spectrum at 4 °C, pH 7.0.

The $\sim 1170\text{ cm}^{-1}$ Arg band appears to consist of, at least, two components (Figure 5.3), which presumably result from Arg side chains in different environments. One component probably arises from Arg side chains completely accessible to water. A fully hydrated Arg would occur, for example, in the extended PPII-like conformation. The other component probably results from a partially hydrated state where the Arg side chain occurs in a more folded state, such as in an α -helix-like conformation. In the α -helix-like conformation, the Arg side chain would be partially shielded from water by the backbone, preventing full hydrogen bonding between water and the guanidinium group. In addition, in the α -helix, the peptide bond carbonyls are intramolecularly hydrogen bonded to the peptide bond NH groups, which prevent hydrogen bonding between the guanidinium and peptide bond carbonyls.

5.3.3 pH and ionic strength dependence of AEP conformations

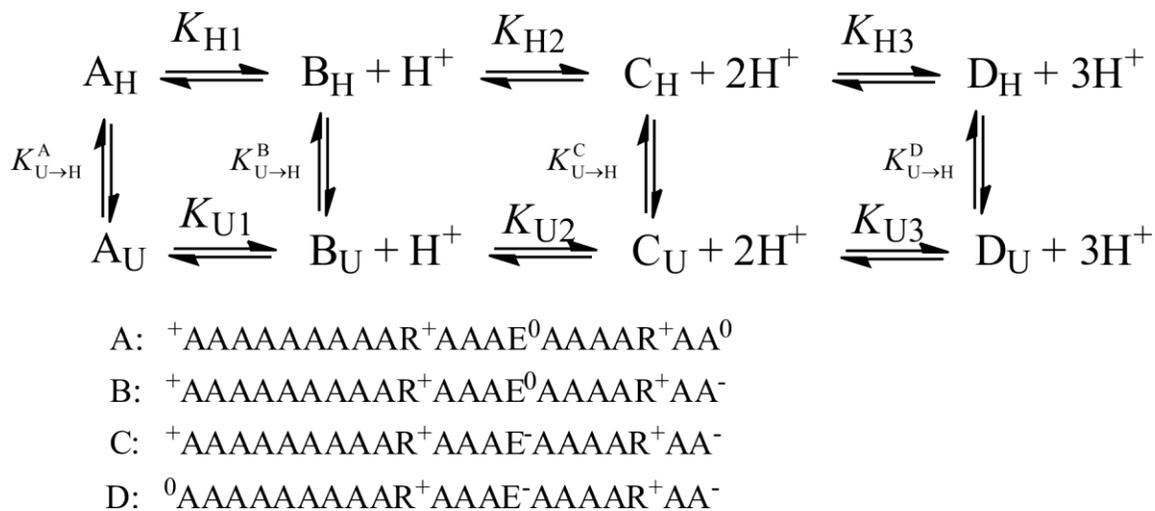
The helix stability is determined by the helix propensities of the amino acids and the possible interactions between the side chains and the macrodipole.⁶⁸ the Gibbs free energy change for the conversion of the unfolded conformation to the α -helix conformation (U \rightarrow H),

$$\Delta G_{U \rightarrow H} = \sum_{AA} \Delta G_{U \rightarrow H}^{AA} + \Delta G_{U \rightarrow H}^{ND} + \Delta G_{U \rightarrow H}^{CD} + \sum_{SD} \Delta G_{U \rightarrow H}^{SD} + \Delta G_{U \rightarrow H}^{CR} + \Delta G_{U \rightarrow H}^{SB} + \sum_{SS} \Delta G_{U \rightarrow H}^{SS} \quad (6)$$

$\sum_{AA} \Delta G_{U \rightarrow H}^{AA}$ is the summation of the α -helix propensity contributions to the α -helix stability of all the residues; $\Delta G_{U \rightarrow H}^{ND}$ is the contribution from the electrostatic interaction between the N-terminal amino group (α -NH₃⁺) and the α -helix macro-dipole; $\Delta G_{U \rightarrow H}^{CD}$ is the contribution from the electrostatic interaction between the C-terminal carboxylate (α -COO⁻) and the α -helix macro-dipole; $\sum_{SD} \Delta G_{U \rightarrow H}^{SD}$ is the summation of the contributions from the electrostatic interactions between charged side chains (Arg-10, Arg-19 and Glu-14, if charged) and the α -helix macro-dipole; $\Delta G_{U \rightarrow H}^{CR}$ is the contribution from the electrostatic interaction between the Arg-19 side chain guanidinium cation and the α -helix C-terminal anion; $\Delta G_{U \rightarrow H}^{SB}$ is the contribution from the possible salt bridge formation between the Arg-10 and Glu-14 side chains, including the contribution of a hydrogen bond from the Arg guanidinium to the Glu carboxylate; $\sum_{SS} \Delta G_{U \rightarrow H}^{SS}$ is the summation of contributions to the α -helix stability from side chain-side chain interactions except for the possible salt bridge.

These electrostatic interactions depend upon the solution pH and ionic strength; the charge states of AEP depend upon pH. These different charge states show different α -helix stabilizations/destabilizations. In order to study the impact of pH on α -helix stability, we

measured the pH dependence of the α -helical fraction of AEP at 4 °C by using CD and UVRR spectroscopy.



Scheme 5. 1. Species in AEP solution

The subscript H and U represent α -helix and unfolded conformations

In AEP, there are three titratable groups: the C-terminal carboxyl, the Glu-14 side chain carboxyl and the N-terminal amine. Therefore, there are four charge states in AEP. The peptide in each charge state will occur in different α -helix/unfolded conformation equilibria due to the different α -helix stabilization/destabilization. Scheme 5. 1 defines these states and the transitions between them.

The pH dependent α -helical fraction of AEP is given by (See Appendix for derivation):

$$f_H = \frac{1 + 10^{pH - pK_{H1}} + 10^{2pH - pK_{H1} - pK_{H2}} + 10^{3pH - pK_{H1} - pK_{H2} - pK_{H3}}}{(1 + 1/K_{U \rightarrow H}^A) + (1 + 1/K_{U \rightarrow H}^B) \times 10^{pH - pK_{H1}} + (1 + 1/K_{U \rightarrow H}^C) \times 10^{2pH - pK_{H1} - pK_{H2}} + (1 + 1/K_{U \rightarrow H}^D) \times 10^{3pH - pK_{H1} - pK_{H2} - pK_{H3}}} \quad (7)$$

The pK_{H1} , pK_{H2} and pK_{H3} are the pK_a of the C-terminal carboxyl, the Glu-14 side chain carboxyl and the N-terminal amine in the α -helix conformations. The α -helical formation equilibrium constants $K_{U \rightarrow H}^A$, $K_{U \rightarrow H}^B$, $K_{U \rightarrow H}^C$ and $K_{U \rightarrow H}^D$ were obtained from the CD and the UVRR data.

At pH 2.0, the C-terminal and Glu-14 side chain carboxyls are protonated and neutral, while the N-terminal amino group is protonated and cationic. This corresponds to state A in Scheme 5. 1 The concentrations of other charged states are negligible. Therefore,

$$K_{U \rightarrow H}^A = \frac{f_H(2.0)}{1 - f_H(2.0)}.$$

At pH 4.5, the C-terminal carboxyl is mainly deprotonated, while the Glu-14 side chain carboxyl and N-terminal amino group remain protonated. This corresponds to state B. Therefore,

$$K_{U \rightarrow H}^B = \frac{f_H(4.5)}{1 - f_H(4.5)}.$$

At pH 7.0, state C with two carboxylates and a protonated N-terminal amino group, is the major AEP component. Thus $K_{U \rightarrow H}^C = \frac{f_H(7.0)}{1 - f_H(7.0)}$.

At pH 11.0, state D dominates, where all titratable groups are deprotonated. Therefore,

$$K_{U \rightarrow H}^D = \frac{f_H(11.0)}{1 - f_H(11.0)}.$$

The values of these α -helical formation equilibrium constants calculated from the CD and UVRR data are listed in Table 5. 2. We fitted the experimental α -helical fraction data calculated from the CD and UVRR spectra, to eq. (6) using the corresponding measured $K_{U \rightarrow H}^A$, $K_{U \rightarrow H}^B$, $K_{U \rightarrow H}^C$ and $K_{U \rightarrow H}^D$ values (Figure 5.4a). The fitting results are listed in Table 5. 2.

In the conversion from state B to state C, the Glu-14 is deprotonated and a salt bridge may form. In state C, there may be an interaction between the Glu-14 charge and the macrodipole. The helix propensity for the state C deprotonated Glu-14 differs from that of the state B protonated Glu-14. Except for the possible salt bridge formation, the interaction between the Glu-14 and the macrodipole and the helix propensity difference, all other interactions not

involving Glu-14 should be identical for states B and C. The states B and C Gibbs free energy changes for the conversion from the unfolded conformation to the α -helix conformation can be written:

$$\Delta G_{U \rightarrow H}^B = \sum_{AA'} \Delta G_{U \rightarrow H}^{AA'} + \Delta G_{U \rightarrow H}^{E^0} + \Delta G_{U \rightarrow H}^{ND} + \Delta G_{U \rightarrow H}^{CD} + \sum_{SD'} \Delta G_{U \rightarrow H}^{SD'} + \Delta G_{U \rightarrow H}^{CR} + \sum_{SS} \Delta G_{U \rightarrow H}^{SS}$$

$$\Delta G_{U \rightarrow H}^C = \sum_{AA'} \Delta G_{U \rightarrow H}^{AA'} + \Delta G_{U \rightarrow H}^E + \Delta G_{U \rightarrow H}^{ND} + \Delta G_{U \rightarrow H}^{CD} + \sum_{SD'} \Delta G_{U \rightarrow H}^{SD'} + \Delta G_{U \rightarrow H}^{E'D} + \Delta G_{U \rightarrow H}^{CR} + \Delta G_{U \rightarrow H}^{SB} + \sum_{SS} \Delta G_{U \rightarrow H}^{SS} \quad (8)$$

Table 5. 2 Equilibrium constants obtained by fitting the CD and Raman pH dependent helical fractions with eq. (7)

α -helical formation equilibrium constant ^a	CD	Raman
$K_{U \rightarrow H}^A$	1.2 \pm 0.2	1.44 \pm 0.02
$K_{U \rightarrow H}^B$	0.68 \pm 0.10	1.04 \pm 0.02
$K_{U \rightarrow H}^C$	0.58 \pm 0.08	1.04 \pm 0.01
$K_{U \rightarrow H}^D$	1.8 \pm 0.3	1.89 \pm 0.02
pK_a^b	CD	Raman
pK_{H1}	3.4 \pm 0.2	4.0 \pm 0.2
pK_{H2}	7.0 \pm 2.0	7.7 \pm 1.9
pK_{H3}	7.4 \pm 0.2	7.4 \pm 0.3
$\Delta pK_a (pK_{H^-} - pK_U)^c$	CD	Raman
ΔpK_{a1}	0.2 \pm 0.1	0.14 \pm 0.01
ΔpK_{a2}	0.1 \pm 0.1	0.00 \pm 0.01
ΔpK_{a3}	-0.5 \pm 0.1	-0.30 \pm 0.01

^a Measured equilibrium constants, which are fixed in fitting. The errors are experimental standard deviations.

^b Errors are standard deviations from fitting.

^c Errors calculated from the errors of α -helical formation equilibrium constants.

The helix propensities of the state B protonated Glu-14 and the state C deprotonated Glu-14 are explicitly separated from those of the other residues, which are denoted as AA'. The summation over SD' denotes the interactions between the Arg-10 and Arg-19 charges and the macro-dipole; $\Delta G_{U \rightarrow H}^{E^D}$ is the contribution from the interaction between the Glu-14 charge and the macro-dipole. By taking the difference between states B and C,

$$\Delta G_{U \rightarrow H}^C - \Delta G_{U \rightarrow H}^B = \Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{E^D} + \left(\Delta G_{U \rightarrow H}^{E^-} - \Delta G_{U \rightarrow H}^{E^0} \right) \quad (9)$$

By rearranging we obtain:

$$\Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{E^D} = \left(\Delta G_{U \rightarrow H}^C - \Delta G_{U \rightarrow H}^B \right) - \left(\Delta G_{U \rightarrow H}^{E^-} - \Delta G_{U \rightarrow H}^{E^0} \right) \quad (10)$$

The two Gibbs free energy differences in the second term are expected to be 1.81 kJ mol⁻¹ for $\Delta G_{U \rightarrow H}^{E^-}$, and 0.941 kJ mol⁻¹ for $\Delta G_{U \rightarrow H}^{E^0}$ (Values for $\Delta G_{U \rightarrow H}^{E^-}$ and $\Delta G_{U \rightarrow H}^{E^0}$ are taken from Table 5.2 of ref.⁶⁸). $\Delta G_{U \rightarrow H}^B$ and $\Delta G_{U \rightarrow H}^C$ can be calculated from the measured α -helical formation equilibrium constant $K_{U \rightarrow H}^B$ and $K_{U \rightarrow H}^C$ for states B and C:

$$\begin{aligned} \Delta G_{U \rightarrow H}^B &= -RT \ln K_{U \rightarrow H}^B \\ \Delta G_{U \rightarrow H}^C &= -RT \ln K_{U \rightarrow H}^C \end{aligned} \quad (11)$$

We, thus, obtain a value $\Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{E^D} = -0.5 \pm 0.5$ kJ mol⁻¹, which indicates that the salt bridge, together with the interaction between Glu-14 charge and macro-dipole, probably has little impact on the α -helix stability.

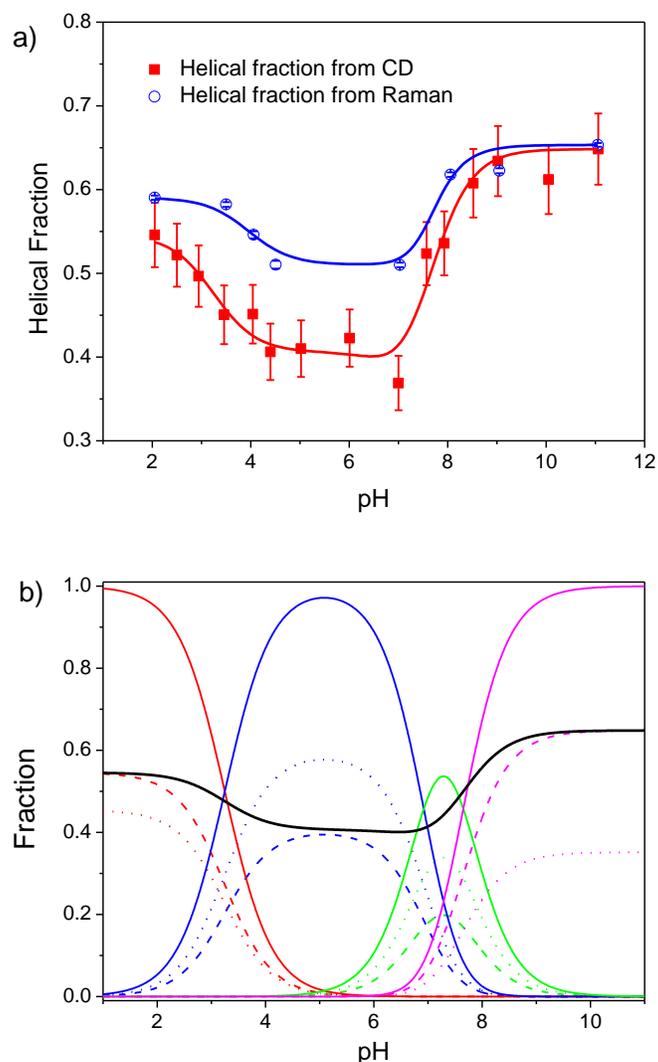


Figure 5. 4. a) pH dependence of helical fraction, f_H of AEP at 4 °C in pure water calculated from the CD and UVRR spectra via eq. (1). The curves show the fit to eq. (7). b) Calculated AEP species fractions from CD spectra (see Scheme 5. 1). Red: state A; Blue: state B, Green: state C, Magenta: state D. Dashed lines: α -helix conformation; Dotted lines: unfolded conformation; Solid lines: total AEP fraction in each state including helix and unfolded conformations. Thick solid black line: total helical fraction.

We can estimate the contribution from the interaction between the C-terminal carboxylate charge and the macro-dipole ($\Delta G_{U \rightarrow H}^{CD}$), plus the interaction between the C-terminal carboxylate charge and the Arg-19 side chain ($\Delta G_{U \rightarrow H}^{CR}$) to the α -helix stability, as well as the contribution

from the interaction between the N-terminal amino charge and the macro-dipole ($\Delta G_{U \rightarrow H}^{ND}$) to the α -helix stability.

$$\begin{aligned}\Delta G_{U \rightarrow H}^{CD} + \Delta G_{U \rightarrow H}^{CR} &= \Delta G_{U \rightarrow H}^B - \Delta G_{U \rightarrow H}^A = -RT(\ln K_{U \rightarrow H}^B - \ln K_{U \rightarrow H}^A) \\ \Delta G_{U \rightarrow H}^{ND} &= \Delta G_{U \rightarrow H}^C - \Delta G_{U \rightarrow H}^D = -RT(\ln K_{U \rightarrow H}^C - \ln K_{U \rightarrow H}^D)\end{aligned}\quad (12)$$

From above equations, we estimate that $\Delta G_{U \rightarrow H}^{CD} + \Delta G_{U \rightarrow H}^{CR} = 1.3 \pm 0.5 \text{ kJ mol}^{-1}$; $\Delta G_{U \rightarrow H}^{ND} = 2.6 \pm 0.5 \text{ kJ mol}^{-1}$. These two values are significantly larger than $\Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{ED}$, indicating that these unfavorable interactions have greater impact on the α -helix stability than does the favorable salt bridge.

These results explain the pH dependences of helical fraction: at low or high pH, AEP has a higher α -helical fraction than at neutral pH. The deprotonation of C-terminal carboxyl, which turns on the interaction between the C-terminal carboxylate charge and the macro-dipole ($\Delta G_{U \rightarrow H}^{CD}$), as well as, the interaction between the C-terminal carboxylate charge and the Arg-19 side chain ($\Delta G_{U \rightarrow H}^{CR}$), destabilize the α -helix. At high pH, the neutralization of the N-terminal amino, which eliminates the interaction between the N-terminal amino charge and the macro-dipole ($\Delta G_{U \rightarrow H}^{ND}$) stabilizes the α -helix. Because the C-terminal carboxyl and N-terminal amino are involved in the interactions that affect the α -helix stability, their pK_a values differ between the unfolded and the α -helix conformations. The pK_a differences between the α -helix conformation and the unfolded conformation ($\Delta pK_a = pK_H - pK_U$) calculated from the α -helical formation equilibrium constants are listed in Table 5.2 (see Appendix for details).

We also can calculate $\Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{ED}$, $\Delta G_{U \rightarrow H}^{CD} + \Delta G_{U \rightarrow H}^{CR}$ and $\Delta G_{U \rightarrow H}^{ND}$ from the UVRR spectra. The C_{α} -H b bands derive from the PPII-like conformations. Thus, the integrated

intensity of the C $_{\alpha}$ -H b band is proportional to the PPII fraction. By assuming a two-state transition, we can calculate the helical fraction from the UVRR spectra. At pH 7.0 and 4 °C, AEP has the identical C $_{\alpha}$ -H b band intensity as does AP (Figure 5.3), indicating identical helical fractions of ~ 0.51 ;^{29,37} at pH 7.0 and 4 °C, AEP has a PPII fraction of ~ 0.49 . We therefore calculate the PPII fractions of AEP at other pH values by the proportionality between the PPII fraction and the C $_{\alpha}$ -H b bands integrated intensity, and subsequently calculate the pH dependent helical fractions.

Figure 5.4a shows the calculated pH dependence of the AEP helical fraction calculated from the CD and Raman spectra. By applying the procedure to the Raman results, we calculated that $\Delta G_{U \rightarrow H}^{SB} + \Delta G_{U \rightarrow H}^{ED} = -0.9 \pm 0.1 \text{ kJ mol}^{-1}$, $\Delta G_{U \rightarrow H}^{CD} + \Delta G_{U \rightarrow H}^{CR} = 0.8 \pm 0.1 \text{ kJ mol}^{-1}$, and $\Delta G_{U \rightarrow H}^{ND} = 1.4 \pm 0.1 \text{ kJ mol}^{-1}$. These values are somewhat different than the values obtained from the CD spectra, due to the fact that the CD is biased towards smaller α -helix fraction as discussed above.

Table 5.1 shows our CD calculated average residue enthalpy changes (Δh) and entropy changes (Δs) (also shown in Figure 5.5b), as well as, T_m . These thermodynamic parameters, Δh and Δs are for the transition from the unfolded to the α -helix conformation at specific pH values and ionic strengths. Figure 5.5b shows that Δh and Δs at pH 7.0 are more positive than at other pH values due to the two α -helix destabilizing terminal charge-macro-dipole (C-terminal-macro-dipole and N-terminal-macro-dipole) interactions. At pH 2.0 and pH 10.0 only a single unfavorable terminal charge-macro-dipole interaction occurs.

The smaller T_m and Δh of AEP compared to those of AP at pH 7.0 can be explained by more α -helix-turn- α -helix conformations in AEP. The peptide bonds at both ends of an α -helix segment can only form single intrapeptide hydrogen bonds. Thus, the average Δh should be more

positive for α -helix-turn- α -helix conformations than for the same total number of α -helical residues within a single helix segment.

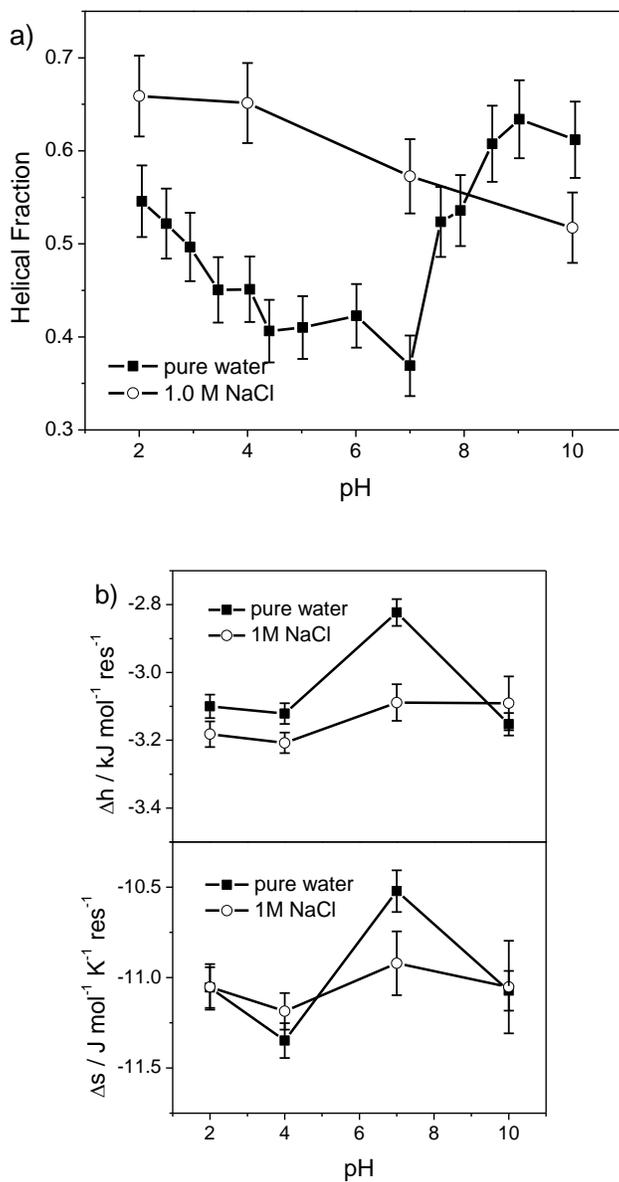


Figure 5.5. a) pH dependence of the α -helical fraction, f_H of AEP at 4 °C in pure water and 1.0 M NaCl. b) pH dependence of average enthalpy change (Δh) and entropy change (Δs) per residue for forming an α -helical peptide bond conformation at 4 °C in pure water and in 1.0 M NaCl. See Table 5.1.

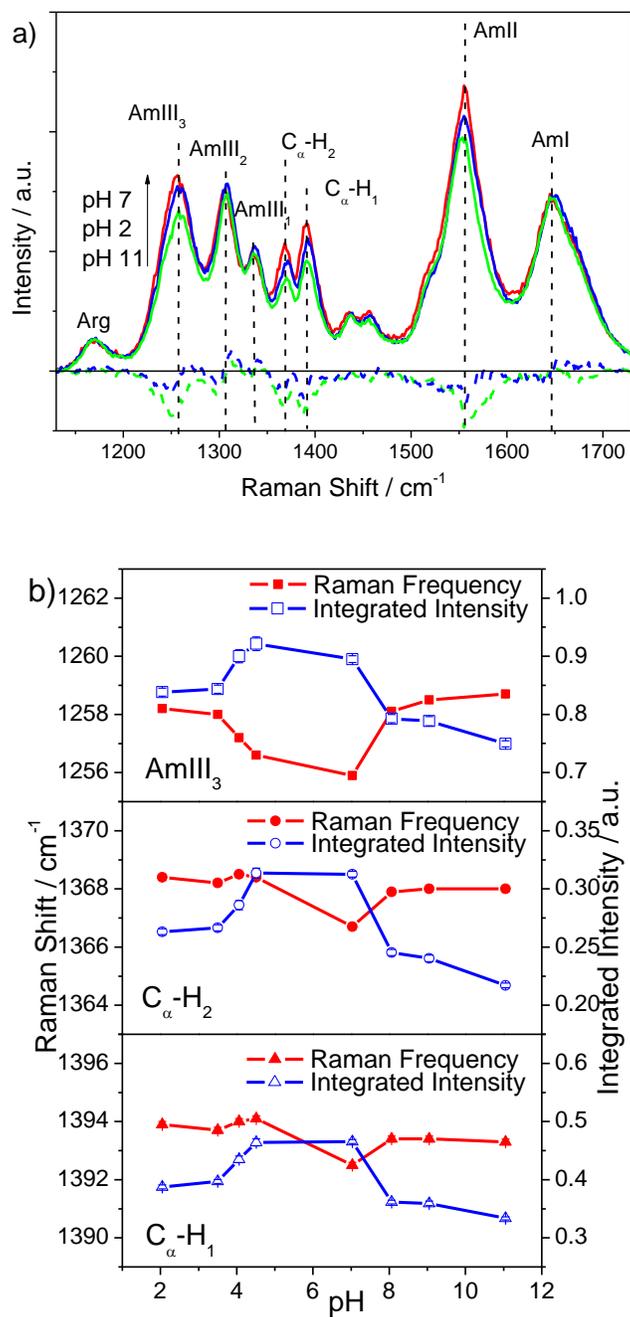


Figure 5. 6. a) pH dependence of AEP UVRR spectra at 4 °C. b) pH dependence of AmIII₃ and C_α-H band frequencies and integrated intensities from deconvolution of UVRR spectra. The error bars shown are standard deviations from the deconvolution.

Figure 5.6a shows the pH dependence of the AEP UVRR spectra at 4 °C. The intensities of the AmIII₃ and C_α-H b bands are larger at pH 7.0 than at pH 2.0 or 11.0, while the AmIII₃ frequency varies by < 4 cm⁻¹ (Figure 5.6b). The increased C_α-H b and AmIII₃ band intensities at pH 7.0 clearly signal the melting of the α-helix conformation. However, the lack of an AmIII₃ band frequency shift indicates that the melting is not to a PPII-like conformation, but the melting occurs to conformations with similar Ramachandran Ψ angles to that of the α-helix.

5.3.4 AEP forms α-helix-turn-α-helix conformation

Both the (pH 2.0 - pH 7.0) and (pH 11.0 - pH 7.0) UVRR difference spectra show a negative trough at ~1251 cm⁻¹, from which we can calculate the Ramachandran Ψ angle.³⁶ Since the hydrogen bonding status of the residues is unknown, we estimate the Ψ angle by using equations (6E) in ref.³⁶ 36 and obtain a value of -27°. A perfect type III turn has an Ψ angle of -30°. Therefore, the negative trough of 1251 cm⁻¹ suggests that at pH 7.0, the AEP α-helix partially melts to a type III turn-like structure. This type III turn-like structure is likely to occur around the Glu-14, because the helix propensity of Glu is lower than ala and Arg.⁶⁸⁻⁶⁹

Figure 5.4a compares the pH dependences of AEP helical fractions obtained from CD and UVRR. At low or high pH, the helical fractions calculated from CD and UVRR spectra are similar. At neutral pH, there is a significant discrepancy between helical fractions calculated from the CD and UVRR spectra. As discussed previously, the similar α-helix content measured by CD and UVRR at low and high pH indicates that at these pH values AEP adopts long helical segments that have large CD ellipticities. The large discrepancy at neutral pH indicates that the α-helix forms multiple short helices that only show small ellipticities and thus bias the CD

measurement towards small α -helical contents. Therefore, we conclude that at low and high pH, AEP has less turn structures and a longer α -helix segment. At neutral pH, AEP has more turn structures and shorter α -helix segments.

5.4 CONCLUSIONS

We compared the polyala peptide AP that has three Arg side chains to the peptide AEP where the Arg-14 is replaced with a Glu. We find that the AEP peptide has an increased concentration of α -helix-turn- α -helix conformations. We examined the nature of the electrostatic interactions that (de)stabilize α -helix conformations. The inter-residue interactions between the Arg-10 and Glu-14 in AEP makes only a minor contribution to the α -helix stabilization. As the temperature increases, the AEP melts from α -helix to PPII-like conformations, while the pH induced melting causes the formation of an α -helix-type III turn- α -helix conformation.

5.5 APPENDIX

5.5.1 Derivation of eq. (7)

In Scheme 5. 1, the α -helical fraction of AEP is a function of pH:

$$f_H = [A_H] + [B_H] + [C_H] + [D_H] \quad (\text{A1})$$

where $[A_H]$, $[B_H]$, $[C_H]$ and $[D_H]$ are the helical fractions in states A, B, C and D. For each charge state, the relationship between the α -helix conformation and the unfolded conformation is given by:

$$\begin{aligned} [A_U] &= \frac{[A_H]}{K_{U \rightarrow H}^A}; \\ [B_U] &= \frac{[B_H]}{K_{U \rightarrow H}^B}; \\ [C_U] &= \frac{[C_H]}{K_{U \rightarrow H}^C}; \\ [D_U] &= \frac{[D_H]}{K_{U \rightarrow H}^D}. \end{aligned} \quad (A2)$$

The peptide concentration is conserved in solution, therefore,

$$[A_H] + [A_U] + [B_H] + [B_U] + [C_H] + [C_U] + [D_H] + [D_U] = 1 \quad (A3)$$

Substituting eq. (A2) into eq. (A3), we eliminate four unknowns:

$$\left(1 + 1/K_{U \rightarrow H}^A\right)[A_H] + \left(1 + 1/K_{U \rightarrow H}^B\right)[B_H] + \left(1 + 1/K_{U \rightarrow H}^C\right)[C_H] + \left(1 + 1/K_{U \rightarrow H}^D\right)[D_H] = 1 \quad (A4)$$

The Henderson-Hasselbalch equation gives relationships between two adjacent states:

$$\begin{aligned} [B_H] &= [A_H] \times 10^{pH - pK_{H1}}; \\ [C_H] &= [B_H] \times 10^{pH - pK_{H2}}; \\ [D_H] &= [C_H] \times 10^{pH - pK_{H3}} \end{aligned} \quad (A5)$$

Combining eq. (A5) and eq. (A4), we obtain expressions for $[A_H]$, $[B_H]$, $[C_H]$ and $[D_H]$:

$$\begin{aligned} [A_H] &= \frac{1}{\left(1 + 1/K_{U \rightarrow H}^A\right) + \left(1 + 1/K_{U \rightarrow H}^B\right) \times 10^{pH - pK_{H1}} + \left(1 + 1/K_{U \rightarrow H}^C\right) \times 10^{2pH - pK_{H1} - pK_{H2}} + \left(1 + 1/K_{U \rightarrow H}^D\right) \times 10^{3pH - pK_{H1} - pK_{H2} - pK_{H3}}} \\ [B_H] &= \frac{10^{pH - pK_{H1}}}{\left(1 + 1/K_{U \rightarrow H}^A\right) + \left(1 + 1/K_{U \rightarrow H}^B\right) \times 10^{pH - pK_{H1}} + \left(1 + 1/K_{U \rightarrow H}^C\right) \times 10^{2pH - pK_{H1} - pK_{H2}} + \left(1 + 1/K_{U \rightarrow H}^D\right) \times 10^{3pH - pK_{H1} - pK_{H2} - pK_{H3}}} \end{aligned}$$

$$[C_H] = \frac{10^{2pH-pK_{H1}-pK_{H2}}}{(1+1/K_{U \rightarrow H}^A) + (1+1/K_{U \rightarrow H}^B) \times 10^{pH-pK_{H1}} + (1+1/K_{U \rightarrow H}^C) \times 10^{2pH-pK_{H1}-pK_{H2}} + (1+1/K_{U \rightarrow H}^D) \times 10^{3pH-pK_{H1}-pK_{H2}-pK_{H3}}}$$

$$[D_H] = \frac{10^{3pH-pK_{H1}-pK_{H2}-pK_{H3}}}{(1+1/K_{U \rightarrow H}^A) + (1+1/K_{U \rightarrow H}^B) \times 10^{pH-pK_{H1}} + (1+1/K_{U \rightarrow H}^C) \times 10^{2pH-pK_{H1}-pK_{H2}} + (1+1/K_{U \rightarrow H}^D) \times 10^{3pH-pK_{H1}-pK_{H2}-pK_{H3}}} \quad (A6)$$

Therefore, explicitly,

$$f_H = \frac{1 + 10^{pH-pK_{H1}} + 10^{2pH-pK_{H1}-pK_{H2}} + 10^{3pH-pK_{H1}-pK_{H2}-pK_{H3}}}{(1+1/K_{U \rightarrow H}^A) + (1+1/K_{U \rightarrow H}^B) \times 10^{pH-pK_{H1}} + (1+1/K_{U \rightarrow H}^C) \times 10^{2pH-pK_{H1}-pK_{H2}} + (1+1/K_{U \rightarrow H}^D) \times 10^{3pH-pK_{H1}-pK_{H2}-pK_{H3}}} \quad (A7)$$

5.5.2 Calculation of pKa difference (ΔpK_a) in Table 5.2

The calculations of ΔpK_{a1} , ΔpK_{a2} and ΔpK_{a3} are similar. Here we only give the calculation of ΔpK_{a1} as an example. Consider the closed thermodynamic cycle of states A and B. The fraction of the unfolded conformation in state B can be calculated:

$$[B_U] = \frac{[B_H]}{K_{U \rightarrow H}^B} = \frac{[A_H] \times 10^{-pK_{H1}}}{[H^+] \times K_{U \rightarrow H}^B} \quad (A8)$$

However,

$$[B_U] = \frac{[A_U]}{[H^+]} \times 10^{-pK_{U1}} = \frac{[A_H] \times 10^{-pK_{U1}}}{[H^+] \times K_{U \rightarrow H}^A} \quad (A9)$$

Therefore,

$$\frac{[A_H] \times 10^{-pK_{H1}}}{[H^+] \times K_{U \rightarrow H}^B} = \frac{[A_H] \times 10^{-pK_{U1}}}{[H^+] \times K_{U \rightarrow H}^A} \quad (A10)$$

Eliminating $[A_H]$ and $[H^+]$ and rearranging eq. (A10):

$$\frac{10^{-pK_{H1}}}{10^{-pK_{U1}}} = \frac{K_{U \rightarrow H}^B}{K_{U \rightarrow H}^A} \quad (A11)$$

$$10^{-(pK_{H1}-pK_{U1})} = \frac{K_{U \rightarrow H}^B}{K_{U \rightarrow H}^A} \quad (A12)$$

Substitute the definition of pK_a difference $\Delta pK_{a1} = pK_{H1} - pK_{U1}$ into eq. (A12):

$$\Delta pK_{a1} = \log_{10} K_{U \rightarrow H}^A - \log_{10} K_{U \rightarrow H}^B \quad (\text{A13})$$

For $f(x) = \log_{10} x$, the uncertainty is $df(x) = \frac{dx}{2.303x}$, therefore the uncertainty of ΔpK_{a1} is:

$$d(\Delta pK_{a1}) = \sqrt{\left(\frac{d(K_{U \rightarrow H}^A)}{2.303K_{U \rightarrow H}^A}\right)^2 + \left(\frac{d(K_{U \rightarrow H}^B)}{2.303K_{U \rightarrow H}^B}\right)^2} \quad (\text{A14})$$

The calculations of ΔpK_{a2} and ΔpK_{a3} can be achieved by examining the closed thermodynamic cycles of states B and C and of states C and D.

5.6 ACKNOWLEDGEMENT

We thank Dr. Bhavya Sharma, Lu Ma and Dr. Edward Gooding for their helpful discussions.

This work is supported by NIH Grant 5RO1 EB002053 and NIH 1RO1 EB009089.

5.7 SUPPORTING INFORMATION

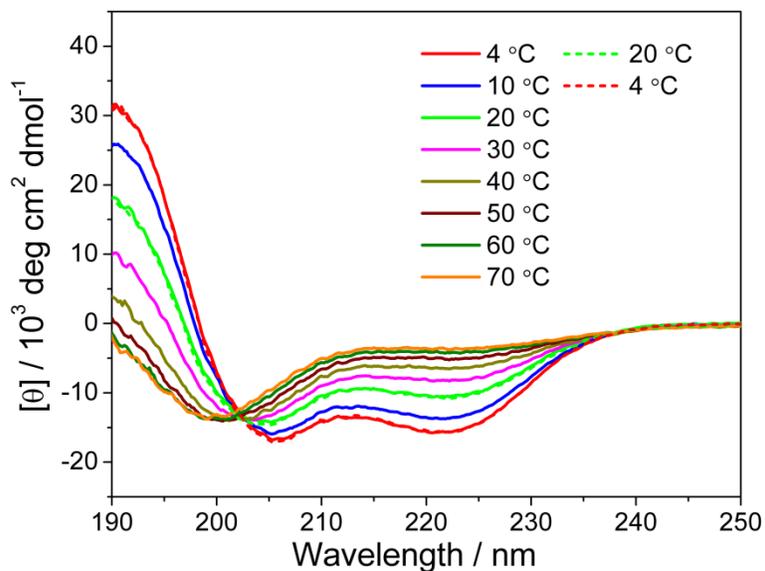


Figure 5. 7. Temperature dependence of CD spectra of AEP at pH 2.0.

The solid lines represent heating process, the dashed lines represent cooling process.

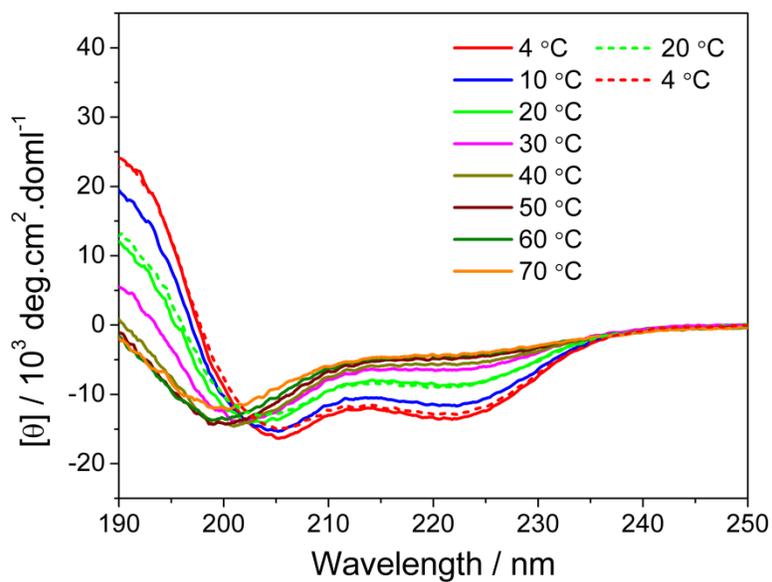


Figure 5. 8. Temperature dependence of CD spectra of AEP at pH 4.0.

The solid lines represent heating process, the dashed lines represent cooling process.

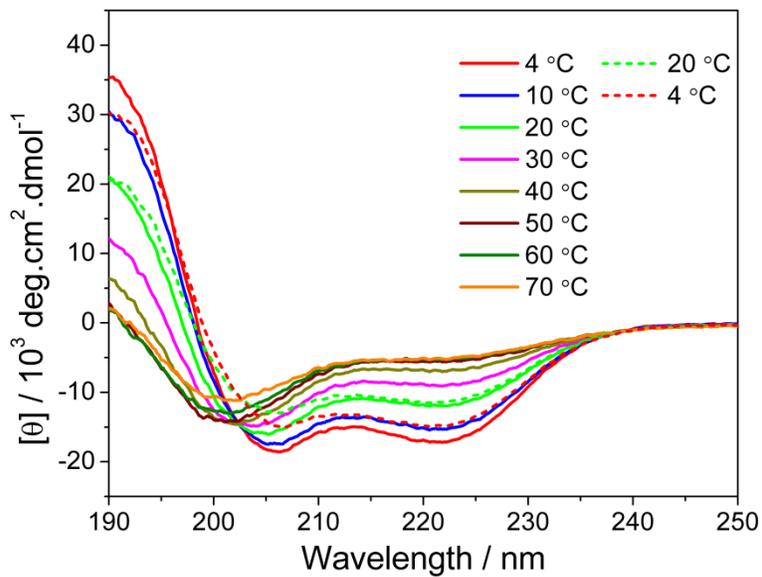


Figure 5. 9. Temperature dependence of CD spectra of AEP at pH 10.0.

The solid lines represent heating process, the dashed lines represent cooling process.

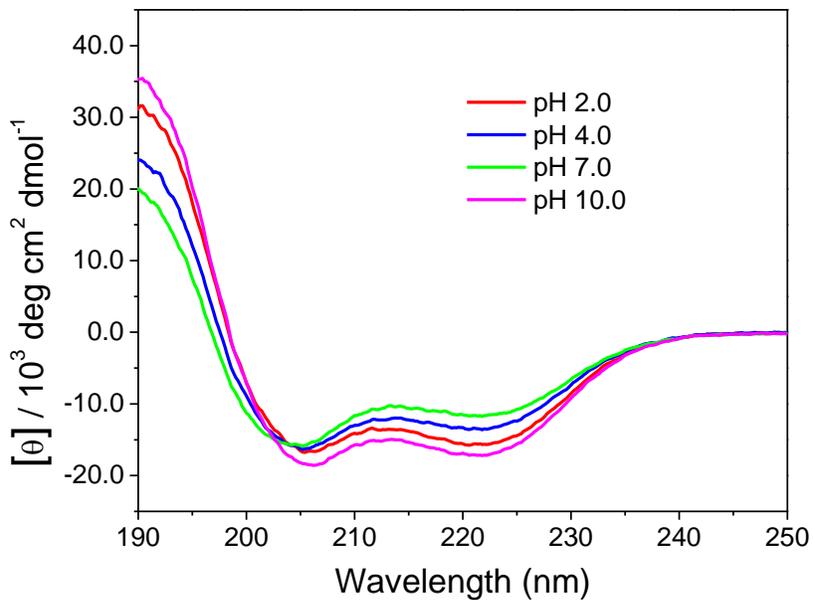


Figure 5. 10. pH dependence of CD spectra of AEP at 4 °C.

5.8 BIBLIOGRAPHY

- (1) Dill, K. A. *Protein Sci.* **1999**, *8*, 1166.
- (2) Dill, K. A.; Bromberg, S.; Yue, K. Z.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S. *Protein Sci.* **1995**, *4*, 561.
- (3) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1983**, *168*, 867.
- (4) Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1990**, *216*, 1031.
- (5) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211.
- (6) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol.* **1995**, *254*, 761.
- (7) Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293*, 1241.
- (8) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8898.
- (9) Stellwagen, E.; Park, S. H.; Shalongo, W.; Jain, A. *Biopolymers* **1992**, *32*, 1193.
- (10) Huyghuesdespointes, B. M. P.; Scholtz, J. M.; Baldwin, R. L. *Protein Sci.* **1993**, *2*, 80.
- (11) Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 9668.
- (12) HuyghuesDespointes, B. M. P.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 1965.
- (13) Smith, J. S.; Scholtz, J. M. *Biochemistry* **1998**, *37*, 33.
- (14) Olson, C. A.; Spek, E. J.; Shi, Z. S.; Vologodskii, A.; Kallenbach, N. R. *Proteins: Struct., Funct., Genet.* **2001**, *44*, 123.
- (15) Errington, N.; Doig, A. J. *Biochemistry* **2005**, *44*, 7553.
- (16) Iqbalsyah, T. M.; Doig, A. J. *J. Am. Chem. Soc.* **2005**, *127*, 5002.
- (17) Park, S. H.; Shalongo, W.; Stellwagen, E. *Biochemistry* **1993**, *32*, 12901.
- (18) Mayne, L.; Englander, S. W.; Qiu, R.; Yang, J. X.; Gong, Y. X.; Spek, E. J.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1998**, *120*, 10643.
- (19) Sindelar, C. V.; Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1998**, *7*, 1898.
- (20) Fezoui, Y.; Braswell, E. H.; Xian, W. J.; Osterhout, J. J. *Biochemistry* **1999**, *38*, 2796.
- (21) Shi, Z. S.; Olson, C. A.; Bell, A. J.; Kallenbach, N. R. *Biopolymers* **2001**, *60*, 366.
- (22) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2782.
- (23) Ghosh, T.; Garde, S.; Garcia, A. E. *Biophys. J.* **2003**, *85*, 3187.
- (24) Iqbalsyah, T. M.; Doig, A. J. *Biochemistry* **2005**, *44*, 10449.
- (25) Asher, S. A. *Anal. Chem.* **1993**, *65*, A201.
- (26) Asher, S. A. *Anal. Chem.* **1993**, *65*, A59.
- (27) Chi, Z. H.; Asher, S. A. *Biochemistry* **1998**, *37*, 2865.
- (28) Chi, Z. H.; Chen, X. G.; Holtz, J. S. W.; Asher, S. A. *Biochemistry* **1998**, *37*, 2854.
- (29) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 8074.
- (30) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775.
- (31) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **2001**, *123*, 2388.
- (32) Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433.
- (33) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. *J. Phys. Chem. B* **2004**, *108*, 19020.

- (34) Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 10943.
- (35) Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2005**, *109*, 3047.
- (36) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2006**, *110*, 1928.
- (37) Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2006**, *128*, 13789.
- (38) Mikhonin, A. V.; Myshakina, N. S.; Bykov, S. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 7712.
- (39) Ma, L.; Ahmed, Z.; Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2007**, *111*, 7675.
- (40) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 5286.
- (41) Merutka, G.; Shalongo, W.; Stellwagen, E. *Biochemistry* **1991**, *30*, 4245.
- (42) Cheng, R., P.; Girinath, P.; Ahmad, R. *Biochemistry* **2007**, *46*, 10528.
- (43) Lyu, P. C. C.; Gans, P. J.; Kallenbach, N. R. *J. Mol. Biol.* **1992**, *223*, 343.
- (44) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. *Appl. Spectrosc.* **2005**, *59*, 1541.
- (45) Kallenbach, N. R.; Lyu, P.; Zhou, H. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996.
- (46) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526.
- (47) Cantor, C. R.; Schimmel, P. R. *Biophysical chemistry*; W. H. Freeman and Company: New York, 1980.
- (48) Scholtz, J. M.; Qian, H.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biopolymers* **1991**, *31*, 1463.
- (49) Yang, J. X.; Zhao, K.; Gong, Y. X.; Vologodskii, A.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1998**, *120*, 10646.
- (50) Scheraga, H. A. *Pure Appl. Chem.* **1973**, *36*, 1.
- (51) Wojcik, J.; Altmann, K. H.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 121.
- (52) Padmanabhan, S.; York, E. J.; Gera, L.; Stewart, J. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 8604.
- (53) Zimm, B. H.; Doty, P.; Iso, K. *Proc. Natl. Acad. Sci. U. S. A.* **1959**, *45*, 1601.
- (54) Ostroy, S. E.; Lotan, N.; Ingwall, R. T.; Scheraga, H. A. *Biopolymers* **1970**, *9*, 749.
- (55) Ptitsyn, O. B. *Pure Appl. Chem.* **1972**, *31*, 227.
- (56) Daggett, V.; Kollman, P. A.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1115.
- (57) Wang, Y.; Purrello, R.; Jordan, T.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359.
- (58) Woody, R. W.; Tinoco Jr., I. *J. Chem. Phys.* **1967**, *46*, 4927.
- (59) Madison, V.; Schellman, J. *Biopolymers* **1972**, *11*, 1041.
- (60) Woody, R. W. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996.
- (61) Ozdemir, A.; Lednev, I. K.; Asher, S. A. *Biochemistry* **2002**, *41*, 1893.
- (62) Mix, G.; Schweitzer-Stenner, R.; Asher, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 9028.
- (63) Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2007**, *111*, 4271.
- (64) Tinoco Jr., I.; Halpern, A.; Simpson, W. T. In *Polyamino Acides, Polypeptides, and Proteins*; Stahmann, M. A., Ed.; University of Wisconsin Press: Madison, 1961.
- (65) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101*, 535.
- (66) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13934.
- (67) Zhang, W.; Lei, H. X.; Chowdhury, S.; Duan, Y. *J. Phys. Chem. B* **2004**, *108*, 7479.

- (68) Chakrabarty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994**, 3, 843.
- (69) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, 250, 646.

6.0 UV RESONANCE RAMAN AND DFT STUDIES OF ARGININE SIDE CHAINS IN PEPTIDES: INSIGHTS INTO ARGININE HYDRATION

This chapter has been published: *J. Phys. Chem. B*, **2013**, *117*, 7145-7156

Authors: Zhenmin Hong, Jonathan Wert and Sanford A. Asher*

Address: Department of Chemistry, University of Pittsburgh, PA 15260

We examined the UV resonance Raman (UVRR) spectra of four models of the Arg side chain, guanidinium (gdn), ethylguanidinium (EG), arginine (Arg) and Ac-Arg-OMe (AAO) in H₂O and D₂O, in order to identify spectral markers that report on the environment of the Arg side chain. To elucidate the resonance Raman enhancement mechanism of the Arg side chain, we used DFT to calculate the equilibrium geometries of the electronic ground state and the first excited state. We determined the vibrational mode frequencies of the ground state and the first derivative of the first electronic excited state potential energy with respect to each vibrational normal mode of the electronic ground state at the electronic ground state equilibrium geometry. The DFT

calculations and the potential energy distributions reveal that, in addition to the gdn group C-N stretching vibrations, the C-N bond stretching vibration of the gdn group-methylene linkage is also strongly resonance enhanced in EG, Arg and AAO. From the UVRR spectra, we find that the Raman cross section and frequency of the $\sim 1170\text{ cm}^{-1}$ vibration of the Arg side chain depends on its hydration state and can be used to determine the hydration state of the Arg side chain in peptides and proteins. We examined the hydration of the Arg side chain in two polyala peptides and found that in the α -helical conformation the Arg side chain in the AEP peptide (sequence: $A_9RA_3EA_4RA_2$) is less hydrated than that in the AP peptide (sequence: $A_8RA_4RA_4RA_2$).

6.1 INTRODUCTION

UV resonance Raman (UVRR) spectroscopy has been shown to be a powerful tool to study the conformations and folding dynamics of peptides and proteins.¹⁻¹⁰ Excitation in the deep UV ($< 230\text{ nm}$) selectively excites protein backbone vibrations and amino acid side chain vibrations. The side chain vibrations are sensitive to hydration and hydrogen bonding. The UVRR spectra of tryptophan,¹¹⁻¹⁸ tyrosine,^{11-13,19} phenylalanine^{11-12,19} and histidine^{18,20} have been extensively studied, while the UVRR spectrum of the arginine (Arg) side chain that is resonance enhanced by $\sim 200\text{ nm}$ excitation has been little investigated.

Hudson and coworkers²¹ systematically studied the electronic and vibrational modes of guanidinium (gdn) and some substituted derivatives. They studied the excitation profiles of gdn, methylguanidinium, and N,N-dimethylguanidinium in both H_2O and D_2O . They calculated the electronic ground and first excited states of these compounds using the semiempirical INDO/S

methods and assigned some of the Raman bands; however, due to their relatively poor spectral resolution and their poor signal-to-noise ratios, they were unable to study ethylguanidinium (EG) and Arg in detail.

The Arg side chain is positively charged up to pH 12 and should be strongly hydrated in water. Early measurements of the Arg amino acid distribution coefficient between water and the vapor phase suggested that Arg is the most hydrated residue; the hydration energy is double that of negatively charged asp or Glu.²² More recent studies implied that Arg might not be as hydrated as previously suggested. Mason et al.²³⁻²⁵, on the basis of neutron diffraction and MD simulations, claimed that gdn could form dimers in aqueous solution. Hunger et al.²⁶ used broadband dielectric relaxation spectroscopy to examine the hydration of gdn and found that gdn does not bind strongly to water molecules. McQuinn et al.²⁷ utilized energy-dependent ESI-MS/MS to study the dehydration of gdn-water and methylated gdn-water nanodroplets and found that the number of water molecules bound to the gdn group decreases upon alkyl group substitution. These studies suggest that the gdn group in Arg might not be very hydrated.

Garica et. al.,²⁸ using MD simulations, calculated that the mean residence time of water in the hydration shells of Arg residues in proteins was comparable, and even slightly less than that of negatively charged asp and Glu; this result indicated that Arg is not more hydrated than the other charged residues. The conclusions of Garcia et. al. were later confirmed by whole-residue hydrophobicity scales obtained by measuring the transfer free energy of a specific residue in a host peptide from water to a bilayer interface or to n-octanol.²⁹ On the whole-residue hydrophobicity scale, Arg is the least hydrated charged residue (among lys, his, Glu and asp).

We here extend the work of Hudson and coworkers to EG, Arg and N-acetyl arginine methyl ester (Ac-Arg-OMe, AAO). The ethyl substituent of EG breaks the gdn symmetry,

turning on some Raman inactive vibrations in gdn to result in a different UVRR spectrum. AAO is an N-acetyl-capped arginine methyl ester, modeling the Arg residue in peptides. Our recent study on the Arg residue Raman excitation profile suggested that the Arg side chain electronic transitions would interact with the peptide backbone amide π - π^* transition.³⁰ Figure 6.1 shows the molecular structures of gdn, EG, Arg and AAO.

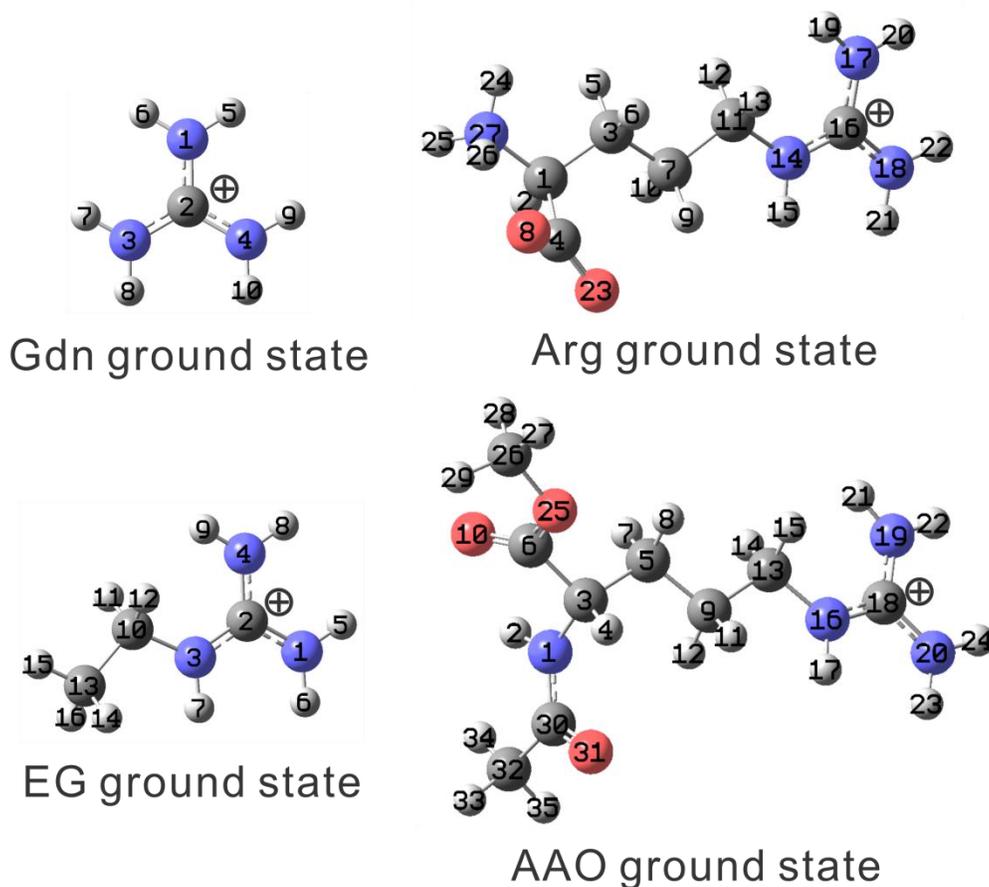


Figure 6. 1. The calculated electronic ground state equilibrium geometries of gdn, EG, Arg and AAO.

We measured the 204 nm excited UVRR spectra of gdn, EG, Arg and AAO and used DFT to calculate their electronic ground and first excited state conformations and their resonance Raman spectra. We also find that the $\sim 1170\text{ cm}^{-1}$ Arg band is sensitive to the hydration of the gdn group. We used this marker band to study the conformational dependence of the hydration of Arg in the peptides AP (sequence: $A_8RA_4RA_4RA_2$) and AEP (sequence: $A_9RA_3EA_4RA_2$).

6.2 EXPERIMENTAL SECTION

6.2.1 Materials

Guanidinium chloride (GC), ethylguanidinium sulfate (98%, EGS), L-arginine chloride (98%), sodium perchlorate (98%, NaClO₄), barium perchlorate (97%, Ba(ClO₄)₂), silver perchlorate (anhydrous, 97%, AgClO₄) and acetonitrile (spectrophotometric grade) were purchased from Sigma-Aldrich Co. and used as received. Ba(ClO₄)₂ and AgClO₄ are explosive, handle with care. Ac-Arg-OMe chloride (AAOC) was purchased from Bachem Americas, Inc. The AP (A₈RA₄RA₄RA₂, >95% purity) and AEP (A₉RA₃EA₄RA₂, >95% purity) peptides were synthesized by Anaspec Inc. D₂O (99.9%) was purchased from Cambridge Isotope Laboratories, Inc. Water (18.2 MΩ·cm) was purified by a NANOPURE Infinity™ ultrapure water purifier. The concentrations of gdn, EG, Arg and AAO for UVRR were 10 mM. 0.2 M NaClO₄ was added as an internal standard for each solution.

6.2.2 Synthesis of ethylguanidinium perchlorate and Ac-Arg-OMe perchlorate

Due to the high charge density of sulfate and chloride anions, EGS and AAOC are insoluble in acetonitrile. We exchanged the sulfate and chloride counter ions to perchlorate in order to make the compounds soluble.

To convert EGS and AAOC into ethylguanidinium perchlorate (EGPC) and Ac-Arg-OMe perchlorate (AAOPC), ~0.1 M Ba(ClO₄)₂ or ~0.1 M AgClO₄ was slowly added to the ~0.1 M EGS or ~0.1 M AAOC aqueous solutions to form insoluble BaSO₄ or AgCl, respectively. The suspension was centrifuged several times to remove the precipitate. The supernatant was

collected and gently heated with stirring to evaporate the water. The resulting EGPC and AAOPC were very light yellow viscous ionic liquids at room temperature. The yellow color derived from impurities. The calculated yields were 93% and 105%, respectively. The fact that the yield was greater than 100% indicated that a small amount of water remained in AAOPC.

6.2.3 204 nm excited UVRR spectra

The UVRR instrumentation has been described in detail elsewhere.³¹ Briefly, the 204 nm excitation beam, generated from the frequency mixing of the fundamental and the third harmonic of a Nd:YLF pumped Ti:Sapphire laser (Photonic Industries, Inc.), was directed to an open flow sample stream. The backscattered Raman light was collected and then dispersed using a partially subtractive double monochromator and finally detected by a liquid nitrogen cooled, Lumogen coated, back-thinned CCD camera (Princeton Instruments Spec-10:400B).

The Raman scattering cross section of a specific band is calculated by:

$$\sigma_s = \frac{\sigma_{IS} c_{IS} I_s \eta_{IS}}{c_s I_s \eta_s} \quad (1)$$

where σ is the Raman scattering cross section, c represents the concentration, I is the integrated intensity of a specific band, and η is the instrument efficiency at the frequency of the specific Raman band. The subscript s denotes the sample, which is EG or AAO, while the subscript IS denotes the perchlorate internal standard. The instrument efficiencies at the sample band (η_s) and the internal standard band (η_{IS}) were assumed to be equal. The Raman scattering cross section of perchlorate was estimated to be $0.12 \times 10^{-26} \text{ cm}^2 \cdot \text{molecule sr}^{-1}$ at 204 nm excitation by extrapolating the Raman cross section measurement of Dudik et al³².

6.2.4 DFT calculations

The electronic ground state equilibrium structures of gdn, EG, Arg and AAO, as well as their electronic ground state vibrational frequencies, were calculated at the B3LYP/aug-cc-pVTZ level in water using the Polarizable Continuum Model (PCM) to model solvent effects with the Gaussian 09 package.³³ The calculated frequencies were not scaled. The potential energy distributions (PED) of the vibrational normal modes were calculated using a program that we wrote.³⁴⁻³⁶ The force along a ground state vibrational normal coordinate in the first electronic excited state at the ground state equilibrium geometry is the negative of the first derivative of the first electronic excited state potential energy function with respect to the ground state normal coordinate at the ground state equilibrium geometry. These first electronic excited state forces at the ground state equilibrium geometry, the first electronic excited state equilibrium geometries and the vibrational frequencies of gdn and EG were calculated by time-dependent DFT at the B3LYP/aug-cc-pVTZ level in water with the PCM using the Gaussian 09 package.³³ The atomic charges were calculated using the atoms-in-molecules (AIM) method using the Bader charge analysis program developed by Andri Arnaldsson, Wenjie Tang and Graeme Henkelman from the University of Texas at Austin.³⁷⁻³⁹

6.3 RESULTS AND DISCUSSIONS

6.3.1 Vibrational assignments

Figure 6.2 shows the 204 nm excited UVRR spectra of gdn, EG, Arg and AAO in both H₂O and D₂O. The gdn UVRR spectra are very similar to those measured by Hudson and coworkers.²¹ The spectrum of gdn in H₂O shows a sharp intense CN₃ symmetric stretching band at 1008 cm⁻¹ in addition to three weak, broad bands at ~1684 cm⁻¹, ~1578 cm⁻¹ and ~1466 cm⁻¹. The 1684 cm⁻¹ and 1578 cm⁻¹ bands result from the coupling between CN₃ asymmetric stretching and NH₂ scissoring vibrations.

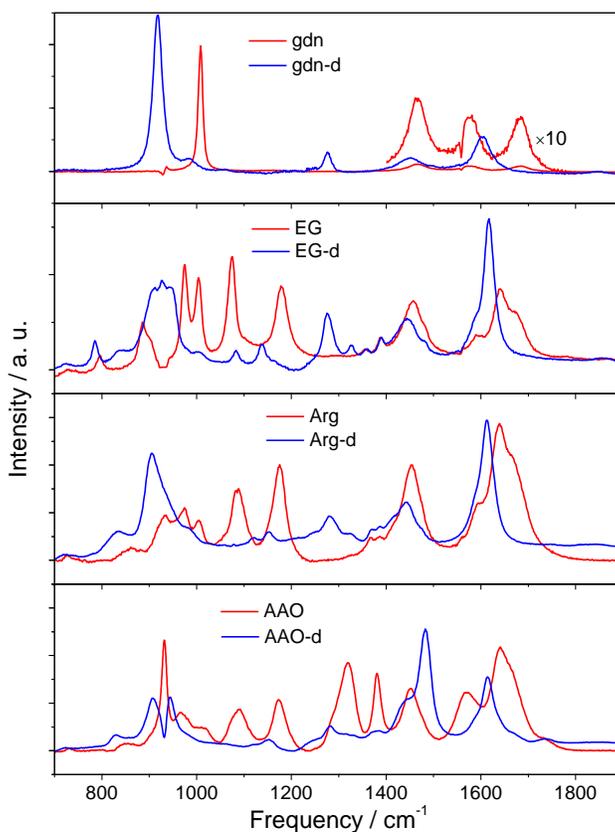


Figure 6. 2. The 204 nm excited UVRR spectra of gdn, EG, Arg, AAO in both H₂O and D₂O.

For gdn-d, the NH₂ hydrogen exchange for deuterium, causing the CN₃ asymmetric stretching and NH₂ scissoring vibrations to decouple and shift to 1604 cm⁻¹ for the CN₃ asymmetric stretching and to 1278 cm⁻¹ for ND₂ scissoring. The 1278 cm⁻¹ band is resonance enhanced due to a significant contribution from CN₃ symmetric stretching as discussed below.

Table 6. 1. Gdn and gdn-d vibrational normal modes involving at least 5% contribution from CN₃ stretching

	Exp (cm ⁻¹)	Calc (cm ⁻¹)	PED (%) [*]
gdn	1684	1694	N3-H ₂ scs (40), N4-H ₂ scs (40), N1-H ₂ scs (12), C2-N ₃ s str (5)
		1674	C2-N ₃ as str' (39), N3-H ₂ scs (26), N4-H ₂ scs (26), C2-N3N4 rock (5)
		1673	N1-H ₂ scs (45), C2-N ₃ as str (38), C2-N3N4 scs (5)
	1578	1592	C2-N ₃ as str (35), N1-H ₂ scs (33), C2-N3N4 scs (11), N4-H ₂ rock (7), N3-H ₂ rock (7)
		1588	C2-N ₃ as str' (35), N3-H ₂ scs (21), N4-H ₂ scs (21), C2-N3N4 rock (10), N1-H ₂ rock (8)
		1126	N4-H ₂ rock (35), N3-H ₂ rock (35), C2-N ₃ as str (28)
		1116	N1-H ₂ rock (45), C2-N ₃ as str' (27), N4-H ₂ rock (13), N3-H ₂ rock (12)
	1008	1023	C2-N ₃ s str (98)
gdn-d	1604	1602	C2-N ₃ as str' (76), C2-N3N4 rock (12), N1-D ₂ rock (5)
		1601	C2-N ₃ as str (75), C2-N3N4 scs (13)
	1278	1302	N3-D ₂ scs (26), N4-D ₂ scs (26), C2-N ₃ s str (26), N1-D ₂ scs (21)
	919	920	C2-N ₃ s str (70), N1-D ₂ scs (7)
		911	N4-D ₂ rock (28), N3-D ₂ rock (28), C2-N ₃ as str (17), C2-N ₃ s str (10), C2-N3N4 scs (8)
		909	N1-D ₂ scs (43), C2-N ₃ as str' (19), N4-D ₂ rock (11), N3-D ₂ rock (11), C2-N3N4 rock (9)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking;

Table 6.1 shows the calculated PED for the normal modes of gdn and gdn-d that contain at least 5% contributions from CN₃ stretching. The 1466 cm⁻¹ gdn band in H₂O and the 1451 cm⁻¹

¹ gdn-d band in D₂O are assigned to the first overtone of the CN₃ out-of-plane deformation mode. The CN₃ out-of-plane deformation mode, calculated to be 727 cm⁻¹ for gdn or 709 cm⁻¹ for gdn-d, is Raman inactive; however, its first overtone is Raman active. The small downshift of this gdn-d band in D₂O indicates that this vibrational mode also contains a small contribution from ND₂ motions. In D₂O, the 919 cm⁻¹ CN₃ symmetric stretching band downshifts from its 1008 cm⁻¹ frequency in H₂O, confirming that CN₃ symmetric stretching is coupled to ND₂ scissoring.

Table 6.2 displays the PED for the normal modes of EG and EG-d. The EG UVRR spectra in both H₂O and D₂O show more features than the gdn UVRR spectra because the ethyl group substitution breaks the gdn symmetry. The CN₃ asymmetric stretching motion couples to NH₂ scissoring and a small amount of CN₂ rocking and/or CN₂ scissoring, giving rise to three moderately resolved bands at ~1674 cm⁻¹, ~1639 cm⁻¹ and ~1590 cm⁻¹. For EG-d, the CN₃ asymmetric stretching and ND₂ scissoring are decoupled, giving rise to a narrow 1617 cm⁻¹ band with a shoulder at ~1588 cm⁻¹.

Table 6. 2. EG and EG-d vibrational normal modes involving at least 5% contribution from CN₃ stretching

	Exp (cm ⁻¹)	Calc (cm ⁻¹)	PED (%) [*]
EG	~1674	1689	N4-H ₂ scs (46), N1-H ₂ scs (32), C2-N ₃ as str (7), C2-N ₃ as str' (5)
		1673	N1-H ₂ scs (35), C2-N ₃ as str' (33), N4-H ₂ scs (9), N3-H ipb (8), C2-N1N4 rock (5), C2-N ₃ as str (5)
	1639	1644	C2-N ₃ as str (43), N1-H ₂ scs (16), N3-H ipb (15), C2-N1N4 scs (7), N4-H ₂ scs (6), N4-H ₂ rock (46)
	1590	1603	N4-H ₂ scs (33), C2-N ₃ as str (16), C2-N ₃ as str' (16), N3-H ipb (11), N1-H ₂ rock (8), C2-N1N4 rock (6), C2-N1N4 scs (5)
		1457	N3-H ipb (36), C2-N ₃ as str' (20), C10-H ₂ wag (18)
		1381	C10-H ₂ wag (66), N3-H ipb (13), C2-N ₃ as str' (5), C13-H ₃ s def (5)

	1180	1180	N3-C10 str (47), C13-H ₃ rock' (12), N1-H ₂ rock (9), C2-N ₃ s str (8)
		1141	N4-H ₂ rock (22), C13-H ₃ rock' (15), C10-C13 str (10), C2-N ₃ as str (9), C10-C13N ₃ scs (9), N1-H ₂ rock (7), C2-N ₃ s str (5), C13-H ₃ rock (5), C2-N ₃ as str' (5)
	1075	1084	C2-N ₃ s str (33), N1-H ₂ rock (27), C2-N ₃ as str (13), C10-C13 str (10), N3-C10 str (9)
		1070	N4-H ₂ rock (45), N1-H ₂ rock (33), C2-N ₃ as str' (6), C13-H ₃ rock' (5)
	975	1016	C10-C13 str (58), C2-N ₃ s str (31), N3-C10 str (5)
		907	C13-H ₃ rock' (24), N3-C10 str (19), C2-N ₃ s str (11), N3-C2C10 ipb (10), C13-H ₃ rock (9), C10-H ₂ wag (6), C2-N1N4 rock (6)
EG-d	1617	1612	C2-N ₃ as str (68), C2-N1N4 scs (11)
	~1585	1593	C2-N ₃ as str' (68), C2-N1N4 rock (14)
	1326	1307	N1-D ₂ scs (34), C2-N ₃ s str (28), N4-D ₂ scs (14), N3-D ipb (12), N3-C10 str (5)
	945	956	N3-D ipb (44), C2-N ₃ s str (32), N3-C10 str (6), C13-H ₃ rock' (5), N4-D ₂ scs (5)
	913	910	N4-D ₂ rock (26), N1-D ₂ rock (24), C2-N ₃ as str (16), C2-N ₃ s str (16), C2-N1N4 rock (7)
		899	C13-H ₃ rock' (20), N3-C10 str (14), N4-D ₂ rock (11), C2-N ₃ s str (9), N3-C2C10 ipb (9), C2-N1N4 rock (8), C13-H ₃ rock (7), C10-H ₂ wag (5)
	785	813	N1-D ₂ rock (46), N4-D ₂ rock (25), N3-C10 str (8), C2-N ₃ s str (5)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking; wag: wagging; twist: twisting; ipb: in-plane bending; s def: symmetric deformation

Hudson and coworkers²¹ proposed that the overtone of the CN₃ out-of-plane deformation mode is Raman active in alkyl substituted gdns. By comparing the ~1457 cm⁻¹ band of EG and the ~1442 cm⁻¹ band of EG-d with the 1466 cm⁻¹ band of gdn and the 1451 cm⁻¹ band of gdn-d, it is reasonable to assign these bands to the first overtones of the CN₃ out-of-plane deformation

modes. This assignment is supported by the calculation, which gives fundamental frequencies of 728 cm^{-1} for EG and 720 cm^{-1} for EG-d. The methylene CH_2 scissoring band appears as a shoulder at $\sim 1482 \text{ cm}^{-1}$ in the spectra of both EG and EG-d. The bands at $\sim 1389 \text{ cm}^{-1}$ and $\sim 1356 \text{ cm}^{-1}$ show little frequency shift between EG and EG-d. Thus, these bands are assigned to the methyl CH_3 umbrella mode and the methylene CH_2 wagging. The band at 1180 cm^{-1} involving C10-N3 stretching is strongly resonance enhanced as shown below. This band is sensitive to the environment of the gdn group.

In EG and EG-d, the CN_3 symmetric stretching motion is distributed over several bands. We are unable to unambiguously assign the bands below 1100 cm^{-1} .

The Arg and Arg-d PED for its vibrational normal modes are given in Table 6.3. The UVRR spectra of Arg and Arg-d are very similar to those of EG and EG-d; however, the Arg bands are somewhat broader than those of EG and EG-d. Similar to the 1180 cm^{-1} band EG, the $\sim 1175 \text{ cm}^{-1}$ band involving C13-N16 stretching is well resolved in the UVRR spectra of Arg. The bands below 1000 cm^{-1} differ significantly from those of EG.

Table 6. 3. Arg and Arg-d vibrational normal modes involving at least 5% contribution from CN_3 stretching

	Exp (cm^{-1})	Calc (cm^{-1})	PED (%) [*]
Arg	~ 1667	1686	N17-H ₂ scs (54), N18-H ₂ scs (25), C16-N ₃ as str (6), C16-N ₃ as str' (6)
		1672	N18-H ₂ scs (37), C16-N ₃ as str' (32), N14-H ipb (8), N17-H ₂ scs (7), C16-N ₃ as str (6), C16-N17N18 rock (5)
	1640	1643	C18-N ₃ as str (46), N18-H ₂ scs (17), N14-H ipb (12), C16-N17N18 scs (8), N17-H ₂ rock (5)
	1596	1598	N17-H ₂ scs (32), C16-N ₃ as str' (18), N14-H ipb (13), C16-N ₃ as str (11), C16-N17N18 rock (7), N18-H ₂ rock (7), N18-H ₂ scs (6)
		1459	N14-H ipb (30), C11-H ₂ wag (29), C16-N ₃ as str' (17)
	1176	1167	N14-C11 str (25), N18-H ₂ rock (12), C16-N ₃ s str (10), N27-H ₃ rock' (8),

			N17-H ₂ rock (7), C16-N ₃ as str (6), C1-C3 str (5)
	1085	1084	C16-N ₃ s str (18), N17-H ₂ rock (17) , C16-N ₃ as str (12), N14-C11 str (7), N18-H ₂ rock (7), N27-H ₃ rock (5), C3-C7 str (5)
		1072	N18-H ₂ rock (36), N17-H ₂ rock (15), C16-N ₃ s str (8), C7-C11 str(7), C3-C7 str (7)
	975	1005	C1-N27 str (20), C16-N ₃ s str (16), C1-C3 str (15), N27-H ₃ rock' (10), N14-C11 str (6)
		966	C1-N27 str (22), C16-N ₃ s str (22), N14-C11 str (13), C3-C7 str (12), C7-C11 str(7)
Arg-d	1612	1613	C16-N ₃ as str (70), C16-N17N18 scs (11)
	~1586	1593	C16-N ₃ as str' (70), C16-N17N18 rock (14)
	1279	1314	N18-D ₂ scs (17), C16-N ₃ s str (17), C3-H ₂ wag (16), C7-H ₂ wag (13), N17-D ₂ scs (8), N14-D ipb (8)
		1297	C1-H rock' (22), C7-H ₂ wag (19), C3-H ₂ wag (13), N18-D ₂ scs (7), C11-H ₂ wag (5), C11-H ₂ twist (5), C16-N ₃ s str (5)
		1294	C1-H rock' (20), C1-H rock (13), C11-H ₂ twist (11), C7-H ₂ twist (7), N18-D ₂ scs (7), C3-H ₂ rock(6), C7-H ₂ wag (5), C16-N ₃ s str (5)
		942	C1-N27 str (18), C1-C4 str (10), C16-N ₃ s str (9), N14-D ipb (8), C11-H ₂ rock (6), C3-H ₂ rock (5)
	907	926	C16-N ₃ s str (31), C1-N27 str (9), N17-D ₂ rock (6), C3-C7 str (6)
		907	N17-D ₂ rock (28), N18-D ₂ rock (24), C16-N ₃ as str (16), C16-N ₃ s str (15), C16-N17N18 scs (7)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking; wag: wagging; twist: twisting; ipb: in-plane bending

The PED for the vibrational normal modes of AAO and AAO-d are shown in Table 6.4. In addition to enhancement of Arg vibrations, the AAO and AAO-d UVRR spectra show enhancement of amide vibrations because the amide π - π^* electronic transition is also in

resonance with 204 nm excitation. The Am I band overlaps with the gdn group CN₃ asymmetric stretching band. The Am II band appears at ~1570 cm⁻¹. The band at 1452 cm⁻¹ in AAO and the shoulder appearing at ~1435 cm⁻¹ in AAO-d are assigned to the overtones of the CN₃ out-of-plane deformation modes. The narrow band at 1381 cm⁻¹ originates from the coupling of the methyl CH₃ bending mode of the acetyl group with the N-H bending mode of the amide bond. The Am III band occurs at 1319 cm⁻¹. The AAO 1170 cm⁻¹ band that involves C13-N16 stretching is well resolved in the UVRR spectrum. This band is equivalent to the 1180 cm⁻¹ band of EG and is also sensitive to the gdn group environment.

Table 6. 4. AAO and AAO-d vibrational normal modes involving at least 5% contribution from CN₃ stretching

	Exp (cm ⁻¹)	Calc (cm ⁻¹)	PED (%) [*]
AAO		1692	N19-H ₂ scs (50), N20-H ₂ scs (28), C18-N ₃ as str (6), C18-N ₃ as str' (5)
		1673	N20-H ₂ scs (39), C18-N ₃ as str' (31), N19-H ₂ scs (8), N16-H ipb (7), C18-N ₃ as str (6), C18-N19N20 rock (5)
		1644	C18-N ₃ as str (43), N16-H ipb (16), N20-H ₂ scs (15), C18-N19N20 scs (7), N19-H ₂ scs (6), N19-H ₂ rock (5)
		1605	N19-H ₂ scs (29), C18-N ₃ as str' (17), C18-N ₃ as str (17), N16-H ipb (11), N20-H ₂ rock (8), C18-N19N20 rock (6), C18-N19N20 scs (5),
		1461	N16-H ipb (32), C13-H ₂ wag (23), C18-N ₃ as str' (19)
	1170	1170	N16-C13 str (39), N20-H ₂ rock (8), C18-N ₃ s str (13), N19-H ₂ rock (8), C18-N ₃ as str (7)
	1091	1084	C18-N ₃ s str (29), N20-H ₂ rock (29), C18-N ₃ as str (11), C9-C13 str (10), N16-C13 str (6)
	~1018	1005	C26-O25 str (29), C3-C6 str (18), C3-C5 str (14), C18-N ₃ s str (6)
		990	C26-O25 str (14), C30-C32 str (11), N1-C3C30 ipb (8), C6-C3-N1 def (6), C18-N ₃ s str (6), C30-N1 str (6), C30-O31 ipb (5)
	966	973	C18-N ₃ s str (22), N16-C13 str (12), C5-C9 str (10), C26-O25 str (6), C9-

			C13 str (6), C3-C6 str (5)
AAO-d	1617	1612	C18-N ₃ as str (69), C18-N19N20 scs (11)
		1593	C18-N ₃ as str' (68), C18-N19N20 rock (14)
	1282	1308	N20-D ₂ scs (25), C18-N ₃ s str (20), N19-D ₂ scs (11), C3-H rock (10), N16-D ipb (8), N16-C13 str (5), C13-H ₂ twist (5)
	944	935	C18-N ₃ s str (23), N1-D ipb (21), C30-N1 str (9), C32-H3 rock' (5)
		931	C18-N ₃ s str (23), N1-D ipb (16), C30-N1 str (12), C32-H3 rock' (5)
	907	910	N19-D ₂ rock (31), N20-D ₂ rock (22), C18-N ₃ as str (16), C18-N ₃ s str (1), C18-N19N20 scs (7)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking; wag: wagging; twist: twisting; ipb: in-plane bending; def: deformation

For AAO-d, the Am I, Am II and Am III bands disappear, while the Am II' band (C-N stretching) appears at 1483 cm⁻¹, revealing the CN₃ asymmetric stretching at 1617 cm⁻¹. The weak, broad bump at ~1376 cm⁻¹ is probably an overlapping methylene CH₂ wagging band.

6.3.2 Calculated electronic ground state structures of gdn and EG

The calculated electronic ground state equilibrium geometry of gdn in water has C₂ symmetry, which is different from those geometries calculated in gas phase.⁴⁰⁻⁴¹ Gobbi and Frenking examined the electronic structure of gdn and concluded that the resonance stabilization of gdn is not strong enough to constrain the molecule to have D_{3h} symmetry.⁴¹

The calculated gdn group electronic ground state geometry of EG is similar to that of gdn; however, in EG the C2-N3 bond is 0.001 Å shorter, while the C2-N1 and C2-N4 bonds are 0.005 Å and 0.002 Å longer than the corresponding C-N bonds of gdn (Table 6.5). The

shortening of the C2-N3 bond is due to electron release by the adjacent ethyl group, allowing the positive charge on the gdn group to extend into the ethyl group (Table 6.6).

Table 6. 5. Calculated bond lengths of gdn and EG in the ground and first excited states*

	Bond length	Ground state (Å)	Excited state (Å)	Difference (Å)
gdn	C2-N1	1.3305	1.3654	0.0349
	C2-N3	1.3305	1.4085	0.0780
	C2-N4	1.3305	1.4085	0.0780
EG	C2-N1	1.3353	1.4101	0.0748
	C2-N3	1.3291	1.3650	0.0359
	C2-N4	1.3325	1.4110	0.0785
	N3-C10	1.4672	1.4655	-0.0017
	C10-C13	1.5192	1.5198	0.0006

*Bond lengths are calculated at B3LYP/aug-cc-pVTZ level using Gaussian 09 package.

The sp^2 hybridization of the N3 nitrogen atom requires the C2, N3, H7 and C10 atoms to be planar; however, the N3, H7, C10 and C13 atoms are not planar due to the sp^3 hybridization of C10 carbon atom.

Table 6. 6. Calculated atomic charges of gdn and EG in the ground and first excited states*

	Atom	Ground state	Excited state	Difference
gdn	N1	-1.18	-1.07	0.11
	C2	1.68	1.21	-0.47
	N3	-1.22	-0.97	0.25
	N4	-1.22	-0.97	0.25
	H5	0.48	0.45	-0.03
	H6	0.48	0.45	-0.03
	H7	0.50	0.47	-0.03
	H8	0.48	0.49	0.01

	H9	0.49	0.48	-0.01
	H10	0.48	0.47	-0.01
EG	N1	-1.22	-0.97	0.25
	C2	1.69	1.17	-0.52
	N3	-1.16	-1.07	0.09
	N4	-1.19	-0.96	0.23
	H5	0.48	0.45	-0.03
	H6	0.49	0.49	0.00
	H7	0.45	0.45	0.00
	H8	0.48	0.48	0.00
	H9	0.47	0.47	0.00
	C10	0.36	0.46	0.10
	H11	0.02	0.04	0.02
	H12	0.03	0.02	-0.01
	C13	0.00	0.05	0.05
	H14	0.02	0.00	-0.02
	H15	0.04	0.02	-0.02
	H16	0.03	-0.01	0.04

*The atomic charges were calculated by using Bader program developed by Andri Arnaldsson, Wenjie Tang and Graeme Henkelman in The University of Texas at Austin. <http://theory.cm.utexas.edu/vtsttools/bader/>

We calculated the effect of changing the torsion angle of the C13-C10-N3-C2 linkage on the energy of EG using DFT. The calculated potential energy surface along the C13-C10-N3-C2 torsion angle (Figure 6.3) is consistent with a previous report⁴² that indicated the trans-conformation (180°) is the global energy minimum. Two local minima corresponding to ±90° conformations exist 2.4 kJ·mol⁻¹ above the global minimum. At 298 K, with the thermal energy correction, which corrects the Gibbs free energy by accounting for the molecular translational

energy, rotational energy and vibrational energy at 298 K and 1 atm, the energy gap is 6.6 $\text{kJ}\cdot\text{mol}^{-1}$. The population fractions of these three conformers are calculated from the Boltzmann distribution to be 0.88, 0.06 and 0.06 respectively. Based on these calculations, EG likely occurs in a conformational distribution. The low energy barriers of $\sim 5 \text{ kJ}\cdot\text{mol}^{-1}$ (Figure 6.3, with thermal energy correction at 298 K and 1 atm, the barrier is $\sim 9 \text{ kJ}\cdot\text{mol}^{-1}$) suggest fast transitions between the conformers.

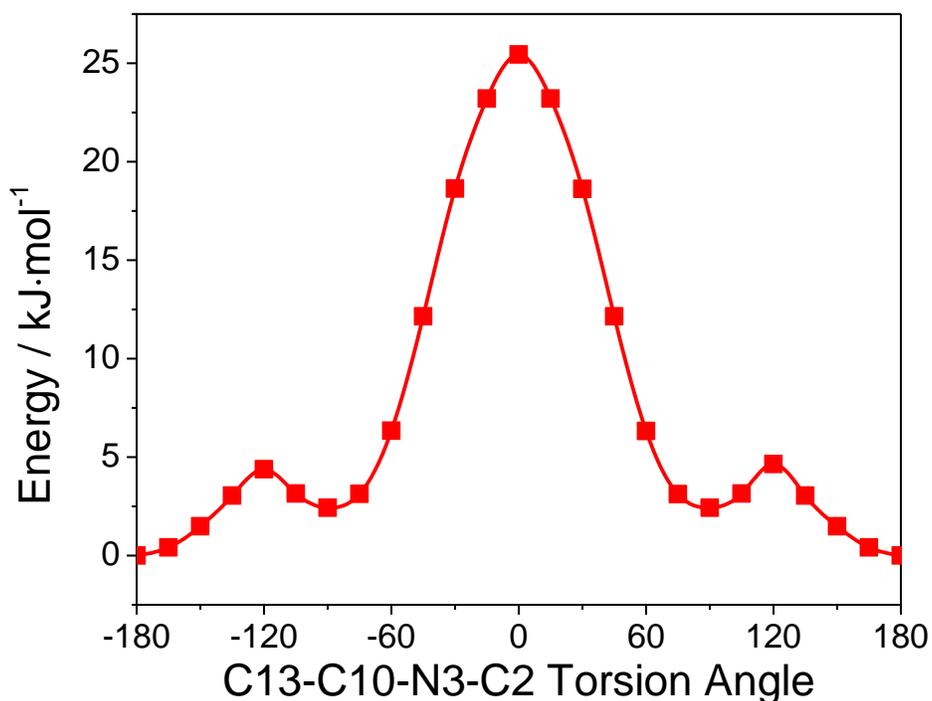
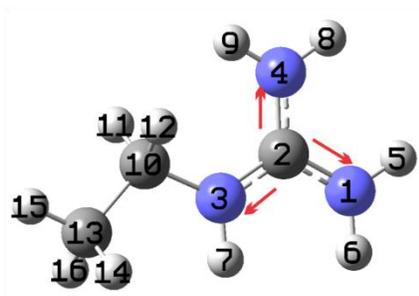


Figure 6. 3. The calculated EG potential energy surface along the C13-C10-N3-C2 torsion angle.

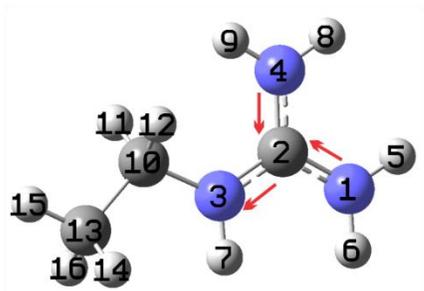
6.3.3 Calculated first electronic excited state structure of EG

The equilibrium geometry displacement of the first electronic excited state, and the first electronic excited state force along each vibrational normal coordinate of the electronic ground

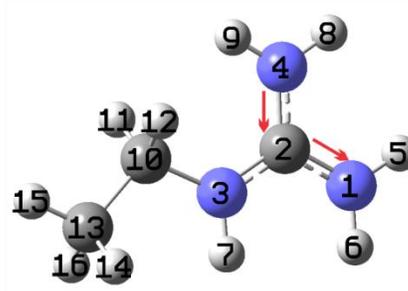
state at the ground state equilibrium geometry are related to the Raman scattering cross section as shown below. Generally, along a vibrational normal coordinate, the larger the vertical transition first electronic excited state force, or the larger the nuclear displacement of the first electronic excited state equilibrium geometry from the electronic ground state equilibrium geometry, the larger the corresponding vibrational mode resonance Raman scattering cross section.



CN₃ symmetric stretching



CN₃ asymmetric stretching



CN₃ asymmetric stretching'

Figure 6. 4. Definitions of CN₃ stretching vibrations in EG, Arg and AAO.

In order to elucidate the resonance Raman enhancement mechanism of the EG vibrations, we calculated the first electronic excited state equilibrium geometry of EG and the first electronic excited state force along each vibrational normal coordinate of the electronic ground state at the electronic ground state equilibrium geometry using the time-dependent DFT method in the Gaussian 09 software package, as described above in the experimental section.

The equilibrium geometry of the first electronic excited state of EG significantly differs from that of the electronic ground state. The C2-N3 bond length increases by 0.036 Å (2.7%); the C2-N1 bond increases by 0.075 Å (5.6%) and the C2-N4 bond increases by 0.078 Å (5.9%) (Table 6.5). The elongations of these CN bonds are presumably responsible for the large resonance Raman enhancement of the CN₃ stretching vibrations (Figure 6.4) in the UVRR spectra of EG. Tables 6.1 – 6.4 list those vibrations that involve more than 5% CN₃ stretching contributions for gdn, EG, Arg and AAO.

Table 6.6 shows the calculated atomic charges of EG in both the electronic ground and first excited states. The atomic charge differences between the electronic ground and first excited states clearly indicate that significant electron transfers occur between the electronic ground state and the first electronic excited state. This electron transfer causes the pyramidalization of the central carbon atom, which is consistent with the previous calculations for gdn by Hudson and coworkers using the semiempirical INDO/S method.²¹ Also, the first electronic excited state structure of AB₃ molecules was first examined and predicted to be somewhat pyramidal by Walsh in the 1950s.⁴³

In the electronic ground state, the C2 carbon atom is sp² hybridized, and the perpendicular p orbital of the C2 carbon overlaps with the three p orbitals of the three nitrogen atoms. In the first electronic excited state, the C2 carbon atom carries more electron density. This increased electron density of the C2 carbon in the excited state causes orbital rehybridization from sp² to sp³-like, which results in pyramidalization. The sp³-like hybridization of the central C2 carbon atom also explains the elongations of the three gdn group CN bonds; σ bonds involving sp³ hybridized orbitals are longer than those involving sp² hybridized orbitals.

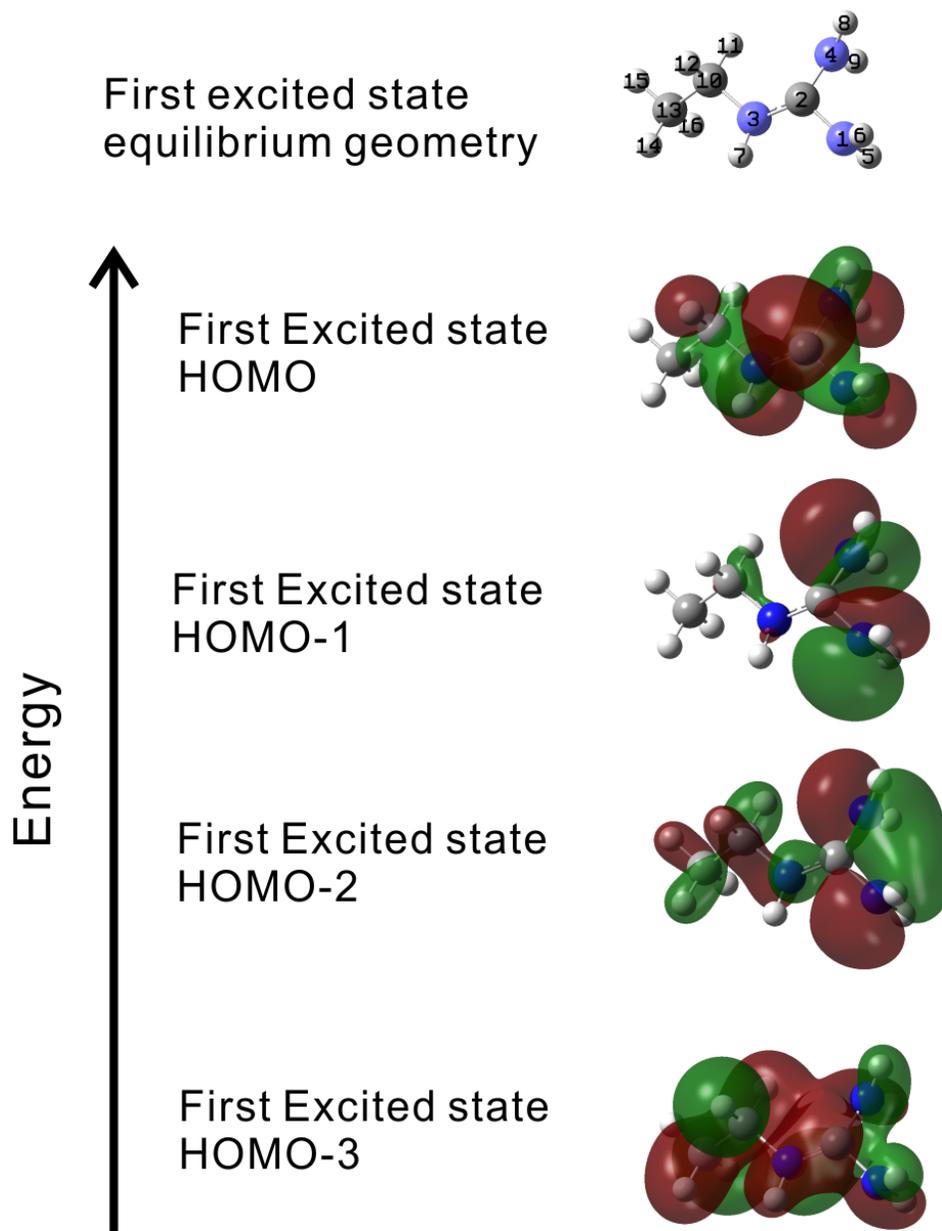


Figure 6. 5. The calculated structure and molecular orbitals of EG in the first electronic excited state.

The pyramidalization of the C2 carbon in the first excited state results in a non-planar CN_3 geometry that gives rise to a large displacement along the CN_3 out-of-plane deformation normal coordinate. The overtone of this CN_3 out-of-plane deformation appears as a strong band in all compounds studied here. Recently, we studied the excitation profiles of Arg and found that

the overtone of the CN₃ out-of-plane vibration is enhanced as the excitation wavelength is tuned into resonance.³⁰

Figure 6.5 shows the calculated EG first electronic excited state structure and HOMOs. In the first electronic excited state, the C2-N1 and C2-N4 bonds are elongated more than the C2-N3 bond (Table 6.5). The elongations of these bonds diminish the overlap of the two perpendicular p orbitals of the N1 and N4 nitrogen atoms with the p orbital of the C2 carbon. The C2-N1 and C2-N4 bonds rotate by ~90° (Figure 6.5), enabling the two p orbitals to partially overlap; however, the two nitrogen atoms are too far away from each other (2.247 Å) to form a normal σ bond. The resulting overlap is not as efficient as the overlap of the C2-N3 p orbitals with normal π-bond character. Therefore, the molecular orbital of this N1-N4 “σ bond” has a higher energy than that of the C2-N3 normal π-bond. The four molecular orbitals formed from the four p orbitals (one from the central carbon and three from the nitrogen atoms) in the first electronic excited state occur in an increasing energy order: the bonding π-orbital between the C2 carbon and the N3 nitrogen, the in-phase overlap between the N1 and N4 nitrogen p orbitals, the out-of-phase overlap between the N1 and N4 nitrogen p orbitals, and the anti-bonding π-orbital between the C2 carbon and the N3 nitrogen.

6.3.4 Calculation of the UVRR spectra of EG and EG-d

Using time-dependent Raman theory,⁴⁴⁻⁴⁷ under the Born-Oppenheimer approximation and the short time dynamics approximation, Heller et. al. predicted that, for the resonance and preresonance Raman cases, the square of the Raman polarizability of the *k*th vibrational normal mode, $|\alpha_k|^2$, is proportional to the square of the force (V_k) on the first electronic excited state

potential energy surface along that particular vibrational normal coordinate k at the electronic ground state equilibrium geometry. The Raman scattered intensity also increases with the excitation frequency to the fourth power, $\omega_1(\omega_1 - \omega_k)^3$:

$$I_k \sim \omega_1(\omega_1 - \omega_k)^3 \frac{V_k^2}{\omega_k} \quad (2)$$

where ω_1 and ω_k are the excitation frequency and the k th vibrational frequency respectively. Using eq. (2), the relative resonance Raman intensity of each vibrational normal mode of EG and EG-d was calculated from the first electronic excited state force along each vibrational normal mode that was obtained from DFT and TD-DFT calculations. Assuming a Lorentzian band shape with a FWHM of 20 cm^{-1} for each mode, the resonance Raman spectra of EG and EG-d were calculated and are shown in the bottom panel of Figure 6.6. For comparison, the experimental spectra are shown in the top panel of Figure 6.6.

The calculated spectra from eq. (2) moderately agree with those measured. In the calculated spectra, the bands involving CN_3 asymmetric stretching between 1500 cm^{-1} and 1700 cm^{-1} and the bands involving CN_3 symmetric stretching between 900 cm^{-1} and 1200 cm^{-1} are resonance enhanced, which is expected from the calculated large bond length increase along the three CN bonds of the gdn group.

The experimentally observed 1466 cm^{-1} EG and 1451 cm^{-1} EG-d overtone bands are missing in the calculated spectra. Eq. (2) only describes the fundamental vibrational band intensities and intrinsically does not include overtone or combination band intensities. It is evident from Figure 6.6 that the 1180 cm^{-1} band, that involves C10-N3 stretching, is resonance enhanced. From our experimental UVRR spectra, we find that this band can be used to determine the hydration state of the Arg side chains in peptides and proteins, as discussed below.

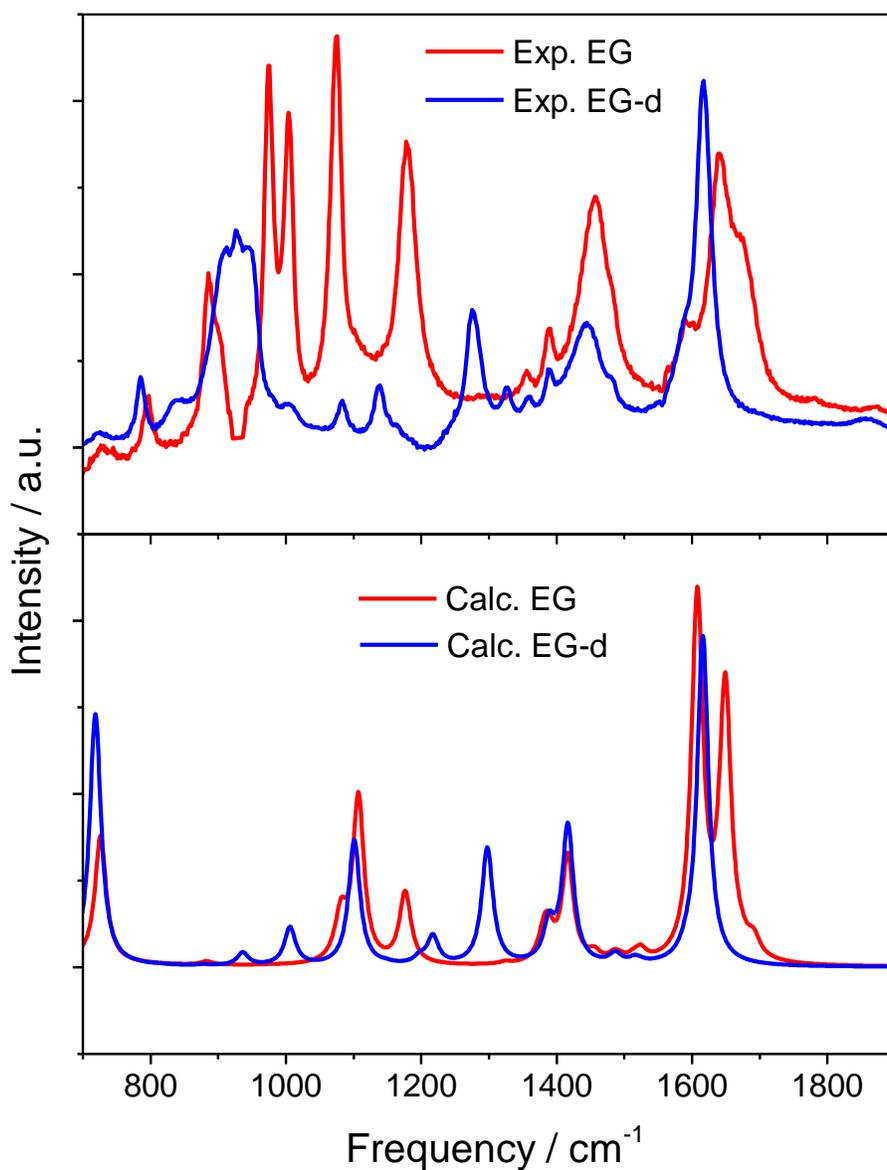


Figure 6. 6. 204 nm excited UVRR spectra of EG and EG-d (top panel) and calculated resonance Raman spectra of EG and EG-d (bottom panel).

The calculated spectra are the weighted average, based on the population fraction of each conformer, of the calculated spectra of the three EG conformers. See text for more details.

If no vibrational frequency change or Duschinsky rotation occurs between the electronic ground state and the first excited state, it is straightforward to rewrite eq. (2) as:

$$I_k \sim \omega_1(\omega_1 - \omega_k)^3(\omega_k^2 \Delta_k^2) \quad (3)$$

where Δ_k is the displacement of the equilibrium geometry of the first electronic excited state from the electronic ground state along the k th dimensionless vibrational normal coordinate. Eq. (3) is widely used in the interpretation and prediction of resonance Raman spectra. For gdn and gdn-d, the strongly resonance enhanced ground state CN₃ symmetric stretching vibration remains almost identical to that in the first electronic excited state (Table 6.7). Thus eq. (3) is valid to analyze the enhancement of this particular mode. However, for EG and EG-d, the excited state vibrational normal modes change significantly from those of the ground state, indicating that a significant Duschinsky rotation is occurring for these vibrations (Table 6.8). Obviously the first electronic excited state vibrational frequencies will differ from those in the electronic ground state. Therefore, the applicability of eq. (3) for EG and EG-d to interpret resonance Raman spectra becomes questionable.

Table 6. 7. First electronic excited state gdn and gdn-d vibrational normal modes involving at least 5% contribution from CN₃ stretching

	Exp (cm ⁻¹)	Calc (cm ⁻¹)	PED (%) [*]
gdn		1631	N1-H ₂ scs (69), C2-N ₃ as str (10), N4-H ₂ scs (8)
		1469	C2-N ₃ as str (59), H7-N3-C2-N4 tor (12), N1-H ₂ scs (6), C2-N3N4 scs (6), H10-N4-C2-N1 tor (6)
		1243	C2-N3N4 rock (35), N1-H ₂ rock (27), C2-N ₃ as str' (21), C2-N ₃ def (9)
		1035	C2-N ₃ as str' (44), N1-H ₂ rock (20), H7-N3-C2-N4 tor (11), H10-N4-C2-N1 tor (7), N4-H ₂ rock (5)
		933	C2-N ₃ s str (66), H10-N4-C2-N1 tor (19), C2-N ₃ s def (7)
gdn-d		1485	C2-N ₃ as str (65), D7-N3-C2-N4 tor (12), C2-N3N4 scs (5), D10-N4-C2-N1 tor (5), N1-D ₂ scs (5)
		1256	C2-N ₃ as str' (25), N4-D ₂ scs (24), N3-D ₂ scs (24), C2-N3N4 rock (11), D7-

			N3-C2-N4 tor (6)
		1090	C2-N3N4 rock (35), C2-N ₃ as str' (17), C2-N ₃ def (9), N4-D ₂ scs (9), N3-D ₂ scs (9), D10-N4-C2-N1 tor (9), N1-D ₂ rock (8)
		975	C2-N ₃ def (35), N4-D ₂ rock (23), D10-N4-C2-N1 tor (11), D7-N3-C2-N4 tor (10), N3-D ₂ rock (9), C2-N ₃ s str (5)
		850	N1-H ₂ rock (42), C2-N ₃ as str' (21), D10-N4-C2-N1 tor (10), C2-N3N4 rock (8), D7-N3-C2-N4 tor (5)
		842	C2-N ₃ s str (61), N4-D ₂ rock (16), N3-D ₂ rock (7)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking; wag: wagging; twist:

twisting; ipb: in-plane bending; as def: asymmetric deformation; s def: symmetric deformation; tor: torsion;

Table 6. 8. First electronic excited state EG and EG-d vibrational normal modes involving at least 5% contribution from CN₃ stretching

	Exp (cm ⁻¹)	Calc (cm ⁻¹)	PED (%) [*]
EG		1542	C2-N ₃ as str (41), N3-H ipb (31), H9-N4-C2-N1 tor (7)
		1430	N3-H ipb (28), C2-N ₃ as str (28), C10-H ₂ wag (21)
		1378	C10-H ₂ wag (59), N3-H ipb (15), C2-N ₃ as str (5), C13-H ₃ s def (5)
		1185	C10-H ₂ rock (19), C2-N ₃ as str' (17), C13-H ₃ rock' (17), C2-N1N4 rock (9), N3-C10 str (9), N3-C2C10 ipb (8)
		1113	C2-N ₃ as str' (52), N3-C10 str (21), C2-N ₃ s str (6)
		1065	C10-C13 str (33), C2-N ₃ as str' (18), N3-C10 str (16), C13-H ₃ rock' (6), C2-N ₃ s str (6), C2-N1N4 rock (5), N3-H ipb (5)
		971	C2-N ₃ s str (41), C10-C13 str (33), C13-H ₃ rock' (7)
		884	C13-H ₃ rock' (25), C2-N ₃ s str (22), N3-C10 str (15), N3-C2C10 ipb (9)
EG-d		2104	D9-N4-C2-N1 tor (47), D6-N1-C2-N4 tor (16), N1-C2D5D6 s def (7), C2-N ₃ s str (5)
		1511	C10-H ₂ scs (66), C13-H ₃ as def' (12), C2-N ₃ as str (5)
		1500	C2-N ₃ as str (59), D9-N4-C2-N1 tor (8), C13-H ₃ as def (6), N3-C10 str (5),

			N3-D ipb (5)
		1486	C13-H ₃ as def (77), C10-H ₂ scs (8), C13-H ₃ rock (7), C2-N ₃ as str (5)
		1241	N1-D ₂ scs (63), C2-N ₃ s str (9), C2-N ₃ as str' (6), N4-D ₂ scs (5), D6-N1-C2-N4 tor (5)
		1057	C2-N ₃ as str' (31), N1-D ₂ scs (14), N3-C10 str (12), C10-C13 str (11), N4-D ₂ scs (7), C2-N1N4 rock (7), D6-N1-C2-N4 tor (5)
		997	C10-C13 str (26), C2-N ₃ s str (15), N1-D ₂ rock (8), N1-C2D5D6 s def (7), N4-D ₂ rock (7), C2-N ₃ as str' (6), D9-N4-C2-N1 tor (6), D6-N1-C2-N4 tor (6), N3-D ipb (5)
		964	H6-N1-C2-N4 tor (24), N4-D ₂ rock (14), N4-D ₂ rock (12), D9-N4-C2-N1 tor (9), C2-N ₃ s def (9), C10-C13 str (6), C2-N ₃ as str' (5), N1-C2D5D6 s def (5)
		896	N3-D ipb (27), C13-H ₃ rock' (24), N3-C10 str (19), C2-N ₃ s str (5)
		842	C2-N ₃ s str (31), C13-H ₃ rock' (12), C10-H ₂ rock (6), C2-N ₃ as str' (6), N3-C2C10 ipb (5), N4-D ₂ rock (5)

* as str: asymmetric stretching; s str: symmetric stretching; scs: scissoring; rock: rocking; wag: wagging; twist: twisting; ipb: in-plane bending; as def: asymmetric deformation; s def: symmetric deformation; tor: torsion;

6.3.5 UVRR spectra of EG in water and in acetonitrile

To study the effect of the environment on the UVRR spectrum of the Arg side chains, we examined the 204 nm excited UVRR spectra of EGPC in water and in acetonitrile (Figure 6.7a). Most of the bands in acetonitrile are significantly more intense than those in water. To understand this phenomenon, we measured the UV absorbance spectra of EG in both water and acetonitrile (Figure 6.7b) and found only small differences. The difference spectrum shows an

additional weak, broad band at ~ 200 nm. At 204 nm, the absorbance of EG in acetonitrile is about 15% greater than that in water.

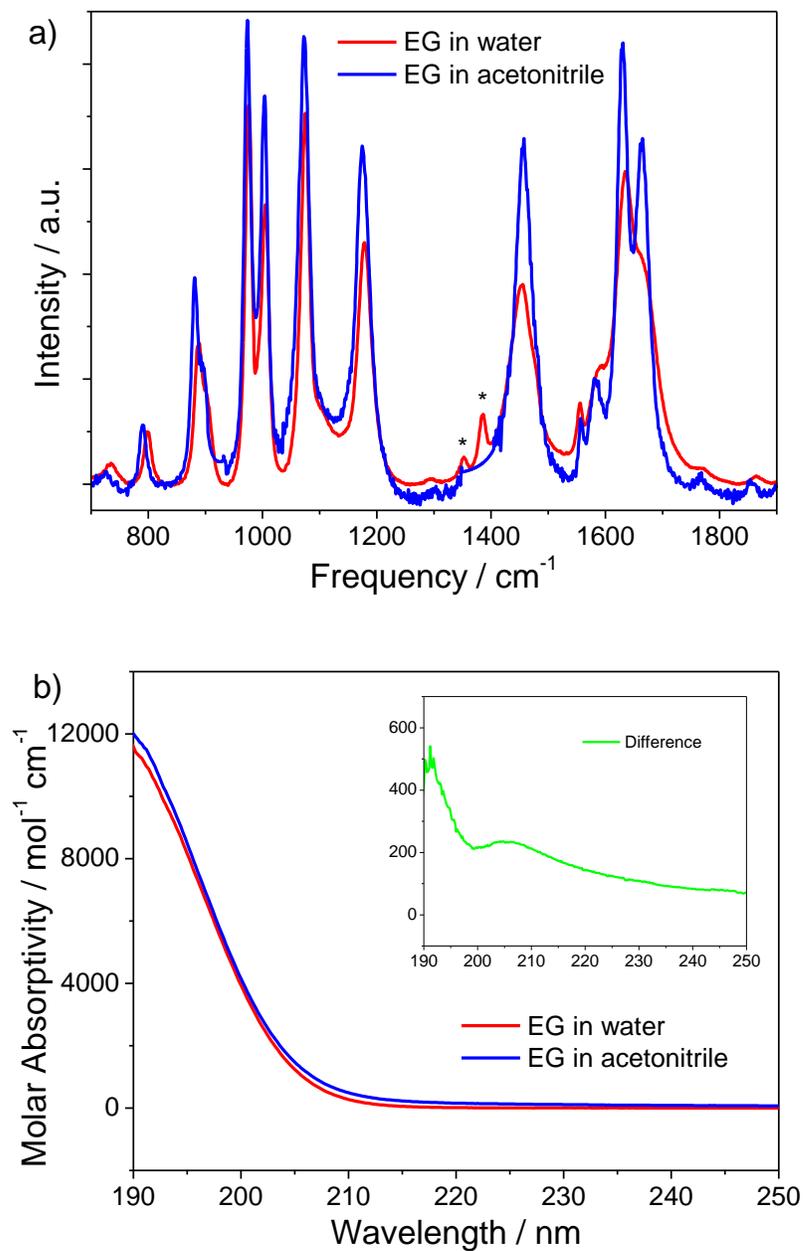


Figure 6. 7. a) 204 nm excited UVRR spectra and b) UV absorbance spectra of EG in water and in acetonitrile. The bands marked with an asterisk in the spectrum of EG in water are missing in the spectrum of EG in acetonitrile due to the subtraction of the solvent.

The Raman band frequencies of EG in acetonitrile are slightly downshifted compared to those in water. Particularly, the 1180 cm^{-1} band in water downshifts to 1174 cm^{-1} in acetonitrile. This Arg side chain band occurs in a region that generally does not overlap with the amide vibrational bands in the UVRR spectra of proteins and peptides. Thus, this band can be conveniently used to probe the local environment and hydrogen bonding state of the Arg side chains. The other Arg side chain bands, such as the CN_3 asymmetric stretching bands between 1500 cm^{-1} to 1700 cm^{-1} , unfortunately, are overlapped by amide bands.

Our observed frequency increase of the UVRR 1174 cm^{-1} band of EG in acetonitrile to 1180 cm^{-1} in water is consistent with the IR study of Braiman et. al.⁴⁸ With acetate as a counter ion, they observed a higher frequency for this band in nonpolar solvents than in polar solvents, such as H_2O , methanol and DMSO. The 1174 cm^{-1} band has contributions from the gdn group NH_2 bending vibrations (Table 6.2). The hydrogen bond between the gdn NH_2 and the hydrogen bonding acceptor molecule makes the NH_2 group more difficult to bend and narrows the potential well, thus increasing the bending force constant. Therefore, stronger NH_2 hydrogen bonding would upshift the frequency of this band.⁴⁹ Acetate is a strong hydrogen bond acceptor and can form a strong hydrogen bond with the EG NH_2 in nonpolar solvents resulting in an increased frequency of the 1174 cm^{-1} band. Therefore, the higher frequency observed in nonpolar solvents compared to polar solvents is due to the interaction of EG with the acetate counter ion, not due to the decreased interaction with the nonpolar solvent molecules themselves.

In polar solvents, they observed a lower frequency of this band compared to nonpolar solvents because the polar solvent molecules form weaker hydrogen bonds with the EG NH_2 than does the acetate counter ion. Since large numbers of polar solvent molecules are present, they compete with the acetate counter ions and prevent the acetate from forming hydrogen bonds with

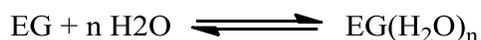
the EG NH₂. Therefore, the lower frequencies they observed in polar solvents most likely reflect the weaker hydrogen bonding between the polar solvent molecules and the EG NH₂. These arguments are supported by the fact that the frequency dependence of this band on the identity of the counter ion only occurs in the solid state or in nonpolar solvents, while in polar solvents this dependence disappears.⁴⁸

In our experiment, the perchlorate counter ion is a weak hydrogen bond acceptor and is expected to have little interaction with the EG NH₂ in either water or acetonitrile; thus, it has little impact on the EG frequency even in the less polar acetonitrile solvent. Therefore, the frequency shift of the 1174 cm⁻¹ band observed here reflects the differences in the interactions of EG with water and acetonitrile.

6.3.6 Hydration of guanidinium group

We measured the 204 nm excited UVRR spectra of EG and AAO in different water-acetonitrile mixtures. For EG, as the water mole fraction increases, the frequency upshifts from 1174 cm⁻¹ to 1180 cm⁻¹, while the cross section monotonically decreases from $6.5 \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$ to $4.2 \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$; for AAO, the frequency upshifts from 1170 cm⁻¹ to 1176 cm⁻¹, while the cross section monotonically decreases from $9.1 \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$ to $5.7 \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$.

We utilize a simple binding model to rationalize the Raman spectral dependences on EG hydration. Upon addition of water into the EG acetonitrile solution, the free EG is bound by water and hydrated:



Assuming ideal solution behavior, by definition, the apparent association constant is:

$$K = \frac{[\text{EG}(\text{H}_2\text{O})_n]}{[\text{EG}][\text{H}_2\text{O}]^n} \quad (4)$$

For a water mole fraction of x , the relative fraction of dehydrated EG is $\eta(x) = \frac{[\text{EG}]}{[\text{EG}]_{\text{Total}}}$; then eq.

(4) is:

$$\eta(x) = \frac{1}{1 + K[\text{H}_2\text{O}]^n} \quad (5)$$

It is possible to obtain an explicit expression of $[\text{H}_2\text{O}]$ in terms of x from the densities and molecular weights of water and acetonitrile and assuming ideal solution behavior:

$$[\text{H}_2\text{O}]/\text{mol}\cdot\text{L}^{-1} = \frac{1.06 \times 10^3 x}{55.4 - 36.2x} \quad (6)$$

The apparent Raman cross section can be treated as a sum of the cross sections of free and bound EG:

$$\sigma(x) = (1 - \eta(x)) \cdot \sigma_{\text{EG}(\text{H}_2\text{O})_n} + \eta(x) \cdot \sigma_{\text{EG}} \quad (7)$$

By combining eqs. (4), (5) and (6), eq. (7) becomes:

$$\sigma = \sigma_{\text{EG}(\text{H}_2\text{O})_n} + \frac{\sigma_{\text{EG}} - \sigma_{\text{EG}(\text{H}_2\text{O})_n}}{1 + K \left(\frac{1.06 \times 10^3 x}{55.4 - 36.2x} \right)^n} \quad (8)$$

Fitting the EG data to eq. (8) gives $\sigma_{\text{EG}} = (6.6 \pm 0.2) \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$, $\sigma_{\text{EG}(\text{H}_2\text{O})_n} = (4.1 \pm 0.2) \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$, $n = 1.8 \pm 0.5$, and $K = 0.19 \pm 0.11$ (Figure 6.8). In the case of AAO, $\sigma_{\text{AAO}} = (8.8 \pm 0.3) \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$, $\sigma_{\text{AAO}(\text{H}_2\text{O})_n} = (5.9 \pm 0.4) \times 10^{-26} \text{cm}^2 \cdot \text{molecule}^{-1} \cdot \text{sr}^{-1}$, $n = 2.1 \pm 0.9$, and $K = 0.09 \pm 0.10$ (Figure 6.8).

This hydration model is highly simplified. This model can only distinguish between hydrated and dehydrated EG in solution. n , representing the phenomenological hydration level

of the molecule, is ~ 1.8 for EG and ~ 2.1 for AAO suggesting that the gdn group binds ~ 2 water molecules. This is consistent with previous reports on the hydration of the gdn compounds.²³⁻²⁷

The formation of gdn dimers^{23-25,50} possibly accounts for this low hydration level.

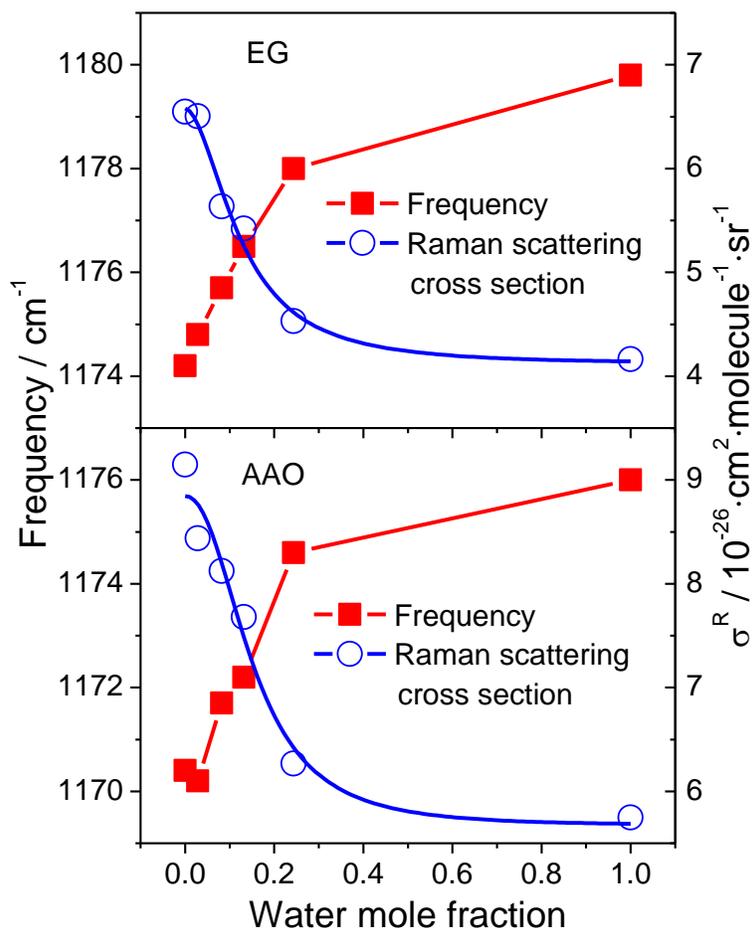


Figure 6. 8. The integrated Raman scattering cross sections and frequencies of the $\sim 1170 \text{ cm}^{-1}$ band measured from the 204 nm excited UVRR spectra of EG and AAO in different acetonitrile-water solutions.

6.3.7 Different environments in peptides

Previously, we found that the $\sim 1170 \text{ cm}^{-1}$ Arg bands in the UVRR spectra of two similar alanine based 21 residue peptides, AP ($A_8RA_4RA_4RA_2$) and AEP ($A_9RA_3EA_4RA_2$), consist of at least

two components ascribed to their different hydration states. (see Figure 5.3)⁵¹ We modeled the hydration dependence of the $\sim 1170 \text{ cm}^{-1}$ Arg UVRR band by assuming that the Raman band shape $S(\nu)$ has contributions from two different hydrated forms:

$$S(\nu) = w_{\text{FH}}S_{\text{FH}}(\nu) + w_{\text{DH}}S_{\text{DH}}(\nu) \quad (9)$$

where w_{FH} and w_{DH} are the relative fractions ($w_{\text{FH}} + w_{\text{DH}} = 1$) of the fully hydrated and dehydrated forms, and $S_{\text{FH}}(\nu)$ and $S_{\text{DH}}(\nu)$ are the Raman spectra of these two forms. The dehydrated form occurs only within the α -helical conformation with a fractional content of ϕ_{DH} :

$$w_{\text{DH}} = f_{\text{H}}\phi_{\text{DH}} \quad (10)$$

where f_{H} is the helical fraction. The fully hydrated form occurs in the PPII-like conformations because the Arg side chains should be fully hydrated in these extended conformations. It also appears that some Arg side chains are fully hydrated in the α -helical form. In the α -helical conformation the fully hydrated fraction of the Arg side chain is $\phi_{\text{FH}} = 1 - \phi_{\text{DH}}$, therefore:

$$w_{\text{FH}} = 1 - w_{\text{DH}} = f_{\text{U}} + f_{\text{H}}\phi_{\text{FH}} = f_{\text{U}} + f_{\text{H}}(1 - \phi_{\text{DH}}) \quad (11)$$

where $f_{\text{U}} = 1 - f_{\text{H}}$ is the fraction of PPII-like conformation.

The $\sim 1170 \text{ cm}^{-1}$ Arg bands of AP and AEP can be fitted with two Gaussians, as shown in Figure 6.9, one of which derives from the fully hydrated component with a higher frequency; a second Gaussian with a lower frequency is found to model the dehydrated component. We globally fitted the Arg bands of AP and AEP with two Gaussians of identical bands widths and frequencies. The fitted dehydrated component appears at 1165 cm^{-1} for both AP and AEP peptides, while the hydrated component is found at 1173 cm^{-1} .

We can estimate the fraction of the dehydrated component, ϕ_{DH} from the ratio:

$$\frac{I_{\text{DH}}}{I_{\text{FH}}} = \frac{w_{\text{DH}}\sigma_{\text{DH}}}{w_{\text{FH}}\sigma_{\text{FH}}} = \frac{f_{\text{H}}\phi_{\text{DH}}\sigma_{\text{DH}}}{[f_{\text{U}} + f_{\text{H}}(1 - \phi_{\text{DH}})]\sigma_{\text{FH}}} \quad (12)$$

Rearranging eq. (12), we get:

$$\phi_{\text{DH}} = \frac{I_{\text{DH}}\sigma_{\text{FH}}}{f_{\text{H}}(I_{\text{FH}}\sigma_{\text{DH}} + I_{\text{DH}}\sigma_{\text{FH}})} \quad (13)$$

where I_{DH} and I_{FH} are the integrated band intensities of the dehydrated and fully hydrated components respectively. $\sigma_{\text{DH}} = \int S_{\text{DH}}(\nu)d\nu$ and $\sigma_{\text{FH}} = \int S_{\text{FH}}(\nu)d\nu$ are the integrated Raman scattering cross sections of the $\sim 1170 \text{ cm}^{-1}$ band, which are assumed to be equal to the cross sections of AAO in acetonitrile and in water, respectively. I_{DH} and I_{FH} are obtained by fitting the measured 1170 cm^{-1} UVRR band with two Gaussian bands.

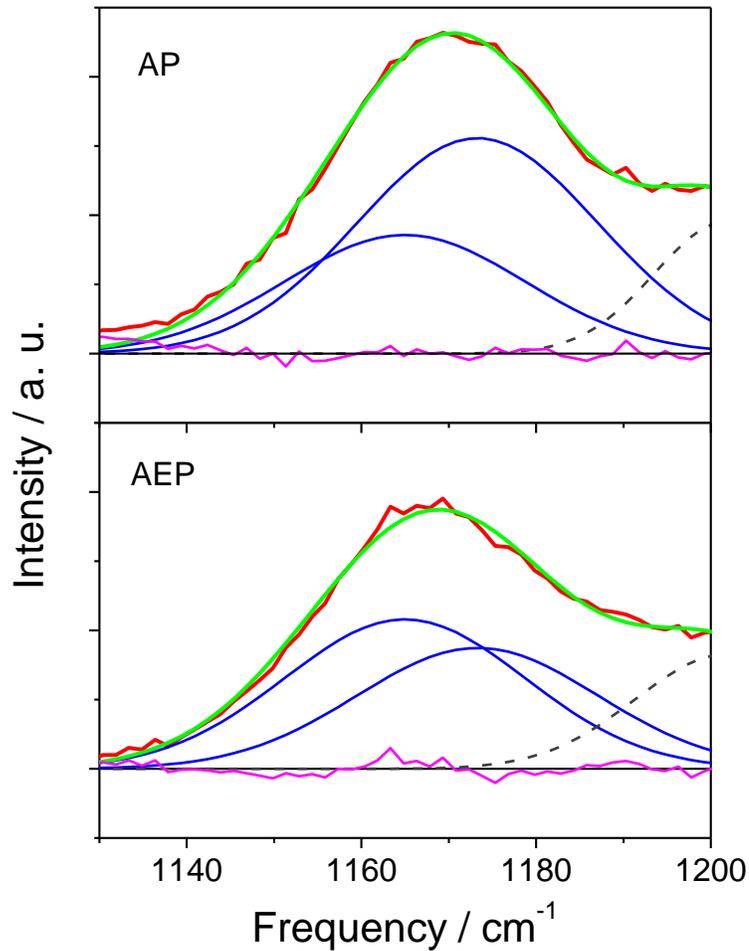


Figure 6. 9. The fitting of the UVRR bands of the Arg side chains in AP and AEP using two Gaussian bands. The dashed bands are tentatively assigned to the backbone AmIII₃ bands of turn structures.

The dehydrated fractions, w_{DH} , are estimated to be 0.27 for AP and 0.45 for AEP. Using our calculated α -helical fraction, f_{H} , of 0.51 at 4 °C for AEP and AP, we further estimate from eq. (13) that, on average, $\phi_{\text{DH}} = 0.53$ for the three Arg side chains of AP, while $\phi_{\text{DH}} = 0.89$ for the two Arg side chains of AEP.

Table 6. 9. Fractions of different Arg hydration states, equilibrium constants and corresponding Gibbs free energies in AP and AEP peptides

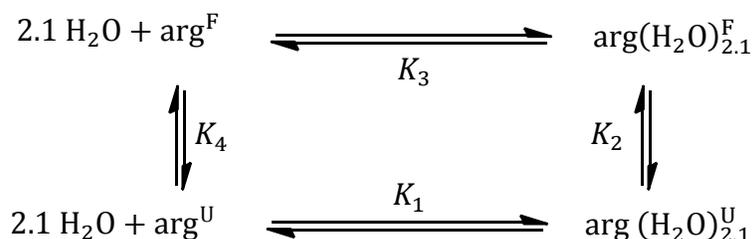
Peptide	AP	AEP
fraction of Arg ^{U^a}	0.0 ± 0.2	0.0 ± 0.5
fraction of Arg (H ₂ O) _{2.1} ^{U^a}	0.49 ^b	0.49 ^b
fraction of Arg(H ₂ O) _{2.1} ^{F^a}	0.24 ± 0.23	0.06 ± 0.5
fraction of Arg ^{F^a}	0.27 ± 0.23	0.45 ± 0.5
$K_1/\text{M}^{-2.1\text{c}}$	0.09 ± 0.10	0.09 ± 0.10
K_2	0.5 ± 0.5	0.1 ± 1
$K_3/\text{M}^{2.1\text{c}}$	$(5 \pm 5) \times 10^3$	$(4 \pm 30) \times 10^4$
K_4	0.004 ± 0.008	0.003 ± 0.03
$\Delta G_1/\text{kJ} \cdot \text{mol}^{-1}$	6.0 ± 2.7	6.0 ± 2.7
$\Delta G_2/\text{kJ} \cdot \text{mol}^{-1}$	1.8 ± 2.4	5 ± 22
$\Delta G_3/\text{kJ} \cdot \text{mol}^{-1}$	-21 ± 2.2	-26 ± 22
$\Delta G_4/\text{kJ} \cdot \text{mol}^{-1}$	13 ± 4	15 ± 31

^a: These Arg states are defined in Scheme 6. 1.

^b: the helical fraction of AP is estimated from ref. 3; the AEP helical fraction is identical to AP, see ref. 51 for more details.

^c: as defined in Scheme 6. 1, K_1 has a unit of $\text{M}^{-2.1}$ due to the binding of 2.1 H₂O molecules; K_3 has a unit of $\text{M}^{2.1}$ due to the releasing of 2.1 H₂O molecules.

We can identify four Arg side chain states in AP or AEP: 1) the dehydrated Arg side chain in the PPII-like unfolded state, Arg^{U} ; 2) the hydrated Arg side chain in the PPII-like unfolded state, $\text{Arg}(\text{H}_2\text{O})_{2.1}^{\text{U}}$; 3) the hydrated Arg side chain in the α -helix folded state, $\text{Arg}(\text{H}_2\text{O})_{2.1}^{\text{F}}$; and 4) the dehydrated Arg side chain in the α -helix folded state, Arg^{F} . Scheme 6. 1 shows the thermodynamic cycle between these side chain states. Table 6.9 gives the fractions of these four Arg side chain states, equilibrium constants and corresponding Gibbs free energies of the transition between these four states.



Scheme 6. 1. Thermodynamic cycle that describes the transition among four Arg states in the peptides AP and AEP

It appears that the Arg side chains in the α -helical conformation are mainly dehydrated. These Arg side chains likely pack against the α -helix backbone where they are shielded from water. This packing also protects the α -helix intrapeptide hydrogen bonds from water, which increases the α -helix stability. This configuration is consistent with Garcia's MD simulations.⁵² This packing process releases bound water molecules from the Arg side chains and eliminates the unfavorable interactions between water and Arg side chains. Thus, it is favorable for the peptides to fold into α -helix conformations. The Arg side chains in AEP are less hydrated than those in AP, most likely due to the presence of more α -helix-turn- α -helix conformations⁵¹ that may further shield the Arg side chains from the access of water.

6.4 CONCLUSION

We used UVRR spectroscopy and DFT calculations to examine resonance Raman enhancement in gdn, EG, Arg and AAO. We examined the calculated ground and first excited state structures and assigned most of the UVRR bands above 1100 cm^{-1} . We have examined the origin of resonance enhancement of the gdn side chain vibrations. The UVRR scattering cross section and the frequency of the $\sim 1170\text{ cm}^{-1}$ Arg band in peptides and proteins provide monitors for the local environment of the gdn group. To rationalize the cross section dependence on hydration, a simple hydration model was applied to EG and AAO. By applying this model to two polyAla peptides, AP and AEP, both containing Arg residues, we were able to identify different hydration states for different conformations. It was found that the AEP Arg side chain in the α -helical conformation is more dehydrated than the AP Arg side chain in its α -helical conformation.

6.5 ACKNOWLEDGEMENT

We thank Dr. Bhavya Sharma for helpful discussions. This work was supported by NIH Grant 5RO1 EB002053 and NIH 1RO1 EB009089, and in part by the University of Pittsburgh Center for Simulation and Modeling through the supercomputing resources provided.

6.6 BIBLIOGRAPHY

- (1) Asher, S. A. *Anal. Chem.* **1993**, *65*, A201.
- (2) Asher, S. A. *Anal. Chem.* **1993**, *65*, A59.
- (3) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 8074.
- (4) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 4076.
- (5) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775.
- (6) Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433.
- (7) Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2006**, *128*, 13789.
- (8) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2006**, *110*, 1928.
- (9) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. *J. Phys. Chem. Lett.* **2011**, *2*, 334.
- (10) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. *Chem. Rev.* **2012**, *112*, 2604.
- (11) Asher, S. A.; Ludwig, M.; Johnson, C. R. *J. Am. Chem. Soc.* **1986**, *108*, 3186.
- (12) Fodor, S. P. A.; Copeland, R. A.; Grygon, C. A.; Spiro, T. G. *J. Am. Chem. Soc.* **1989**, *111*, 5509.
- (13) Chi, Z. H.; Asher, S. A. *J. Phys. Chem. B* **1998**, *102*, 9595.
- (14) Sweeney, J. A.; Asher, S. A. *J. Phys. Chem.* **1990**, *94*, 4784.
- (15) Miura, T.; Takeuchi, H.; Harada, I. *J. Raman Spectrosc.* **1989**, *20*, 667.
- (16) Miura, T.; Takeuchi, H.; Harada, I. *Biochemistry* **1991**, *30*, 6074.
- (17) Takeuchi, H. *Biochemistry* **1997**, *36*, 10993.
- (18) Takeuchi, H. *Biopolymers* **2003**, *72*, 305.
- (19) Hildebrandt, P. G.; Copeland, R. A.; Spiro, T. G.; Otlewski, J.; Laskowski, M.; Prendergast, F. G. *Biochemistry* **1988**, *27*, 5426.
- (20) Asher, S. A.; Murtaugh, J. L. *Appl. Spectrosc.* **1988**, *42*.
- (21) Sension, R. J.; Hudson, B.; Callis, P. R. *J. Phys. Chem.* **1990**, *94*, 4015.
- (22) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849.
- (23) Mason, P. E.; Dempsey, C. E.; Neilson, G. W.; Brady, J. W. *J. Phys. Chem. B* **2005**, *109*, 24185.
- (24) Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Barnes, A. C.; Cruickshank, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 4557.
- (25) Mason, P. E.; Neilson, G. W.; Kline, S. R.; Dempsey, C. E.; Brady, J. W. *J. Phys. Chem. B* **2006**, *110*, 13477.
- (26) Hunger, J.; Niedermayer, S.; Buchner, R.; Hefter, G. *J. Phys. Chem. B* **2010**, *114*, 13617.
- (27) McQuinn, K.; McIndoe, J. S.; Hof, F. *Chem. Eur. J.* **2008**, *14*, 6483.
- (28) Garc á, A. E.; Stiller, L. *J. Comput. Chem.* **1993**, *14*, 1396.
- (29) White, S. H.; Wimley, W. C. *Annu. Rev. Biophys. Biomolec. Struct.* **1999**, *28*, 319.
- (30) Sharma, B.; Asher, S. A. *J. Phys. Chem. B* **2011**, *115*, 5659.

- (31) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. *Appl. Spectrosc.* **2005**, *59*, 1541.
- (32) Dudik, J. M.; Johnson, C. R.; Asher, S. A. *J. Chem. Phys.* **1985**, *82*, 1732.
- (33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2010.
- (34) Groner, P. In *Handbook of Vibrational Spectroscopy*; John Wiley & Sons, Ltd, 2006; Vol. III.
- (35) McIntosh, D. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **2009**, *125*, 177.
- (36) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.
- (37) Tang, W.; Sanville, E.; Henkelman, G. *J. Phys.: Condens. Matter* **2009**, *21*, 084204.
- (38) Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. *J. Comput. Chem.* **2007**, *28*, 899.
- (39) Henkelman, G.; Arnaldsson, A.; Jónsson, H. *Comput. Mater. Sci.* **2006**, *36*, 354.
- (40) Magalhães, A. L.; Gomes, J. A. N. F. *Int. J. Quantum Chem.* **1997**, *61*, 725.
- (41) Gobbi, A.; Frenking, G. *J. Am. Chem. Soc.* **1993**, *115*, 2362.
- (42) Mak, M. L.; Salpietro, S. J.; Enriz, R. D.; Csizmadia, I. G. *Can. J. Chem.* **2000**, *78*, 626.
- (43) Walsh, A. D. *J. Chem. Soc.* **1953**, 2301.
- (44) Heller, E. J. *J. Chem. Phys.* **1975**, *62*, 1544.
- (45) Heller, E. J.; Sundberg, R.; Tannor, D. *J. Phys. Chem.* **1982**, *86*, 1822.
- (46) Tannor, D. J.; Heller, E. J. *J. Chem. Phys.* **1982**, *77*, 202.
- (47) Lee, S.-Y.; Heller, E. J. *J. Chem. Phys.* **1979**, *71*, 4777.
- (48) Braiman, M. S.; Briercheck, D. M.; Kriger, K. M. *J. Phys. Chem. B* **1999**, *103*, 4744.
- (49) Vanderkooi, J. M.; Dashnau, J. L.; Zelent, B. *Biochim. Biophys. Acta. - Protein Proteomics* **2005**, *1749*, 214.
- (50) No, K. T.; Nam, K.-Y.; Scheraga, H. A. *J. Am. Chem. Soc.* **1997**, *119*, 12917.
- (51) Hong, Z.; Ahmed, Z.; Asher, S. A. *J. Phys. Chem. B* **2011**, *115*, 4234.
- (52) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2782.

7.0 SDS MONOMERS INDUCE XAO PEPTIDE POLYPROLINE II TO α -HELIX TRANSITION

This chapter has been submitted to *J. Phys. Chem. B*

Authors: Zhenmin Hong, Krishnon Damodaran and Sanford A. Asher

Address: Department of Chemistry, University of Pittsburgh, PA 15260

XAO peptide (Ac-X₂A₇O₂-NH₂, X: diaminobutyric acid, -CH₂CH₂NH₃⁺, O: ornithine, -CH₂CH₂CH₂NH₃⁺) in aqueous solution shows a predominantly polyproline II (PPII) conformation without any detectable α -helix-like conformations. Here we demonstrate by using CD, UV resonance Raman (UVRR) and NMR spectroscopy that sodium dodecyl sulfate (SDS) monomers bind to XAO and induce formation of α -helix-like conformations. The stoichiometry and the association constants of SDS and XAO were determined from the XAO-SDS diffusion coefficients measured by pulsed field gradient NMR. We developed a model for the formation of XAO-SDS aggregate α -helix-like conformations. Using UVRR spectroscopy, we calculated the Ramachandran ψ angle distributions of aggregated XAO peptides. We resolved α -, π - and 3_{10} -

helical conformations and a turn conformation. XAO nucleates SDS aggregation at SDS concentrations below the SDS critical micelle concentration. The XAO₄-SDS₁₆ aggregates have four SDS alkyl chains bound to each XAO to neutralize the four side chain cationic charges. We propose that the SDS alkyl chains partition into a hydrophobic core to minimize the hydrophobic area exposed to water. Neutralization of the flanking XAO charges enables α -helix formation. Four XAO-SDS₄ aggregates form a complex with an SDS alkyl chain dominated hydrophobic core and a more hydrophilic shell where one face of the α -helix peptide contacts the water environment.

7.1 INTRODUCTION

The aggregation and fibrillation of some intrinsically disordered proteins (IDPs)¹ appear to be involved in the development of neurodegenerative disorders such as Parkinson's disease, Alzheimer's disease and type II diabetes.²⁻⁴ It has been proposed that α -helix-like conformations are important intermediates during amyloid fibril formation.⁵⁻⁷ Some IDPs form α -helix-like conformations upon association with membranes through a mechanism that involves further aggregation and fibrillation. For example, α -synuclein (α S, associated with Parkinson's disease) upon binding to membranes or surfactant micelles forms α -helices that appears to mediate α S aggregation and fibrillation.⁸⁻¹¹ For amyloid β -peptide ($A\beta$, associated with Alzheimer's disease), the monomer peptide binds to surfactant micelles¹²⁻¹³ and lipid membranes,¹⁴⁻¹⁵ forming α -helix-like conformations that facilitate aggregation and fibrillation.¹⁶ Similar phenomena occur

for other IDPs, such as medin¹⁷ and islet amyloid polypeptide (IAPP, associated with type II diabetes).¹⁸⁻¹⁹

IDPs protofibrils or oligomeric aggregates are more toxic than are the insoluble fibrils.^{2,20} Recently, it was shown that the toxic IDPs protofibrils or oligomeric aggregates form pores in membranes, apparently sharing a common mechanism with antimicrobial peptides (AMPs).²¹⁻²³ When AMPs are associated with anionic lipid membranes, many of them also form α -helix-like conformations²⁴⁻²⁸ that display antimicrobial activity.²⁹⁻³¹ An understanding of the α -helix formation mechanism of these peptides would help the understanding the overall fibrillation mechanism of these IDPs, as well as help in the design of AMPs.

In this paper, we study the interactions between an undecapeptide XAO (Ac-X₂A₇O₂-NH₂, MW=985, X: diaminobutyric acid, Dab, side chain: -CH₂CH₂NH₃⁺, O: ornithine, Orn, side chain: -CH₂CH₂CH₂NH₃⁺) with sodium dodecyl sulfate (SDS). The XAO peptide was chosen because: (1) it has a large content of hydrophobic residues (seven Ala) and multiple cationic residues (two Dab and two Orn), similar to many of those IDPs and AMPs. (2) In aqueous solution, monomeric XAO predominantly adopts a polyproline II (PPII) conformation,³²⁻³³ similar to that of some IDPs³⁴⁻³⁶ and AMPs³⁷. XAO does not exhibit any detectable α -helical content, presumably due to its short polyala repeat lengths and electrostatic repulsion between its flanking cationic Dab and Orn side chains. SDS was chosen due to its amphiphilic molecular structure. SDS micelles have been widely used as a model membrane to study interactions between IDPs / AMPs and lipid membranes.^{9,12-13,38-45}

We showed here that XAO adopts α -helix-like conformations upon aggregating with SDS. We used UV resonance Raman (UVRR) spectroscopy to determine the structure of the aggregated XAO.⁴⁶⁻⁴⁷ We measured the diffusion coefficients of XAO and SDS by employing

pulsed field gradient NMR (PFG-NMR). Based on the stoichiometry and association constants extracted from the diffusion coefficients, we propose a two-state model for the aggregation between XAO and SDS. This study demonstrates that SDS monomers bind to XAO, forming XAO-SDS aggregates that induce XAO α -helix-like conformations.

7.2 EXPERIMENTAL SECTION

7.2.1 Materials

The undecapeptide XAO was prepared by the Pittsburgh Peptide Facility by using a solid-state peptide synthesis method. SDS (99%), dodecyltrimethylammonium chloride (DTAC, 99%) and sodium perchlorate (NaClO_4 , 98%) were purchased from Sigma-Aldrich Co. Sodium decylsulfate (SDeS, 99%), sodium octylsulfate (SOS, 99%) and sodium pentylsulfate (SPS, 99%) were purchased from Alfa Aesar. Sodium ethylsulfate (SES, 98%) was purchased from Tokyo Chemical Industry Co., Ltd. (TCI). Dodecylphosphocholine (DPC, 99%) and dodecyl- β -D-maltoside (DDM, 99%) were purchased from Avanti polar lipids, Inc. All chemicals were used as received unless specified otherwise. Water (18.2 $\text{M}\Omega$ cm) was purified by a NANOPURE Infinity™ ultrapure water purifier. All samples contained 0.1 M NaClO_4 unless otherwise noted.

7.2.2 Circular dichroism measurements

CD spectra were measured by using a Jasco J-710 spectrometer. We used a temperature-controlled quartz cuvette with a 0.2 mm path length for 1.0 mg/ml (1.0 mM) samples. The temperature was controlled at 20 °C by an external water bath. The CD spectra were collected at 0.2 nm intervals and averaged over 5 scans and smoothed by the second order Savitzky-Golay method over 15 data points.

7.2.3 204 nm excited UVRR spectra

The UVRR instrumentation was described in detail previously.⁴⁸ Briefly, the third harmonic of a Nd:YAG laser (Coherent Infinity) was anti-Stokes Raman shifted five harmonics in 40 psi hydrogen gas to 204 nm. The 204 nm light is in resonance with the peptide bond first allowed π - π^* electronic transition. The excitation beam was directed to a spinning quartz NMR tube containing the room temperature (~20 °C) sample solutions. The backscattered Raman light was collected and dispersed by a partially subtractive double monochromator and detected by a Lumogen coated back-thinned CCD detector (Princeton Instruments Spec-10:400B) that is cooled by liquid nitrogen. The sample UVRR spectra were obtained by subtracting appropriate amounts of water and the empty quartz NMR tube UVRR from the raw spectra and normalized to the intensity of the 0.1 M NaClO₄ internal standard.

7.2.4 NMR diffusion measurements

All diffusion coefficient measurements were performed using a Bruker Avance III 600 MHz spectrometer with a BBFO Plus probe. The temperatures were controlled to ± 1 °C accuracy using a Bruker BVT3000 temperature control system. The spectrometer was calibrated against the water self-diffusion coefficient at 25 °C. The samples were prepared in D₂O and the residual HOD resonance at $\delta = 4.70$ was used as an internal chemical shift standard. Diffusion coefficients were determined using a stimulated echo pulsed field gradient pulse sequence with bipolar gradients.⁴⁹⁻⁵⁰ To obtain the diffusion coefficient, the peak intensity vs. gradient strength data were fit to the Stejskal-Tanner equation⁵¹⁻⁵²

$$I = I_0 e^{-4\pi^2 D_s \gamma^2 G^2 \xi^2 \left(\Xi - \frac{\xi}{3}\right)} \quad (1)$$

where I and I_0 are the intensities of a specific band (or the intensity at a specific chemical shift) with and without the magnetic field gradient; γ is the gyromagnetic ratio and equals 42.58 MHz · T⁻¹ for ¹H nuclei; ξ and Ξ are the pulse width and pulse interval, D_s is the diffusion coefficient of the species in D₂O; and G is the gradient strength applied to the sample. 16 spectra were taken with gradient strengths that varied between 0 and 50 G/cm, while the duration of the gradient ξ was held constant throughout the experiment.

To plot the results as two dimensional diffusion-ordered spectroscopy (DOSY) contours with respect to the chemical shift and diffusion coefficient, these 16 spectra were fitted to eq. (1) to obtain $D_s(\delta)$ and the intensity $I_0(\delta)$ at each chemical shift δ . A normal distribution of intensities was generated along the diffusion coefficient axis at each chemical shift. The peak height, center position and standard deviation of the normal distribution are labeled as $I_0(\delta)$, $D_s(\delta)$ and $\Delta D_s(\delta)$ (standard error of $D_s(\delta)$), respectively⁵³

$$I(\delta, D) = I_0(\delta) e^{-\frac{(D-D_s(\delta))^2}{2\Delta D_s(\delta)^2}} \quad (2)$$

The integrated intensities of resonances from the same species were globally fit to obtain a single diffusion coefficient for qualitative analysis by using eq. (1).

7.3 RESULTS AND DISCUSSIONS

7.3.1 CD spectra of XAO in different SDS concentration solutions

Figure 7.1A shows the CD spectra of XAO as a function of SDS concentration in a 0.1 M NaClO₄ aqueous solution at 20 °C. In the absence of SDS, the CD spectrum of XAO is similar to previously reported CD spectra of XAO in pure water or in low salt concentration solutions.^{33,54-}

⁵⁵ The CD spectra show characteristic PPII features such as the intense negative band at ~197 nm. These features indicate that XAO adopts mainly a PPII-like conformation in 0.1 M NaClO₄ aqueous solutions in the absence of SDS. As the SDS concentration increases, the CD spectrum begins to show α -helix features, with two negative bands at 222 nm and 206 nm and a strong positive band at ~190 nm. A well defined isodichroic point occurs at 202 nm, indicating that XAO apparently undergoes a two-state transition from a PPII-like conformation to an α -helix-like conformation.

The negative mean residue ellipticity is proportional to the fraction of α -helix-like conformations

$$[\theta] = f_H([\theta]_H - [\theta]_P) + [\theta]_P \quad (3)$$

To calculate the α -helical fraction f_H , the 222 nm mean residue ellipticities of the α -helical state $[\theta]_H$ and the PPII state $[\theta]_P$ need to be determined. $[\theta]_P$ is the mean residue ellipticity of the XAO PPII-like conformation. This value can be obtained from the CD measurement of XAO in the absence of SDS.

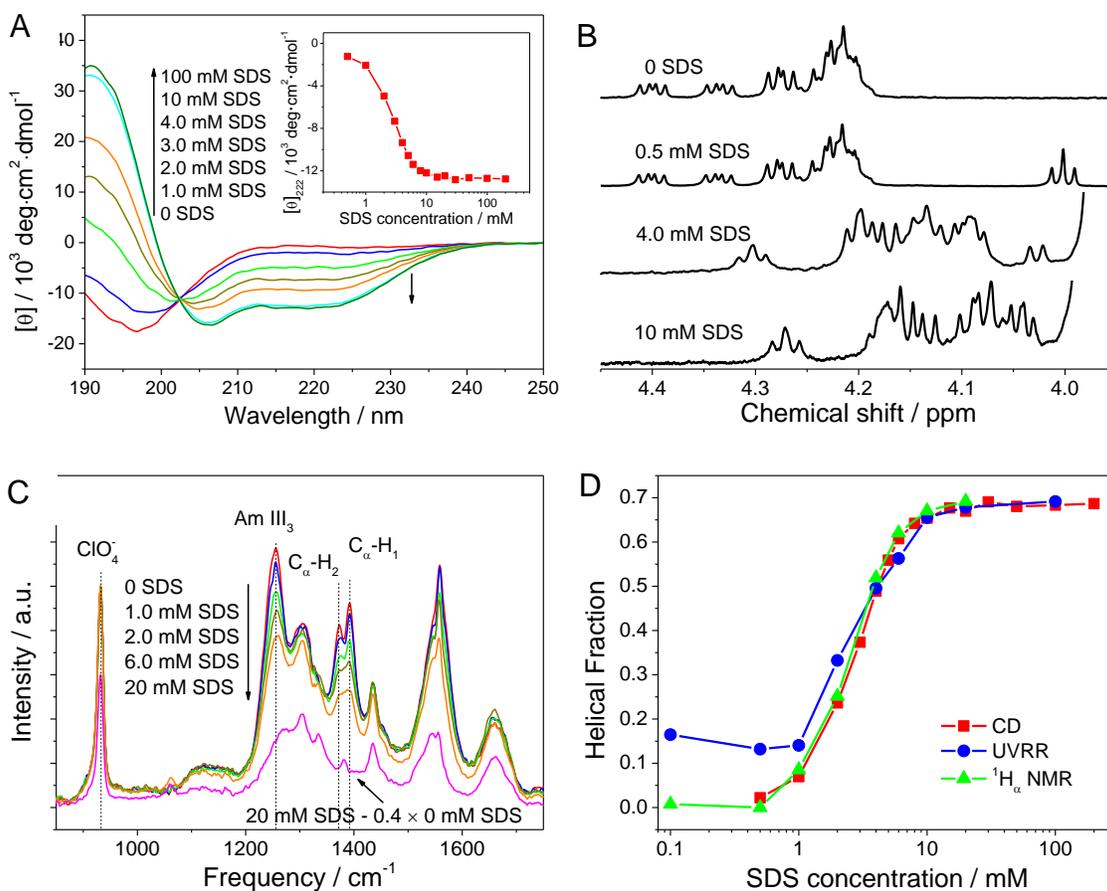


Figure 7. 1. (A) CD spectra of 1.0 mM XAO in 0.1 M aqueous NaClO₄ solution at different SDS concentrations.

The inset shows the dependence on the SDS concentration of the CD spectra 222 nm ellipticity. (B) ¹H_α NMR spectra of XAO in 0.1 M NaClO₄ at different SDS concentrations. (C) 204 nm excited UVRR spectra of XAO in 0.1 M NaClO₄ at different SDS concentrations. The difference spectrum calculated emphasizes the Ala methyl group umbrella bending band at 1382 cm⁻¹. (D) SDS concentration dependence of the α -helical fraction of 1.0 mM XAO in 0.1 M NaClO₄ calculated from the CD, ¹H_α NMR and UVRR spectra.

The determination of $[\theta]_{\text{H}}$ is challenging, because it is difficult to force XAO to adopt a pure α -helix state. The commonly used strong α -helix promoting solvent trifluoroethanol (TFE) does not induce substantial amounts of α -helix. In a previous study,⁵⁵ the 222 nm mean residue ellipticity of XAO in TFE was measured to be $-7000 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$, which is less negative than XAO in 10 mM SDS. The CD spectrum of XAO in TFE does not resemble the pure α -helix CD spectra, indicating that XAO does not exist in its pure α -helix conformation in TFE.

Another common approach⁵⁶⁻⁵⁷ to estimate $[\theta]_{\text{H}}$ uses eq. (4)

$$[\theta]_{\text{H}} = [\theta]_{\infty} (n - k)/n \quad (4)$$

where, $[\theta]_{\infty} = -40000 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ is the mean residue ellipticity at 222 nm of an infinitely long α -helix segment; n is the number of residues ($n = 11$ for XAO); k is a parameter that accounts for the end effects, since the end residues cannot effectively form α -helix conformations. k can vary from 4.6 to 6.3.⁵⁷ For the 21-residue polyAla peptide AP, k was reported to be 7.6.⁵⁸ The large variance of k results in a large uncertainty of $[\theta]_{\text{H}}$ estimated for short peptides.

The inset of Figure 7.1A shows the dependence of the XAO 222 nm ellipticity on the SDS concentration in 0.1 M NaClO₄ at 20 °C. The sigmoid shaped titration curve clearly indicates that SDS concentrations <1 mM have little impact on the XAO conformation. XAO is dominated by PPII-like conformations and shows a 222 nm mean residue ellipticity $[\theta]_{\text{P}} = -850 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$. The PPII to α -helix transition midpoint occurs at ~3 mM SDS. Above 10 mM SDS, the α -helix conformation “saturates” with a 222 nm mean residue ellipticity of $-12800 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$. We assigned this value to a helical fraction of 0.69 that we calculated from the UVRR spectra (see below), and calculate $[\theta]_{\text{H}} = -18200 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$. This value is much more negative than that found for XAO in TFE,⁵⁵ indicating that, as

expected, the maximum helical state is not achieved in TFE. The k parameter estimated from eq. (4) is 6.0.

7.3.2 XAO $^1\text{H}_\alpha$ NMR spectra

Figure 7.1B shows the one-dimensional $^1\text{H}_\alpha$ NMR spectra of XAO at different SDS concentrations in 0.1 M NaClO_4 aqueous solutions. The $^1\text{H}_\alpha$ NMR peaks of XAO overlap. In the absence of SDS, the $^1\text{H}_\alpha$ chemical shifts range from 4.18 – 4.12. The addition of 0.5 mM SDS does not affect the resonances of XAO (The triplet at ~ 4.0 ppm arises from SDS). As the SDS concentration increases, the peaks shift further upfield with a significantly changed pattern.

The $^1\text{H}_\alpha$ XAO chemical shifts are sensitive to secondary structure.⁵⁹⁻⁶¹ For the "random coil" conformation, the reference chemical shift is 4.26 for the $^1\text{H}_\alpha$ of both Ala and Lys.⁶¹ Presumably the chemical shifts are somewhat larger for Dab and Orn than Lys due to the fewer methylene groups between the amine group and C_α in Dab and Orn. The $^1\text{H}_\alpha$ chemical shift for the PPII conformation differs little from "random coil".⁶² In α -helices, the $^1\text{H}_\alpha$ chemical shifts of Ala and Lys decrease to 4.03 and 3.99 respectively.⁶¹ Our observed upfield shifts of XAO $^1\text{H}_\alpha$ resonances upon addition of SDS indicate a transition of PPII to an α -helix-like conformation, supporting the conclusions obtained from the CD spectra.

7.3.3 UVRR spectra of XAO dependence on SDS concentration

The UVRR spectra of XAO with and without 0.1 M NaClO_4 are essentially identical (Figure 7.9), indicating that 0.1 M NaClO_4 has negligible impact on the XAO conformation. We

measured the 204 nm excited UVRR spectra of 1.0 mM XAO at different SDS concentrations with 0.1 M NaClO₄ at room temperature.

Figure 7.1C shows the 204 nm excited XAO UVRR spectra. Without SDS, the XAO UVRR spectrum shows an Am I band at ~1660 cm⁻¹, an Am II band at ~1550 cm⁻¹, overlapped with an interfering molecular oxygen band at 1556 cm⁻¹, two C_α-H bending bands at ~1392 cm⁻¹ and ~1373 cm⁻¹, and an Am III₃ band at ~1255 cm⁻¹. As the SDS concentration increases, the Am III₃ band frequency slightly upshifts, while the intensities of the Am III₃ band and the C_α-H bending bands significantly decrease, indicating that XAO undergoes a transition from a PPII-like conformation to α-helix-like conformations that show hypochromism. These XAO spectra can be modeled as the sum of XAO spectra without SDS and XAO spectra with 100 mM SDS, confirming that this transition appears spectroscopically as a two-state system. The UVRR results are fully consistent with the CD and NMR results.

The C_α-H bending vibrations are resonance enhanced due to their coupling with N-H bending in the PPII-like conformation.⁶³ In α-helix-like conformations, the C_α-H bending decouples from the N-H bending and disappears.⁶⁴ Therefore, the C_α-H bending band intensities are mainly contributed by the PPII-like conformation. Consequently, the C_α-H bending bands can be used to calculate the PPII and α-helical fractions based on two-state modeling.⁵⁸

7.3.4 α-Helical fractions calculated from UVRR, CD and ¹H_α NMR

The difference spectrum, shown in Figure 7.1C, between the UVRR spectra of XAO at 20 mM SDS and 0 mM SDS shows a band at 1382 cm⁻¹, that can be assigned to the methyl group

umbrella bending of the Ala side chains.⁶⁵⁻⁶⁷ We expect that the relative intensity of this methyl group umbrella bending band at 1382 cm⁻¹ is independent of the XAO conformation.

We fit the XAO C_α-H bending region (from 1350 cm⁻¹ to 1420 cm⁻¹) of the UVRR spectra at different SDS concentrations with three Lorentzians. Two Lorentzian bands at 1371 cm⁻¹ and 1394 cm⁻¹ model the C_α-H bending band spectra of the 11 residues. A Lorentzian band at 1382 cm⁻¹ models the Ala methyl group umbrella bending. Assuming there are only two states, the α -helix-like conformation and the PPII conformation, the PPII fraction can be calculated from the 1371 cm⁻¹ and 1394 cm⁻¹ Lorentzian band intensities at different SDS concentrations. The α -helical fractions of 1.0 mM XAO at different SDS concentrations calculated from UVRR spectra are shown in Figure 7.1D.

When the SDS concentration is above 10 mM, XAO shows a "saturated" helical fraction of ~0.69 calculated from the UVRR. The "saturated" α -helical fraction of ~0.69 indicates that, on average, around 7.6 of the 11 residues in XAO peptide are α -helical. The 7.6 residue long helical segment most likely occurs as a single segment in the middle of the peptide. We used the "saturated" helical fraction of ~0.69 to estimate $[\theta]_H$ via eq. (3). With this $[\theta]_H$ value, we can calculate the XAO helical fraction at any SDS concentration from the CD spectra. The calculated XAO helical fractions are shown in Figure 7.1D as well.

The chemical shifts of the XAO ¹H_α, as discussed above, depend upon the XAO conformation. We calculated the first moment, which is the intensity weighed average chemical shift of the XAO ¹H_α NMR bands at various SDS concentrations. The first moment decreases from 4.26 ppm in the absence of SDS to 4.12 ppm in 20 mM SDS. As shown by the CD and UVRR spectra, the XAO α -helix-PPII transition appears spectroscopically to be two-state. If the first moment of the PPII conformation is at 4.26 ppm, and the first moment of the SDS

"saturated" XAO conformation (69 % α -helix-like and 31 % PPII) is at 4.12 ppm, assuming these two conformations are in the fast exchange limit, we can calculate the α -helical fractions of XAO in SDS aqueous solutions from the first moments of the XAO $^1\text{H}_\alpha$ NMR bands using

$$f_{\text{H}} = \frac{\delta - \delta_{\text{P}}}{\delta_{\text{H}} - \delta_{\text{P}}} \quad (5)$$

where δ is the first moment of the XAO $^1\text{H}_\alpha$ NMR band at a given SDS concentration. $\delta_{\text{P}} = 4.26$ ppm is the XAO PPII conformation $^1\text{H}_\alpha$ NMR chemical shift and $\delta_{\text{H}} = \frac{4.12 - 4.26 \times 0.31}{0.69}$ ppm = 4.06 ppm is the XAO α -helix-like conformation $^1\text{H}_\alpha$ NMR chemical shift.

The calculated α -helical fractions from the first moments of the XAO $^1\text{H}_\alpha$ NMR spectra (eq. (5)) are shown in Figure 7.1D. These α -helical fractions are similar to those calculated from the CD and UVRR spectra.

7.3.5 The effects of long alkyl chains and negatively charged surfactant head groups

Figure 7.2A shows the dependence of the 222 nm mean residue ellipticities of XAO on the concentrations of surfactants that have the same anionic sulfate head group as SDS but different alkyl chain lengths. SES and SPS that possess relatively short alkyl chains have negligible impact on the XAO conformation even up to 0.7 – 1.0 M concentrations. Figure 7.2A shows that the surfactant concentration required to induce the XAO PPII to α -helix transition decreases as the surfactant alkyl chain length increases. The critical micelle concentrations (*cmcs*) of these surfactants are similarly dependent on chain length. The PPII to α -helix transition of XAO induced by SOS occurs between ~50 mM to ~120 mM with a midpoint at ~80 mM, while the

transition induced by SDeS occurs between ~10 mM to ~40 mM with a midpoint at ~17 mM.

The transition induced by SDS occurs between ~1 mM to ~10 mM with a midpoint at ~3 mM.

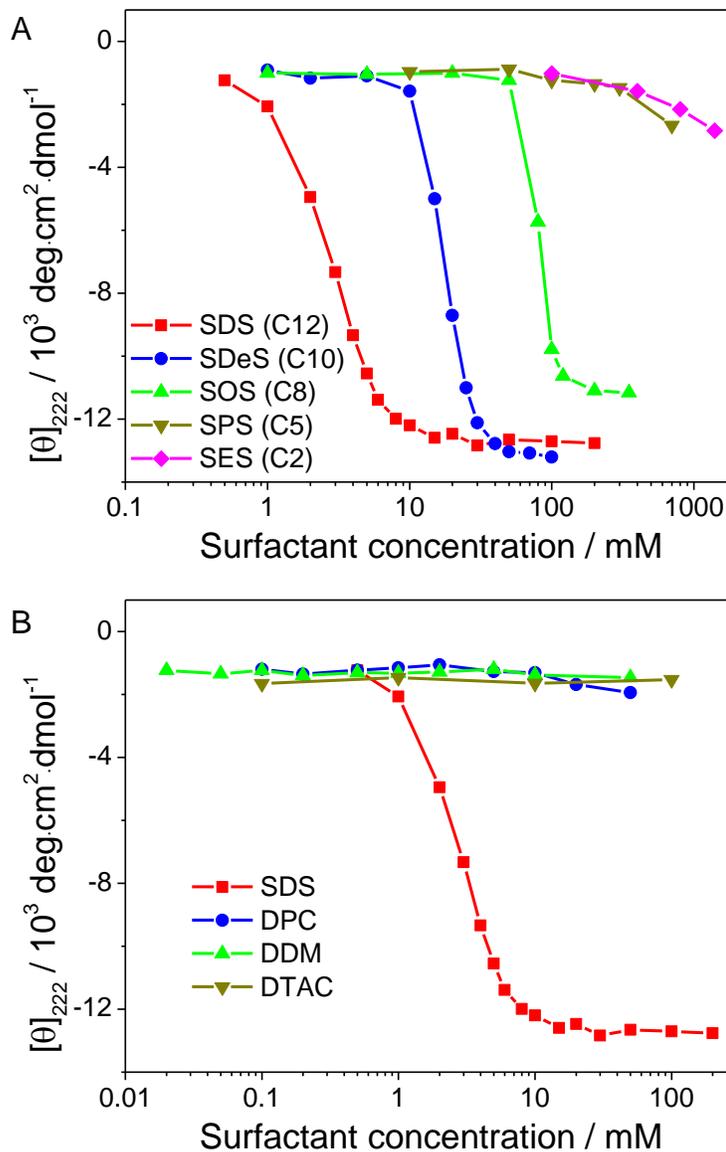
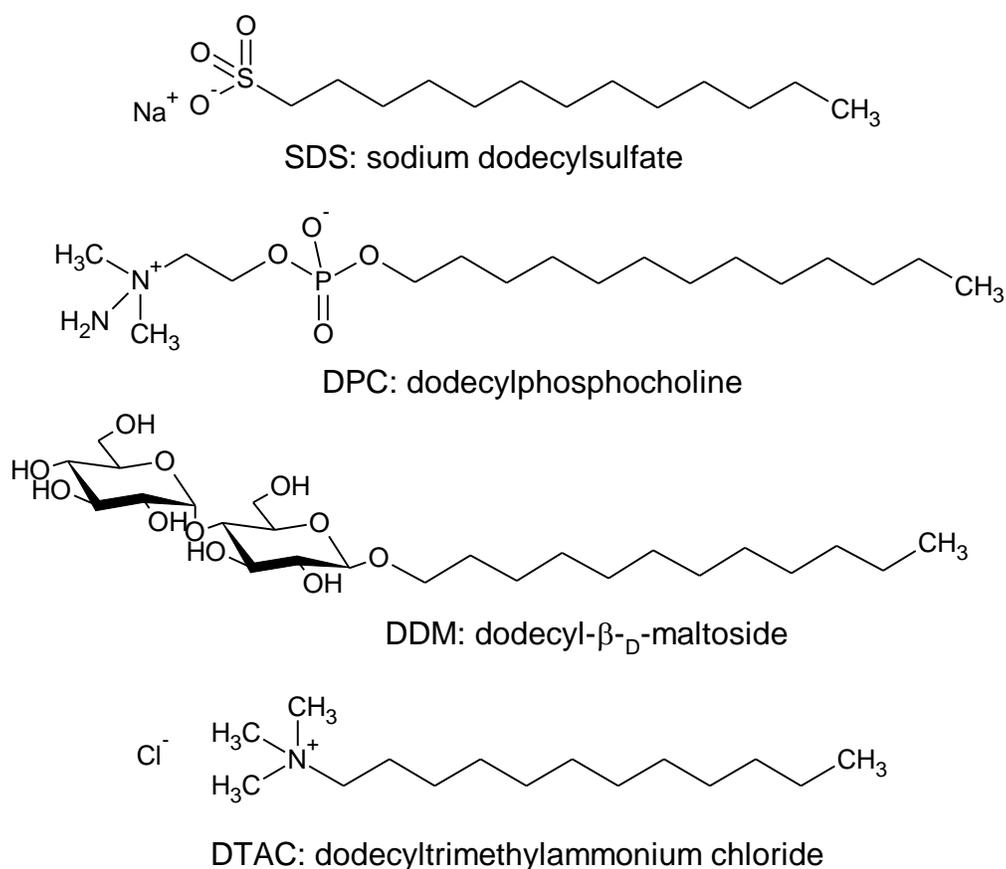


Figure 7. 2. The 222 nm ellipticities of 1.0 mM XAO in the presence of different surfactants.

(A) Surfactants with different alkyl chain lengths. SDS: sodium dodecylsulfate; SDeS: sodium decylsulfate; SOS: sodium octylsulfate; SPS: sodium pentylsulfate; SES: sodium ethylsulfate. (B) Surfactants with different head groups. SDS: sodium dodecylsulfate; DPC: dodecylphosphocholine; DDM: dodecyl- β -D-maltoside; DTAC: dodecyltrimethylammonium chloride. Scheme 7. 1 shows the molecular structures of these surfactants.

Figure 7.2B compares the dependence of the XAO 222 nm mean residue ellipticities on the concentrations of surfactants with identical alkyl chain lengths but different head groups. These surfactants include anionic SDS, zwitterionic DPC, non-ionic DDM and cationic DTAC (Scheme 7. 1). Only SDS induces XAO to form α -helices, while the other three surfactants have little impact on the XAO conformation, even at concentrations much higher than their *cmc*s. This suggests that the anionic SDS interacts with cationic XAO to form α -helix conformations through electrostatic interactions with the four positively charged XAO side chains.



Scheme 7. 1. Molecular structures of different charged head group surfactants

7.3.6 Diffusion coefficients measured by ¹H-NMR DOSY experiment

We used NMR DOSY to study the size and stoichiometry of the XAO-SDS aggregates by measuring the dependencies of the XAO and SDS diffusion coefficients on the SDS concentration.⁶⁸ The NMR DOSY spectra of 1.0 mM XAO at 0, 0.5 mM, 4.0 mM and 10 mM SDS concentrations are shown in Figure 7.3. The intensity of the resonance in the absence of a magnetic field gradient (I_0 , that was calculated via eq. (1)) was plotted as a function of chemical shift at the top of contour plots. Along the right is a pseudo chromatograph, indicating the calculated diffusion coefficient associated with each chemical shift. The contour at $\delta = \sim 4.7$ ppm with $D = \sim 2 \times 10^{-5} \text{ cm}^2 \cdot \text{s}^{-1}$ derives from HOD in D₂O.

In the absence of SDS, Figure 7.3A shows XAO proton resonances at $D = 2.0 \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$. Upon addition of SDS, the NMR spectrum shows SDS proton resonances. At 0.5 mM SDS, monomer SDS peak appears in the diffusion chromatograph at $D = 4.3 \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$ (Figure 7.3B). As the SDS concentration increases (Figure 7.3C, 3D), the diffusion coefficients of both the SDS and XAO species decrease, The XAO ¹H_α resonances shift towards lowfield, indicating α-helix formation as discussed above.

Figure 7.4A shows the dependence of the diffusion coefficients of 1.0 mM XAO and SDS in 0.1 M NaClO₄ on the SDS concentration. Also shown are the SDS diffusion coefficients in 0.1 M NaClO₄ (without XAO) as a function of SDS concentration. The SDS diffusion coefficient in the absence of XAO begins to decrease at ~ 1.8 mM, indicating formation of micelles at a value similar to the *cmc* of SDS in 0.1 M NaCl (1.62 mM).⁶⁹

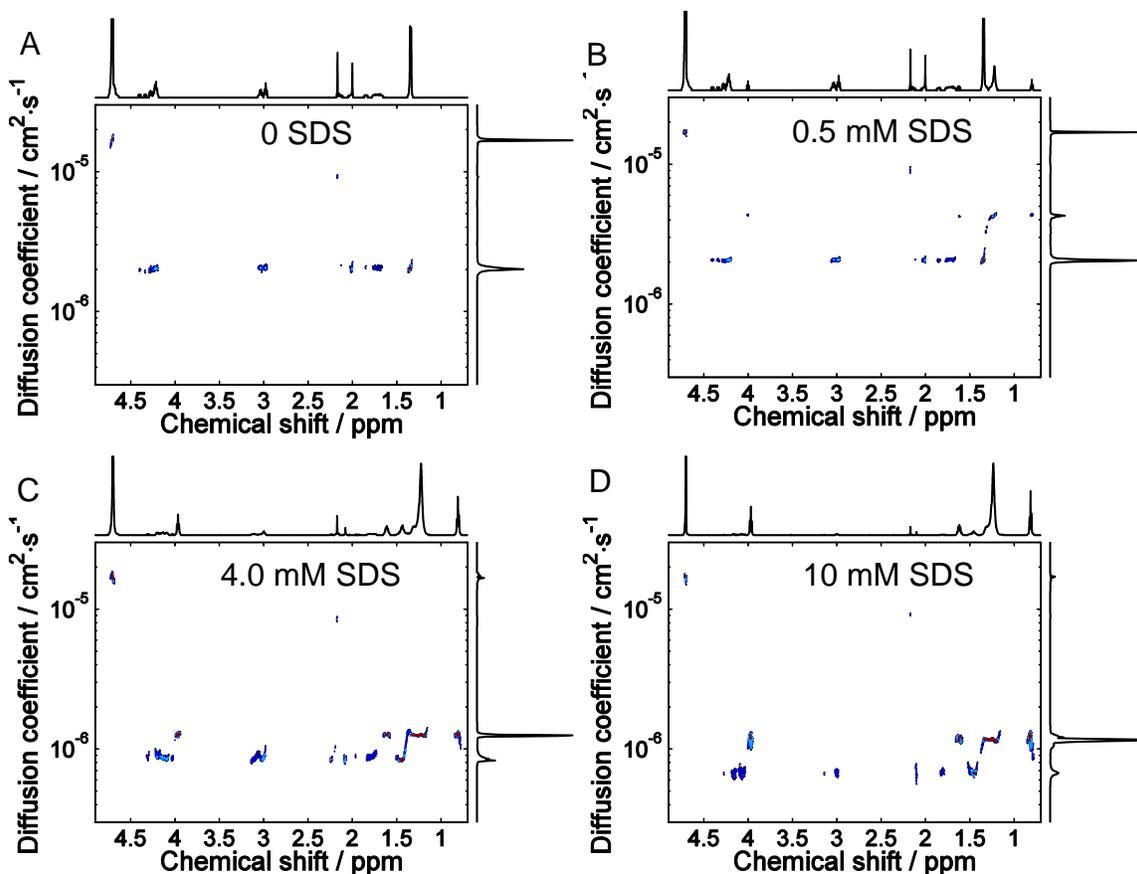


Figure 7. 3. ^1H NMR DOSY spectra of XAO

in (A) 0 mM, (B) 1.0 mM, (C) 4.0 mM and (D) 10 mM SDS solutions. All solutions contain 0.1 M NaClO_4 . The calculated one-dimensional NMR spectrum without a magnetic field gradient is shown at the top of each figure.

Along the right are pseudo chromatographs indicating the concentration of species as a function of diffusion coefficient.

Solutions containing 0.1 M NaClO_4 , 1.0 mM XAO and < 0.5 mM SDS, show constant XAO and SDS diffusion coefficients. The SDS diffusion coefficient is identical to that of 0.5 mM SDS in 0.1 M NaClO_4 . These results indicate that XAO and SDS remain monomeric with $D_{\text{XAO}} = (2.02 \pm 0.01) \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$ and $D_{\text{SDS}} = (4.30 \pm 0.03) \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$, respectively (Table 1).

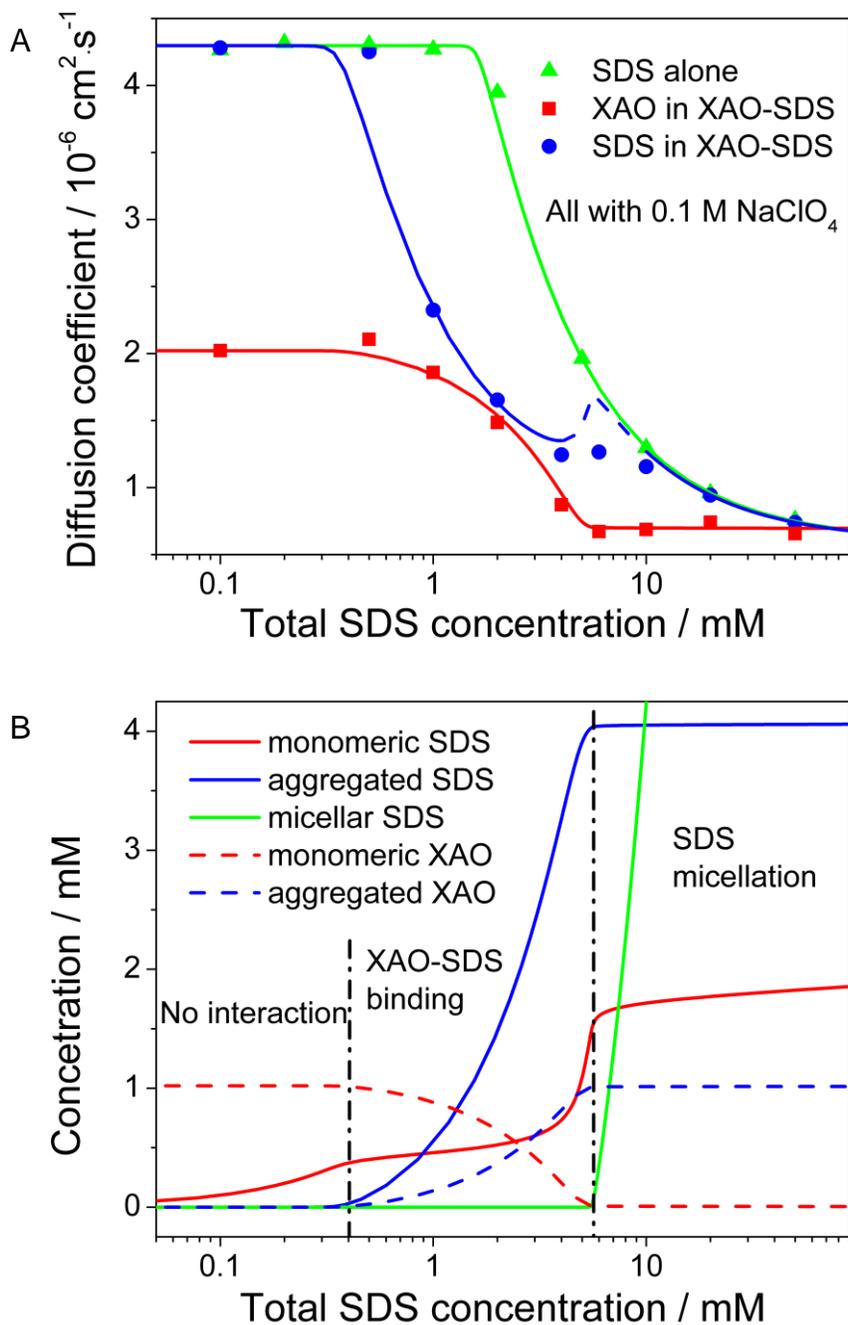


Figure 7. 4. (A) Diffusion coefficients of 1.0 mM XAO and SDS in XAO-SDS solution and the diffusion coefficients of pure SDS in aqueous solution. All solutions contain 0.1 M NaClO_4 . The data were experimentally measured by Pulsed Field Gradient NMR; the curves are calculated by using parameters obtained from experimental data. (B) Calculated concentrations of different species in XAO-SDS solutions by using parameters obtained from experimental data.

Above 0.5 mM SDS, the XAO and SDS diffusion coefficients decrease, indicating the formation of XAO-SDS aggregates. This 0.5 mM SDS concentration is significantly lower than the *cmc* of SDS in 0.1 M NaClO₄ without XAO, indicating that the XAO-SDS aggregates are formed between XAO monomers and SDS monomers. To confirm that the XAO-SDS aggregate formation does not require SDS micelles, we measured the dependence of the XAO CD spectra on SDS concentration in the absence of NaClO₄. Removal of NaClO₄ increases the SDS *cmc* to ~ 8 mM.⁶⁹ If the XAO-SDS aggregation requires SDS micelle, it should show a different XAO conformational dependence on SDS concentration than in the presence of NaClO₄. In contrast, we observe a very similar SDS concentration dependence (Figure 7.10). This similar dependence confirms that XAO-SDS aggregation occurs between XAO monomers and SDS monomers.

Table 7. 1. Diffusion coefficients and hydrodynamic radius of species in XAO-SDS solution

species	diffusion coefficient ^a / 10 ⁻⁶ cm ² ·s ⁻¹	Hydrodynamic radius ^b / nm
XAO monomer	2.02 ± 0.01	0.98 ± 0.01
XAO-SDS aggregate	0.69 ± 0.02	2.9 ± 0.1
SDS monomer	4.30 ± 0.03	0.46 ± 0.01
SDS micelle	0.60 ± 0.03 ^c	3.3 ± 0.1

^ameasured at 25 °C in D₂O in the presence of 0.1 M NaClO₄ by PFG-NMR.

^bcalculated from diffusion coefficient using Stokes-Einstein equation $D = \frac{k_B T}{6\pi\eta r}$, where $\eta = 1.098$ mPa·s is the D₂O viscosity⁷⁰ at 25 °C.

^cobtained by extrapolating the SDS diffusion coefficient to infinite SDS concentration in 0.1 M NaClO₄ without XAO.

At SDS concentration greater than ~5 mM, the XAO diffusion coefficient remains constant at $(6.9 \pm 0.2) \times 10^{-7}$ cm² · s⁻¹. This constant XAO diffusion coefficient indicates that the XAO-SDS aggregates do not grow with increasing SDS concentration. The SDS diffusion

coefficient decreases above ~10 mM SDS due to the formation of SDS micelles ($D_{\text{Mic}} = (6.0 \pm 0.3) \times 10^{-7} \text{ cm}^2 \cdot \text{s}^{-1}$).

Thus, aggregation between XAO and SDS monomers occurs at SDS concentrations between ~0.5 mM to ~5 mM SDS. Above 10 mM SDS, micelle formation occurs among SDS itself as evidenced by the decreased SDS diffusion coefficient. The diffusion coefficients of the XAO monomer, the SDS monomer, the XAO-SDS aggregate and the SDS micelle are listed in Table 1 along with the hydrodynamic radii estimated using the Stokes-Einstein equation that assumes spherical species.

7.3.7 Modeling of XAO and SDS diffusion coefficients

For a system in fast exchange, the apparent diffusion coefficient is a weighted average of the diffusion coefficients of all fast exchanging species

$$D^{\text{App}} = \frac{\sum_{i=1}^n q_i C_i D_i}{C^{\text{Tot}}} \quad (6)$$

where q_i , C_i and D_i are the stoichiometry, the concentration and diffusion coefficient of the i th species; $C^{\text{Tot}} = \sum_{i=1}^n q_i C_i$ is the total concentration. From the diffusion coefficients of the XAO monomer, the SDS monomer, the XAO-SDS aggregate and the SDS micelle listed in Table 1, we can use eq. (6) to calculate the concentrations of XAO monomer (C_{XAO}) and SDS monomer (C_{SDS}), XAO-SDS aggregates (C_{Agg}) and SDS micelles (C_{Mic}) at each SDS concentration.

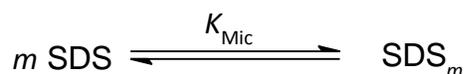
The binding equilibrium between XAO and SDS monomers, assuming only one aggregate species, can be described as



where K_{Agg} is the equilibrium constant for the XAO-SDS aggregate, a and b are the numbers of XAO and SDS molecules in the XAO-SDS aggregate.

Fitting this model with species concentrations calculated using eq. (6) at $\text{SDS} < 5 \text{ mM}$ finds that the ratio $b/a = 3.9 \pm 0.3$, $a = 4.3 \pm 0.1$ and $K_{\text{Agg}} = (3.8 \pm 0.1) \times 10^4 \text{ mM}^{-20.5}$. The ratio b/a suggests that in the aggregate, each XAO molecule is bound to four SDS monomers. The binding sites are most likely the four positively charged XAO side chains. Each XAO-SDS aggregate contains ~ 4 XAO molecules and ~ 16 SDS molecules.

As the SDS concentration increases, the XAO monomer becomes depleted. In contrast, the SDS monomer concentration increases until it reaches the *cmc* and micelles form. If m SDS monomers form a micelle



where K_{Mic} is the micelle formation constant. Fitting this model with species concentrations calculated using eq. (6) for SDS concentrations from 6 mM to 50 mM finds an aggregation number, $m = 38 \pm 8$ and $K_{\text{Mic}} = (1.1 \pm 5.5) \times 10^{-10} \text{ mM}^{-37}$.

Simple statistical mechanics predicts $\text{cmc} = K_{\text{Mic}}^{-\frac{1}{m-1}}$.⁷¹ Thus, we estimate $\text{cmc} = 1.8 \pm 0.1 \text{ mM}$, identical to the *cmc* of SDS measured in 0.1 M NaClO_4 solution (Figure 7.4A). This is close to the reported 1.62 mM *cmc* of SDS in 0.1 M NaCl .⁶⁹ The presence of 1.0 mM XAO does not appear to affect the SDS *cmc* equilibrium. However, XAO binds to SDS monomers and lowers the effective SDS monomer concentration, thus increasing the apparent total SDS concentration required to form SDS micelles.

7.3.8 Mechanism of SDS-induced XAO α -helix formation

The diffusion coefficients (Figure 7.4A) and the solution species concentrations (Figure 7.4B) as a function of total SDS concentration were modeled using the parameters obtained above. The SDS concentration dependencies of solution species shown in Figure 7.4B can be divided into three regions: no interactions (< 0.5 mM total SDS), XAO-SDS binding ($0.5 - 5$ mM total SDS) and SDS micellation (> 5 mM total SDS).

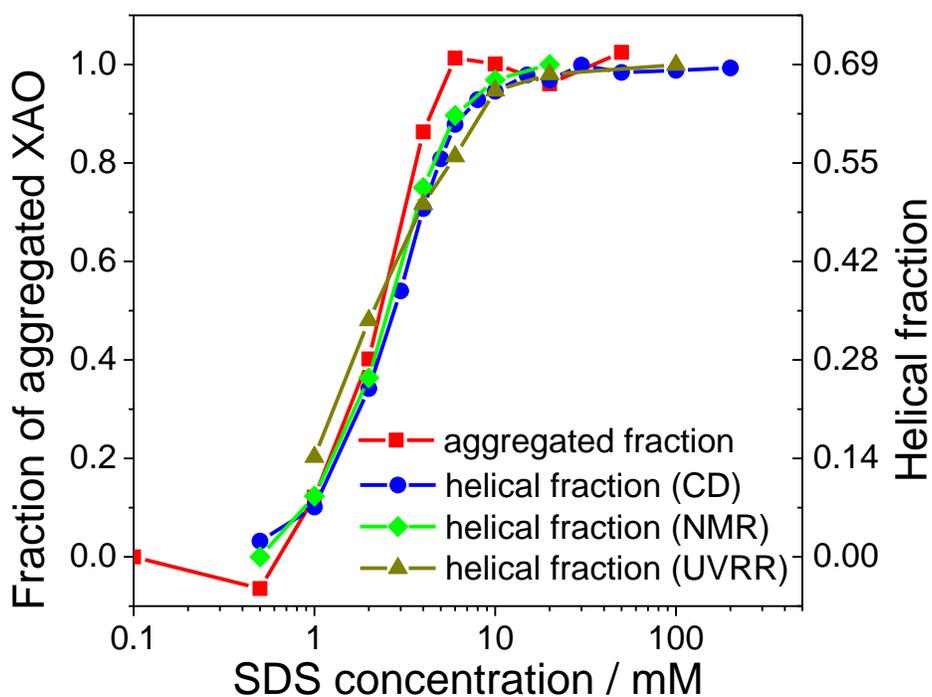


Figure 7. 5. Comparison between fractions of aggregated XAO and α -helical XAO as determined by CD, NMR and UVRR.

Below 0.5 mM SDS, SDS does not bind to XAO or form micelles. The XAO monomer concentration remains constant as the SDS monomer concentration increases. Between 0.5 to 5 mM total SDS concentrations, SDS binds to XAO forming XAO-SDS aggregates prior to SDS micellation. The concentrations of aggregated XAO and aggregated SDS increase at the expense

of the XAO monomer. The SDS monomer concentration also increases, however, it remains below the *cmc* of SDS until ~ 5 mM total SDS concentration. Above ~ 5 mM total SDS, the micellation is triggered and all species remain at constant concentrations, except the SDS micelle concentration increases.

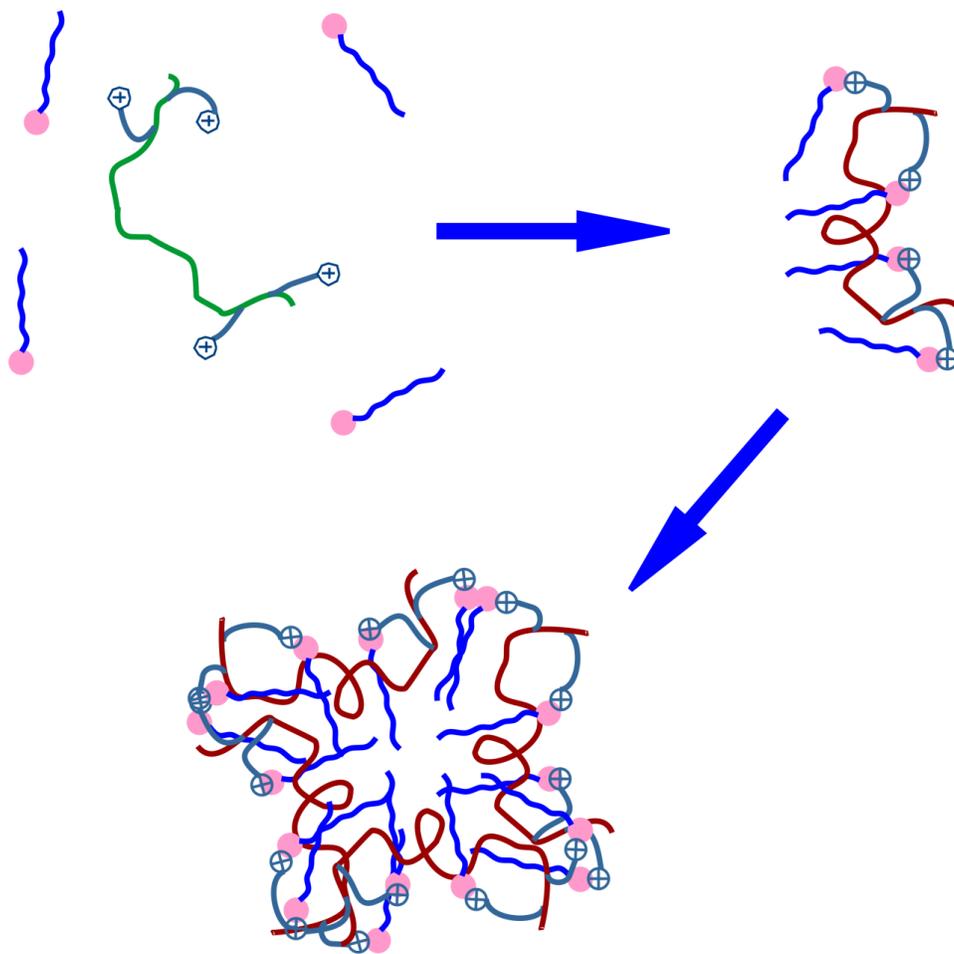


Figure 7. 6. Mechanism of the XAO₄-SDS₁₆ aggregation and the formation of XAO α -helix-like conformations. The four SDS alkyl chains bound to each XAO partition into a hydrophobic core that minimizes the hydrophobic area exposed to water. Neutralization of the flanking XAO charges enables α -helix formation such that one α -helix peptide face is buried within the hydrophobic core while the other face is in contact with the water environment.

Figure 7.5 compares the fraction of the aggregated XAO calculated from the diffusion coefficients, to the α -helical fractions calculated from the CD, $^1\text{H}_\alpha$ -NMR and UVRR spectra.

Upon scaling by 0.69, the fraction of the aggregated XAO exactly tracks the α -helical fraction, indicating that the XAO α -helix-like conformation correlates directly with XAO-SDS aggregation. However, in the XAO-SDS aggregates, some of the XAO residues, most likely the flanking Orn and Dab, remain in a PPII conformation.

Figure 7.6 shows a schematic model for the interactions between XAO and SDS that lead to α -helix formation. The negatively charged head groups of the SDS monomers electrostatically bind to the four XAO positively charged side chains. This significantly reduces electrostatic repulsion between Dab and Orn side chains and allows XAO to adopt the more compact α -helix-like conformations from the extended PPII conformation. The XAO-SDS₄ aggregates further associate with each other to form XAO₄-SDS₁₆ aggregates. This results in a hydrophobic inner core comprised of the SDS alkyl groups, while one face of the α -helix backbone is probably exposed to water medium. The other face of the XAO α -helix backbone is buried within the hydrophobic inner core. This stabilizes the α -helix formed by protecting the backbone hydrogen bonds.

7.3.9 Ramachandran ψ angle distributions of XAO in SDS solution

The UVRR spectra contain information on the aggregated XAO conformations. By subtracting appropriate amounts of the measured PPII-like conformation UVRR spectrum, we can calculate the pure α -helix-like spectrum associated with the aggregated XAO at each SDS concentration. The pure α -helix-like spectra at different SDS concentrations are essentially identical (Figure 7.11), although there is somewhat less inhomogeneous broadening at higher SDS concentrations.

We thus averaged over these calculated pure α -helix-like spectra at different SDS concentrations ($S_{\alpha}(v, C_{\text{SDS}})$) weighted by their respective α -helical fractions ($f_{\alpha}(C_{\text{SDS}})$):

$$S_{\alpha}(v) = \frac{\sum_{C_{\text{SDS}}} f_{\alpha}(C_{\text{SDS}}) S_{\alpha}(v, C_{\text{SDS}})}{\sum_{C_{\text{SDS}}} f_{\alpha}(C_{\text{SDS}})} \quad (7)$$

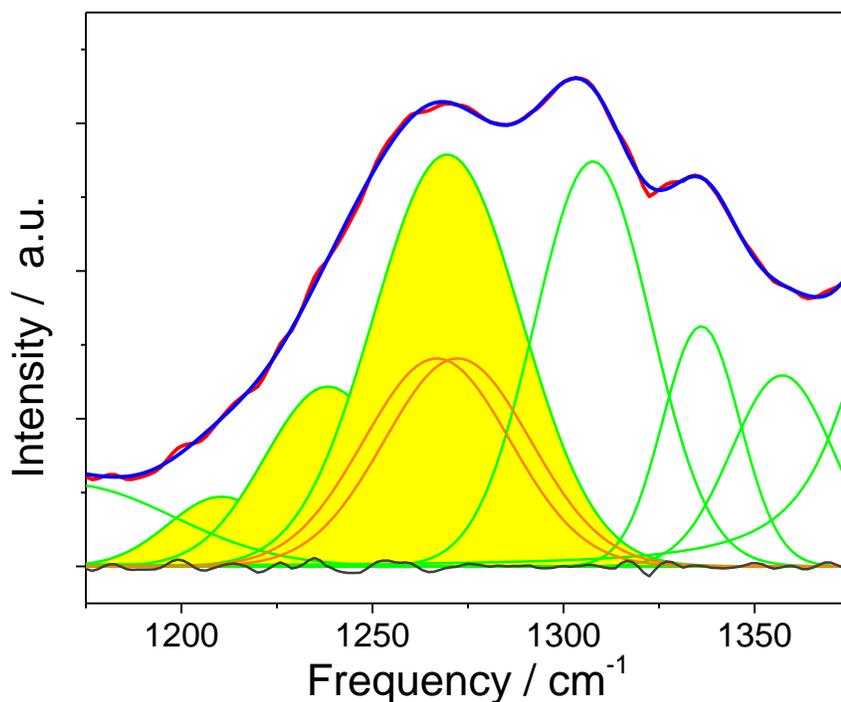


Figure 7.7. Bands resolved Am III region of the room temperature pure α -helix-like UVRR spectrum of XAO-SDS aggregates. This pure α -helix-like XAO-SDS aggregate UVRR spectrum is calculated by averaging the essentially identical pure α -helix-like XAO UVRR spectra at different SDS concentrations as weighted by their respective α -helical fractions. The α -helix-like UVRR spectrum at each SDS concentration is calculated by subtracting the PPII spectrum contribution. The Am III₃ deconvoluted bands are shown in yellow. The two presumed identical ~ 1270 cm^{-1} Am III₃ Gaussian bands derive from the C-terminal and N-terminal peptide bonds in the α -helix segment.

The averaged α -helix-like UVRR spectrum ($S_{\alpha}(v)$) of aggregated XAO is shown in Figure 7.7. The Am III₃ band of this pure α -helix-like UVRR spectrum can be modeled by three Gaussian bands at ~ 1210 cm^{-1} , ~ 1238 cm^{-1} and ~ 1270 cm^{-1} . The ~ 1210 and ~ 1238 cm^{-1} bands

probably resulted from turn structures, while the $\sim 1270\text{ cm}^{-1}$ band likely originates from α -helix-like conformations.

At 20 °C, the Am III₃ band frequency depends on both the ψ Ramachandran dihedral angle and peptide bond hydrogen bonding.^{63,72}

$$\nu = \nu_0 - 54\text{ cm}^{-1} \sin(\psi + 26^\circ) \quad (8)$$

where ν_0 depends on the peptide bond hydrogen bonding. To calculate the ψ angle, ν_0 must be determined from the state of the peptide bond hydrogen bonding. Mikhonin et. al. specified ν_0 for a number of peptide bond hydrogen bonding states.⁷³ Table 2 lists the relevant peptide bond hydrogen bonding states in an α -helix segment and their respective ν_0 values.

Table 7. 2. Hydrogen bonding states of peptide bonds in aqueous solution

peptide bond type	number of peptide bonds	hydrogen bond donor	hydrogen bond acceptor	ν_0 / cm^{-1} (at 20 °C)*
N-terminus in an α -helix	4	H ₂ O	CO	1251.4
center in an α -helix	$n-8$	NH	CO	1244
C-terminus in an α -helix	4	NH	H ₂ O	1246
in PPII-like conformation	-	NH	H ₂ O	1253.8
		H ₂ O	CO	
turns	-	unknown	unknown	1248.8

* values are estimated from ref.⁷³.

For aggregated XAO, we calculated above from the CD and UVRR that 7.6 residues occur in α -helix-like conformations. Thus, there are essentially no peptide bonds where both the carbonyl and amino groups form intra-peptide hydrogen bonds. In the XAO α -helix-like conformations, there are approximately 4 N-terminal and 4 C-terminal peptide bonds. At room

temperature (~ 20 °C), the ν_0 difference between N-terminal and C-terminal peptide bonds is 5.4 cm^{-1} (Table 2) due to their different hydrogen bonding states.⁷³ Therefore, we fit the ~ 1270 cm^{-1} AmIII₃ band to two Gaussian bands that are separated by 5.4 cm^{-1} with identical band areas and identical widths. The low frequency band (~ 1267 cm^{-1}) simulates the C-terminal peptide bonds, while the high frequency (~ 1272 cm^{-1}) band simulates the N-terminal peptide bonds. (Figure 7.7)

We previously developed a method⁷³⁻⁷⁵ to calculate the peptide ψ angle distribution from the measured Am III₃ band. The Am III₃ band of the UVRR spectra from single crystal peptides show a homogenous bandwidth of 15 cm^{-1} .⁷⁴ We deconvoluted the Am III₃ band into a set of Lorentzian bands that have a FWHM of 15 cm^{-1} . The corresponding ψ angle associated with the center frequency (ν_L) of each of these Lorentzian bands was calculated from eq. (8).

The probability of an XAO peptide bond occurring at a particular ψ angle can be calculated from the intensity of the Lorentzian band at this ψ angle, $I(\psi)$:

$$P(\psi) = \frac{I(\psi)}{\Delta\psi \sum_{\psi} I(\psi)} \quad (9)$$

where $\Delta\psi$ is the resolution interval between ψ angles that are determined from the spacing ($\Delta\nu_L$) between deconvoluted Lorentzian frequencies ν_L . Rewriting eq. (8):

$$\Delta\psi(\nu_L) = \frac{\Delta\nu_L}{\sqrt{(54 \text{ cm}^{-1})^2 - (\nu_L - \nu_0)^2}} \quad (10)$$

In the absence of SDS, XAO peptide bonds exist in PPII-like conformations that are fully hydrogen bonded to water. Therefore, the deconvoluted Am III₃ band in this XAO UVRR spectrum can be used to calculate the ψ angle distributions of the PPII-like XAO-SDS aggregate conformations by using eq. (8) with $\nu_0 = 1253.8$ cm^{-1} at ~ 20 °C (Table 2).

The ψ angle distribution of the α -helix-like conformation can be calculated from the Am III₃ band of the XAO pure α -helix-like UVRR spectrum. To reflect the different hydrogen

bonding states of peptide bonds, different ν_0 (Table 2) are used in eq. (8) to calculate the ψ angles from the Figure 7.7 deconvolved different Am III₃ bands.

The aggregated XAO ψ angle distributions shown in Figure 7.8 contain a PPII-like conformation centered at 153° and an α -helix-like region centered at ~ -43°. The α -helix-like region ψ angle distributions deviate from a single Gaussian distribution, indicating multiple α -helix-like conformational states.

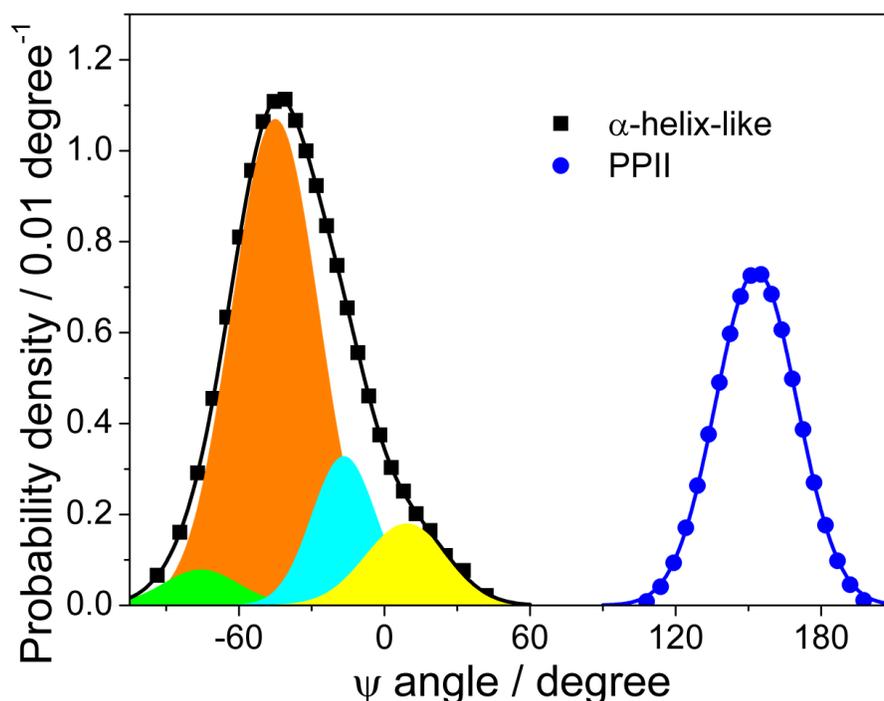


Figure 7. 8. Ramachandran ψ angle distributions calculated for XAO-SDS aggregates from room temperature 204 nm excited UVRR spectra.

The black envelope curve derive from the α -helix-like ψ angle distributions, that were fit to four Gaussians corresponding to π - (filled with green), α - (filled with orange) and 3_{10} - (filled with cyan) helices and turn structures (filled with yellow). The blue curve indicates PPII-like ψ angle distribution in the XAO-SDS aggregates.

Previously, Mikhonin resolved π -, α - and 3_{10} - helices of the AP peptide.⁷⁵ We thus fit the α -helix-like region ψ angle distributions of aggregated XAO to four normal distributions,

modeling the ψ angle distributions as π -, α - and 3_{10} - helices and a turn structure respectively. (Figure 7.8) Three conformational distributions are found at $\psi = -76^\circ$, -45° and -17° , in agreement with the expected α -helix-like ψ angles. An additional conformation is found at 9° , suggesting an additional turn conformation. Table 3 lists the conformations that we resolved in the aggregated XAO.

Table 7. 3. Conformational states resolved in aggregated in the aggregated XAO

conformation	center ψ angle / deg	Standard deviation* / deg
π -helix	-76 ± 30	15 ± 8
α -helix	-45 ± 1	18 ± 5
3_{10} -helix	-17 ± 3	13 ± 3
turn structure	9 ± 6	17 ± 2
PPII	153 ± 1	17 ± 1

* Note torsional force constant is related to standard deviation of the distribution by $f = \frac{RT}{\sigma^2}$.

7.4 CONCLUSIONS

We used CD, NMR and UVRR spectroscopy to demonstrate that XAO, which is predominately PPII conformation in aqueous solution, forms α -helix-like conformations in XAO-SDS aggregates. From the UVRR spectra, we calculated the ψ angle distributions of XAO-SDS aggregates. We resolved α -, π - and 3_{10} - helix conformations, a turn structure and a PPII-like conformation. We studied the XAO-SDS aggregation by measuring the diffusion coefficients of the XAO monomer, the SDS monomer and the XAO-SDS aggregate by using PFG-NMR

spectroscopy. The XAO-SDS aggregates occur at SDS concentrations below the *cmc*. The SDS monomers bind to the XAO peptide and neutralize the four positively charged side chains. Four XAO-SDS₄ further associate to form a stable aggregate of stoichiometry XAO₄-SDS₁₆ where the XAO adopts a predominantly α -helix-like conformation, which also contains turn and π and 3_{10} helices, rather than the predominant PPII conformation of the XAO monomers in aqueous solution. Our work suggests that the XAO-SDS aggregates form a hydrophobic core that minimizes the hydrophobic surface area in contact with water. Neutralization of the flanking charged side chains decreases the electrostatic penalty for α -helix formation. The exterior of the aggregate exposes the faces of the α -helix-like XAO peptide to water. This self assembly mechanism may be relevant to the assembly of antimicrobial peptides—and intrinsically disordered peptides into membranes.

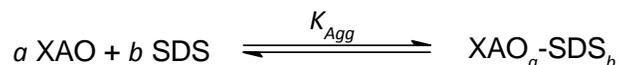
7.5 ACKNOWLEDGEMENT

We are grateful to Dr. Lu Ma, Jonathan Wert and David Punihaole for helpful discussions. This work was supported in part by NIH grant 1RO1 EB009089

7.6 SUPPORTING INFORMATION

7.6.1 Modeling of XAO and SDS interactions

The binding equilibrium between XAO and SDS monomer, assuming only one aggregate species, can be described as



where K_{Agg} is the equilibrium constant for the XAO-SDS aggregate, a and b are the numbers of XAO and SDS molecules in the XAO-SDS aggregate. The equilibrium constant is

$$K_{\text{Agg}} = \frac{C_{\text{Agg}}}{C_{\text{XAO}}^a C_{\text{SDS}}^b} \quad (\text{S1})$$

Applying eq. (6) to diffusion coefficients of XAO and SDS at SDS concentrations < 5 mM where the concentration of SDS micelles is negligible,

$$D_{\text{XAO}}^{\text{App}} = \frac{C_{\text{XAO}} D_{\text{XAO}} + a C_{\text{Agg}} D_{\text{Agg}}}{C_{\text{XAO}}^{\text{Tot}}} \quad (\text{S2})$$

$$D_{\text{SDS}}^{\text{App}} = \frac{C_{\text{SDS}} D_{\text{SDS}} + b C_{\text{Agg}} D_{\text{Agg}}}{C_{\text{SDS}}^{\text{Tot}}} \quad (\text{S3})$$

Combining eq. (S2) and eq. (S3), one can calculate the ratio between b and a ,

$$\frac{b}{a} = \frac{C_{\text{SDS}}^{\text{Tot}} D_{\text{SDS}}^{\text{App}} - C_{\text{SDS}} D_{\text{SDS}}}{C_{\text{XAO}}^{\text{Tot}} D_{\text{XAO}}^{\text{App}} - C_{\text{XAO}} D_{\text{XAO}}} \quad (\text{S4})$$

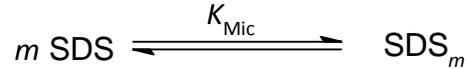
We estimated that $b/a = 3.9 \pm 0.3$ from the data between 1 mM to 4 mM SDS. With $b/a = 4$, we rewrite eq. (S1) as

$$a C_{\text{Agg}} = a K_{\text{Agg}} C_{\text{XAO}}^a C_{\text{SDS}}^{4a} \quad (\text{S5})$$

Recognizing that aC_{Agg} can be calculated from eq. (S2), we could fit the data with eq. (S5) and estimate $a = 4.3 \pm 0.1$ and $K_C = (3.8 \pm 0.1) \times 10^4 \text{mM}^{-20.5}$.

7.6.2 Modeling of SDS Micellation

We modeled the SDS micelle as a cooperative aggregation of m SDS monomers



where K_{Mic} is the micelle formation constant

$$K_{\text{Mic}} = \frac{C_{\text{Mic}}}{C_{\text{SDS}}^m} \quad (\text{S6})$$

Above 10 mM SDS, the SDS species existing in solution consist of SDS monomers, XAO-SDS aggregates and SDS micelles. Applying eq. (6) gives

$$D_{\text{SDS}}^{\text{App}} = \frac{C_{\text{SDS}}D_{\text{SDS}} + bC_{\text{Agg}}D_{\text{Agg}} + mC_{\text{Mic}}D_{\text{Mic}}}{C_{\text{SDS}}^{\text{Tot}}} \quad (\text{S7})$$

At SDS concentrations > 6 mM, the relative XAO monomer concentration is negligible. The contribution from the XAO-SDS aggregate to the diffusion coefficient of SDS is constant. The concentrations of SDS monomer and SDS micelle can be calculated from eq. (S7).

Similar to eq. (S5), we rewrite eq. (S6) as

$$mC_{\text{Mic}} = mK_{\text{Mic}}C_{\text{SDS}}^m \quad (\text{S8})$$

Fitting the data from 6 mM to 50 mM to eq. (S8) gives the aggregation number $m = 38 \pm 8$ and $K_{\text{Mic}} = (1.1 \pm 5.5) \times 10^{-10} \text{mM}^{-37}$.

7.6.3 Supplementary figures

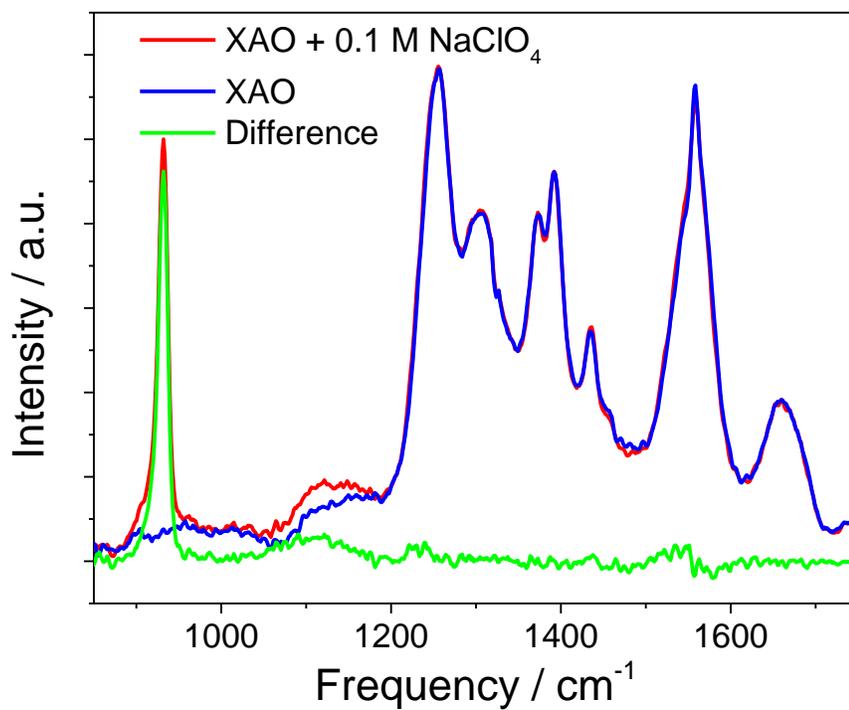


Figure 7. 9. 204 nm excited UVRR spectra of XAO in the presence and in the absence of 0.1 M NaClO₄ and the difference spectrum.

The spectra are almost identical, indicating that the 0.1 M NaClO₄ does not impact the XAO conformation.

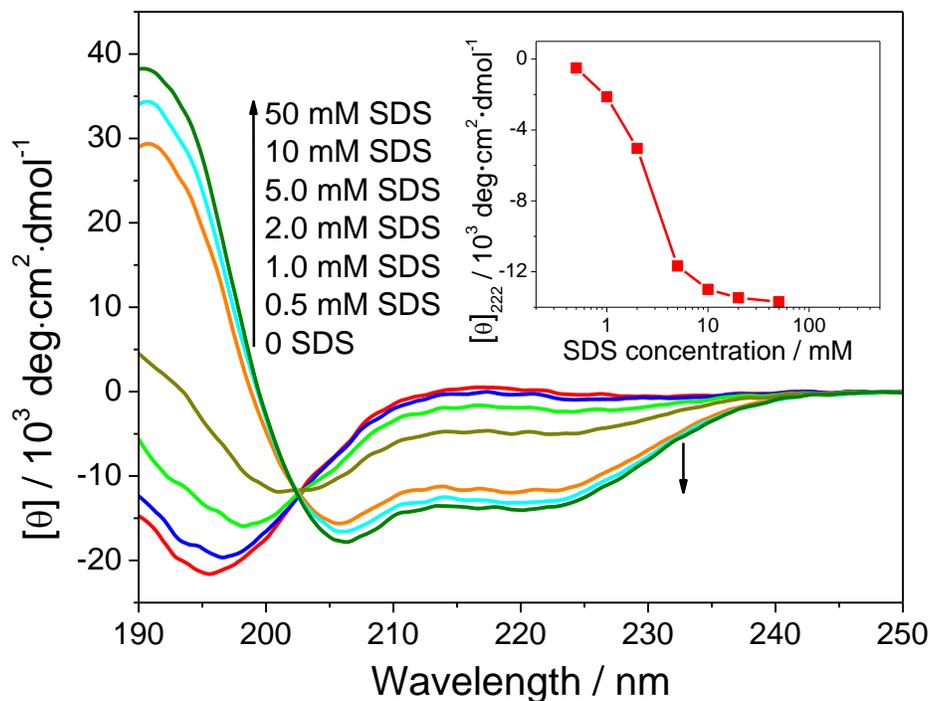


Figure 7. 10. CD spectra of 1.0 mM XAO in the absence of NaClO_4 at different SDS concentration.

The inset shows the dependence on the SDS concentration of the CD spectra 222 nm ellipticity. These XAO CD spectra and their dependence on SDS concentration are essentially identical to those of 1.0 mM XAO in the presence of 0.1 M NaClO_4 , indicating that 0.1 M NaClO_4 has little impact on the XAO conformation. The addition of NaClO_4 decreases the SDS *cmc* significantly from ~8 mM to ~2 mM, however, the XAO conformation dependence on SDS concentration remains similar. SDS micelles are not required for inducing the α -helix-like XAO conformations.

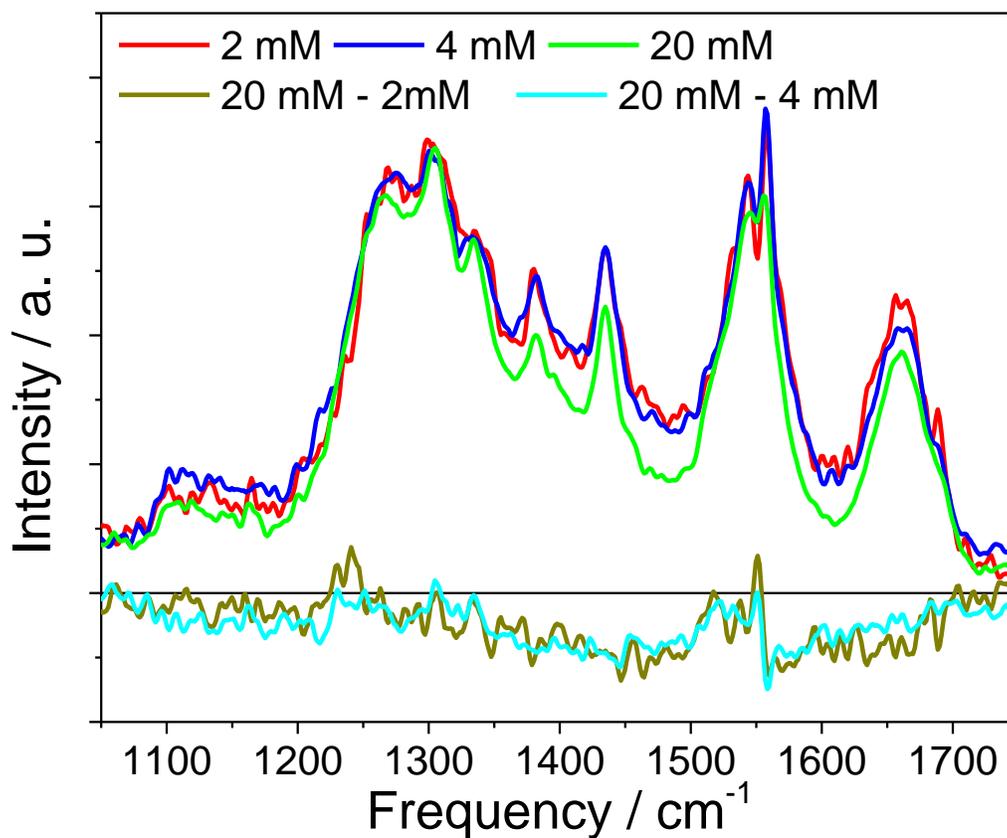


Figure 7. 11. Pure α -helix-like conformation UVR spectra of XAO with 2 mM, 4 mM and 20 mM SDS and their difference spectra.

These pure α -helix-like conformation UVR spectra are similar, indicating that the α -helix-like conformations found at different SDS concentrations are similar.

7.7 BIBLIOGRAPHY

- (1) Uversky, V. N. *Eur. J. Biochem.* **2002**, 269, 2.
- (2) Caughey, B.; Lansbury, P. T. *Annu. Rev. Neurosci.* **2003**, 26, 267.
- (3) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, 75, 333.
- (4) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, 37, 215.
- (5) Abedini, A.; Raleigh, D. P. *Protein Eng., Des. Sel.* **2009**, 22, 453.

- (6) Abedini, A.; Raleigh, D. P. *Phys. Biol.* **2009**, *6*, 015005.
- (7) Butterfield, S. M.; Lashuel, H. A. *Angew. Chem. Int. Ed.* **2010**, *49*, 5628.
- (8) Eliezer, D.; Kutluay, E.; Bussell Jr, R.; Browne, G. *J. Mol. Biol.* **2001**, *307*, 1061.
- (9) Ahmad, M. F.; Ramakrishna, T.; Raman, B.; Rao, C. M. *J. Mol. Biol.* **2006**, *364*, 1061.
- (10) Giehm, L.; Oliveira, C. L. P.; Christiansen, G.; Pedersen, J. S.; Otzen, D. E. *J. Mol. Biol.* **2010**, *401*, 115.
- (11) Dikiy, I.; Eliezer, D. *Biochim. Biophys. Acta. - Biomembranes* **2012**, *1818*, 1013.
- (12) Marcinowski, K. J.; Shao, H.; Clancy, E. L.; Zagorski, M. G. *J. Am. Chem. Soc.* **1998**, *120*, 11082.
- (13) Shao, H.; Jao, S.-c.; Ma, K.; Zagorski, M. G. *J. Mol. Biol.* **1999**, *285*, 755.
- (14) Terzi, E.; Hützemann, G.; Seelig, J. *Biochemistry* **1997**, *36*, 14845.
- (15) Okada, T.; Wakabayashi, M.; Ikeda, K.; Matsuzaki, K. *J. Mol. Biol.* **2007**, *371*, 481.
- (16) Matsuzaki, K. *Biochim. Biophys. Acta. - Biomembranes* **2007**, *1768*, 1935.
- (17) Olofsson, A.; Borowik, T.; Gröbner, G.; Sauer-Eriksson, A. E. *J. Mol. Biol.* **2007**, *374*, 186.
- (18) Jayasinghe, S. A.; Langen, R. *Biochemistry* **2005**, *44*, 12113.
- (19) Knight, J. D.; Hebda, J. A.; Miranker, A. D. *Biochemistry* **2006**, *45*, 9496.
- (20) Lashuel, H. A.; Hartley, D.; Petre, B. M.; Walz, T.; Lansbury, P. T. *Nature* **2002**, *418*, 291.
- (21) Last, N. B.; Miranker, A. D. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6382.
- (22) Lashuel, H. A.; Lansbury, P. T. *Q. Rev. Biophys.* **2006**, *39*, 167.
- (23) Soscia, S. J.; Kirby, J. E.; Washicosky, K. J.; Tucker, S. M.; Ingelsson, M.; Hyman, B.; Burton, M. A.; Goldstein, L. E.; Duong, S.; Tanzi, R. E.; Moir, R. D. *PLoS ONE* **2010**, *5*, e9505.
- (24) Oren, Z.; Shai, Y. *Pept. Sci.* **1998**, *47*, 451.
- (25) Dathe, M.; Wieprecht, T. *Biochim. Biophys. Acta. - Biomembranes* **1999**, *1462*, 71.
- (26) Tossi, A.; Sandri, L.; Giangaspero, A. *Pept. Sci.* **2000**, *55*, 4.
- (27) Powers, J.-P. S.; Hancock, R. E. W. *Peptides* **2003**, *24*, 1681.
- (28) Zelezetsky, I.; Tossi, A. *Biochim. Biophys. Acta. - Biomembranes* **2006**, *1758*, 1436.
- (29) Johansson, J.; Gudmundsson, G. H.; Rottenberg, M. n. E.; Berndt, K. D.; Agerberth, B. *J. Biol. Chem.* **1998**, *273*, 3718.
- (30) Park, C. B.; Yi, K.-S.; Matsuzaki, K.; Kim, M. S.; Kim, S. C. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8245.
- (31) Park, Y.; Park, S.-C.; Park, H.-K.; Shin, S. Y.; Kim, Y.; Hahm, K.-S. *Pept. Sci.* **2007**, *88*, 199.
- (32) Shi, Z. S.; Chen, K.; Liu, Z. G.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877.
- (33) Shi, Z. S.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190.
- (34) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527.
- (35) Kjaergaard, M.; Nørholm, A. B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. *Protein Sci.* **2010**, *19*, 1555.
- (36) Adzhubei, A. A.; Sternberg, M. J. E.; Makarov, A. A. *J. Mol. Biol.* **2013**, *425*, 2100.
- (37) Otvos, J. L. *CMLS, Cell. Mol. Life Sci.* **2002**, *59*, 1138.
- (38) Ferreón, A. C. M.; Gambin, Y.; Lemke, E. A.; Deniz, A. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5645.
- (39) Jalili, S.; Akhavan, M. *Biophys. Chem.* **2011**, *153*, 179.

- (40) Nanga, R. P. R.; Brender, J. R.; Vivekanandan, S.; Ramamoorthy, A. *Biochim. Biophys. Acta. - Biomembranes* **2011**, *1808*, 2337.
- (41) Schibli, D. J.; Hwang, P. M.; Vogel, H. J. *Biochemistry* **1999**, *38*, 16749.
- (42) Rozek, A.; Friedrich, C. L.; Hancock, R. E. W. *Biochemistry* **2000**, *39*, 15765.
- (43) Park, S.; Park, S.-H.; Ahn, H.-C.; Kim, S.; Kim, S. S.; Lee, B. J.; Lee, B.-J. *FEBS Lett.* **2001**, *507*, 95.
- (44) Li, X.; Li, Y.; Han, H.; Miller, D. W.; Wang, G. *J. Am. Chem. Soc.* **2006**, *128*, 5776.
- (45) Japelj, B.; Zorko, M.; Majerle, A.; Pristovšek, P.; Sanchez-Gomez, S.; de Tejada, G. M.; Moriyon, I.; Blondelle, S. E.; Brandenburg, K.; Andrä J.; Lohner, K.; Jerala, R. *J. Am. Chem. Soc.* **2007**, *129*, 1022.
- (46) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. *J. Phys. Chem. Lett.* **2011**, *2*, 334.
- (47) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. *Chem. Rev.* **2012**, *112*, 2604.
- (48) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. *Appl. Spectrosc.* **2005**, *59*, 1541.
- (49) Cotts, R. M.; Hoch, M. J. R.; Sun, T.; Markert, J. T. *J. Magn. Reson.* **1989**, *83*, 252.
- (50) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson. A* **1995**, *115*, 260.
- (51) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288.
- (52) Tanner, J. E. *J. Chem. Phys.* **1970**, *52*, 2523.
- (53) Nilsson, M.; Morris, G. A. *Anal. Chem.* **2008**, *80*, 3777.
- (54) Makowska, J.; Rodziewicz-Motowidło, S.; Bagińska, K.; Vila, J. A.; Liwo, A.; Chmurzyński, L.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1744.
- (55) Makowska, J.; Rodziewicz-Motowidło, S.; Baginska, K.; Makowski, M.; Vila, J. A.; Liwo, A.; Chmurzynski, L.; Scheraga, H. A. *Biophys. J.* **2007**, *92*, 2904.
- (56) Chen, Y.-H.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350.
- (57) Gans, P. J.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. *Biopolymers* **1991**, *31*, 1605.
- (58) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 8074.
- (59) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311.
- (60) Mielke, S. P.; Krishnan, V. V. *Prog. NMR. Spec.* **2009**, *54*, 141.
- (61) Wishart, D. S. *Prog. NMR. Spec.* **2011**, *58*, 62.
- (62) Lam, S. L.; Hsu, V. L. *Biopolymers* **2003**, *69*, 270.
- (63) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775.
- (64) Wang, Y.; Purrello, R.; Jordan, T.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359.
- (65) Lee, S.-H.; Krimm, S. *Biopolymers* **1998**, *46*, 283.
- (66) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. *J. Phys. Chem. B* **2004**, *108*, 19020.
- (67) Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2005**, *109*, 3047.
- (68) Orfi, L.; Lin, M.; Larive, C. K. *Anal. Chem.* **1998**, *70*, 1339.
- (69) Rosen, M. J. *Surfactants and Interfacial Phenomena*; 3rd ed.; Wiley, 2004.
- (70) Dean, J. A. *Lange's Handbook of Chemistry*; 15 ed.; McGraw-Hill, Inc: New York, 1999.
- (71) Dill, K. A.; Bromberg, S. *Molecular Driving Forces*; Garland Science: New York, 2003.
- (72) Myshakina, N. S.; Ahmed, Z.; Asher, S. A. *J. Phys. Chem. B* **2008**, *112*, 11873.
- (73) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2006**, *110*, 1928.

- (74) Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433.
- (75) Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2006**, *128*, 13789.

8.0 DEPENDENCE OF RAMAN AND RESONANCE RAMAN INTENSITIES ON SAMPLE SELF ABSORPTION

This chapter has been submitted to *Appl. Spectros.*

Authors: Zhenmin Hong and Sanford A. Asher

Address: Department of Chemistry, University of Pittsburgh, PA 15260

Resonance Raman cross sections are generally larger than normal or preresonance Raman cross sections. Thus, higher Raman intensities are expected for resonance excitation, especially for backscattering measurements. However, self absorption decreases the observed Raman intensities. In the work here we examine the impact of self absorption on the observed preresonance and resonance Raman intensities. For the simplest case where a single electronic transition dominates the Raman scattering, and where the resonance enhancement scales with the square of the molar absorptivity of the absorption band, theory predicts that for close to resonance excitation the observed Raman intensities monotonically increase as resonance is approached. In the case that an impurity absorbs, the observed Raman intensities may decrease as excitation moves close to resonance for particular conditions of impurity absorption

bandwidths and frequency offsets. Impurity absorption also causes decreases in observed Raman intensities for the more slowly increasing preresonance excitation.

8.1 INTRODUCTION

Raman spectroscopy involves inelastic light scattering.¹⁻³ The scattered light is collected by a collection optic and this scattered light is then spectrally interrogated for its intensity, frequency and polarization to get molecular information on the sample. The vibrational Raman scattering efficiency is generally low with normal analyte Raman cross sections of $\sim 10^{-30} \text{ cm}^2 \cdot \text{molc}^{-1} \cdot \text{str}^{-1}$. The cross sections will significantly increase if there is preresonance Raman enhancement; much larger Raman cross sections can occur with resonance excitation, especially in the deep UV.³⁻⁴ The largest resonance Raman cross sections observed to date are for polycyclic aromatic hydrocarbons such as pyrene where cross sections as large as $\sim 10^{-22} \text{ cm}^2 \cdot \text{molc}^{-1} \cdot \text{str}^{-1}$ occur.⁵

The Raman intensities observed depend upon the values of the Raman cross sections. However, the intensities observed are dramatically impacted by phenomena such as sample self absorption that determines the depth of penetration of the excitation beam and the transmission of the Raman scattering out of the sample.⁶⁻⁸

Typical absorption cross sections are far larger ($\sim 10^{-16} \text{ cm}^2 \cdot \text{molc}^{-1}$) than Raman cross sections, and even resonance Raman cross sections. For strong absorptions the excitation light is attenuated by the sample such that the sample thickness traversed can be limited to only tens of nm, limiting the number of Raman photons generated and observed. Attenuation of the incident light, as well as the Raman scattered light can also result from sample elastic scattering due to

the refractive index inhomogeneities that scatter the light outside of the volume elements that can be imaged into the spectrometer. These phenomena can significantly decrease the Raman intensities observed, which decreases the resulting spectral S/N.

In the work here we examine the impact of self absorption on the observed Raman intensities. These intensities ultimately determine the spectral S/N ratios available in Raman measurements. Optimization of spectral S/N is of great importance in the use of Raman spectroscopy, especially for challenging applications such as standoff detection of hazardous species.

We examine the dependence of Raman intensities on sample absorption in the simplest case of pure samples where a single resonant electronic transition dominates both the Raman scattering as well as the sample absorbance. In this case we show for the lowest order in theory the Raman scattering dispersion is expected to approximately follow the absorption bandshape for optically thick samples. Thus, the maximum Raman spectral S/N measurement will occur for excitation at the maximum of the resonance Raman excitation profiles, which generally occur close to the absorption maximum.

In contrast, if additional absorption bands occur, from impurities or chromophores of the molecule not involved in the resonance Raman enhancement of a particular vibrational mode, these additional absorptions will attenuate the Raman intensities leading to decreased S/N measurements. Depending on the particular conditions, preresonance Raman measurements can show higher intensities and S/N than resonance Raman measurements.

8.2 RESULTS AND DISCUSSION

8.2.1 Dependence of Raman Intensities on Self Absorption Where the Raman Intensities Derive from Preresonance and Resonance Raman Enhancement from a Single Resonant Absorption Band

We assume that all the Raman intensities derive solely from preresonance or resonance excitation from a single dominating electronic transition. We consider a simple backscattering geometry (Figure 8.1) similar to that previously considered by Shriver and Dunn⁶ where an incident excitation beam excites an extended planar sample normal to its surface. For simplicity we assume that we collect the Raman scattering only in the directly backscattered direction. This decreases the self absorption attenuation of the Raman scattered backscattered light at angles $< 180^\circ$, because it neglects the longer pathlengths traversed by this Raman scattering.

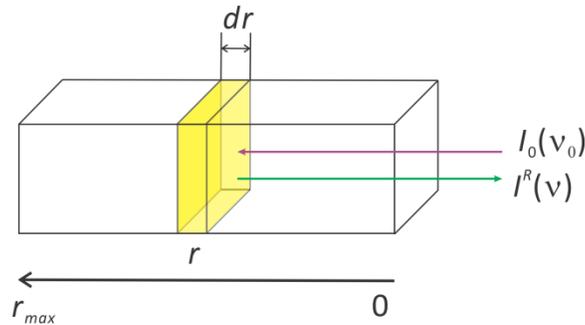


Figure 8. 1. Raman backscattering geometry

Consider a volume element at depth r from the flat sample surface of thickness dr excited by a laser beam of frequency ν_0 with an illuminated cross sectional area of A . The Raman light scattered by this volume element is imaged into a Raman spectrometer and spectrally dispersed and detected. The back-scattered Raman intensity observed at frequency ν is:

$$dI_{\text{ob}}^{\text{R}}(r, \nu_0, \nu) \sim I_0(r, \nu_0) \cdot \sigma^{\text{R}}(\nu_0, \nu) \cdot c \cdot N_{\text{A}} \cdot \frac{\Omega}{4\pi} \cdot K \cdot F(\nu_0, \nu, r, \varepsilon(\nu_0), \varepsilon(\nu)) \cdot A \cdot dr \quad (8.1)$$

where $I_0(r, \nu_0)$ is the incident excitation beam intensity at r , $\sigma^{\text{R}}(\nu_0, \nu)$ is the absolute Raman scattering cross section excited at ν_0 for a Raman band that occurs at ν , c is the analyte concentration, N_{A} is Avagadro's number and Ω is the solid angle of the Raman scattered intensity collected by the spectrometer. K is the spectrometer efficiency and $F(\nu_0, \nu, r, \varepsilon(\nu_0), \varepsilon(\nu))$ is the attenuation factor at the excitation and the Raman scattered light frequencies. For samples that do not elastically scatter light due to refractive index inhomogeneities this factor results solely from absorption by the sample. $\varepsilon(\nu_0)$ and $\varepsilon(\nu)$ are the effective molar absorptivities at the excitation and Raman scattered frequencies. The exciting light is attenuated as:

$$I(r, \nu_0) = I_0(\nu_0) \cdot e^{-2.303 \cdot \varepsilon(\nu_0) \cdot c \cdot r} \quad (8.2)$$

The back-scattering Raman light is attenuated by its transmission through the distance r through the sample:

$$dI_{\text{ob}}^{\text{R}}(r, \nu_0, \nu) = dI^{\text{R}}(r, \nu_0, \nu) \cdot e^{-2.303 \cdot \varepsilon(\nu) \cdot c \cdot r} \quad (8.3)$$

where $dI^{\text{R}}(r, \nu_0, \nu)$ is the backscattered Raman intensity generated at position r in the sample.

Thus, the attenuation factor for the contribution of a volume element at depth r is:

$$F(\nu_0, \nu, r, \varepsilon(\nu_0), \varepsilon(\nu)) = e^{-2.303 \cdot (\varepsilon(\nu_0) + \varepsilon(\nu)) \cdot c \cdot r} \quad (8.4)$$

Therefore, the total Raman intensity detected is the integral of $dI_{\text{ob}}^{\text{R}}(r, \nu_0, \nu)$ over the sample thickness illuminated by the incident excitation beam:

$$\begin{aligned}
I_{\text{ob}}^{\text{R}}(\nu_0, \nu) &\sim \int_0^{r_{\text{max}}} dI_{\text{ob}}^{\text{R}}(r, \nu_0, \nu) dr \\
&= I_0(\nu_0) \cdot A \cdot \sigma^{\text{R}}(\nu_0, \nu) \cdot c \cdot N_{\text{A}} \cdot \frac{\Omega}{4\pi} \cdot K \\
&\cdot \int_0^{r_{\text{max}}} e^{-2.303 \cdot (\varepsilon(\nu_0) + \varepsilon(\nu)) \cdot c \cdot r} dr
\end{aligned} \tag{8.5}$$

For a sample of thickness r_{max} ,

$$I_{\text{ob}}^{\text{R}}(\nu_0, \nu) \sim \frac{I_0(\nu_0) \cdot A \cdot \sigma^{\text{R}}(\nu_0, \nu) \cdot N_{\text{A}} \cdot \Omega \cdot K}{2.303 \cdot 4\pi \cdot (\varepsilon(\nu_0) + \varepsilon(\nu))} (1 - e^{-2.303 \cdot (\varepsilon(\nu_0) + \varepsilon(\nu)) \cdot c \cdot r_{\text{max}}}) \tag{8.6}$$

The observed Raman intensity increases with the Raman cross section and decreases with the absorbance at the excitation and Raman scattering frequencies. The Raman intensities observed are relatively independent of concentration⁶ but are decreased for sample thicknesses less than that necessary to fully attenuate the excitation beam.

The Raman scattering cross sections are proportional to the fourth power of the excitation frequency times the square of Raman polarizability:⁹

$$\sigma^{\text{R}}(\nu_0, \nu) \sim \nu_0^4 |\alpha(\nu_0, \nu)|^2 \tag{8.7}$$

Using Kramer-Heisenberg-Dirac second order perturbation theory we can write down an expression for the resonance Raman polarizability in the lowest order in Raman theory as detailed through the Albrecht A-term expression:⁹

$$\alpha(\nu_0, \nu) \sim \langle g | \hat{P} | e \rangle \langle e | \hat{P} | g \rangle \cdot FC \cdot D(\nu_e, \nu_0, \nu) \tag{8.8}$$

where the transition moment integrals occur over the electric dipole operator, \hat{P} between the electronic ground state and the resonant electronic excited state. FC is the factor associated with the Franck-Condon integrals for the vibrational mode selection factor of the Raman polarizability, while $D(\nu_e, \nu_0, \nu)$ is the resonant excitation energy denominator that determines the dispersion of the Raman cross section due to preresonance and resonance effects.

The square of the electronic transition moment is roughly *proportional* to the maximum molar absorptivity, $\varepsilon(\nu_e)$, where ν_e is the frequency at which the electronic transition occurs. Thus, the resonance Raman cross section for the vibration that scatters at frequency ν is proportional to the square of the molar absorptivity.⁹

$$\sigma^R(\nu_0) \sim k \cdot \nu_0^4 \cdot \varepsilon(\nu_0)^2 \quad (8.9)$$

In strict resonance the Raman cross section dispersion will approximately track the absorption bandshape with higher frequency dependence if the absorption is dominated by inhomogeneous broadening. In this case the absorption bandshape results from a set of molecules with a distribution of transition frequencies. The resonance Raman excitation profile derives from the sum of the contributions of the excitation profiles from each of these transitions.

If the absorption bandshape is dominated by homogeneous broadening, the relationship between the absorption and the resonance Raman excitation profile bandshape becomes somewhat more complex. However, the following conclusions are expected to approximately hold for this case as well.

By approximating $\varepsilon(\nu_0) = \varepsilon(\nu)$, (for deep UV Raman excitation, the Stokes shift frequency is small compared to the excitation frequency) we can simplify eq. (8.6),

$$I_{\text{ob}}^R(\nu_0) \sim \frac{k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \varepsilon(\nu_0)^2 \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot \varepsilon(\nu_0)} (1 - e^{-4.606 \cdot \varepsilon(\nu_0) \cdot c \cdot r_{\text{max}}}) \quad (8.10)$$

If the product of the molar absorptivity $\varepsilon(\nu_0)$, the concentration c , and the sample pathlength is large, the exponential term becomes zero. Eq. (8.10) becomes:

$$I_{\text{ob}}^R(\nu_0) \sim \frac{k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \varepsilon(\nu_0) \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi} \quad (8.11)$$

This expression indicates that even in the presence of self absorption the largest Raman intensities will occur at the absorption spectral maximum. Resonance Raman excitation will

always give rise to higher intensities and higher S/N than will normal and preresonance Raman measurements. The back-scattered Raman intensities become proportional to the molar absorptivity $\varepsilon(\nu_0)$ and are independent of concentration.

The observed Raman intensities track the analyte absorption band, but are biased to higher frequency by the ν_0^4 scattering dependence. This is the behavior expected from close to resonance excited Raman bands whose preresonance intensities are enhanced by a single electronic transition that is situated at higher frequency.

An example of this type of enhancement occurs for the symmetric NO_3^- stretch at $\sim 1050 \text{ cm}^{-1}$ in NH_4NO_3 and in NaNO_3 , whose entire preresonance and resonance enhancement derives from its $\sim 220 \text{ nm}$ lowest energy absorption band.¹⁰⁻¹² In this case all of the absorption oscillator strength of molecule is located in the NO_3^- moiety in aqueous solution. Water absorption and Na^+ and NH_4^+ counterion absorptions lie deeper in the UV and negligibly contribute for $\lambda_0 < 200 \text{ nm}$.

If the self absorption is small (conditions where $\varepsilon(\nu_0) \cdot c \cdot r_{\text{max}}$ is small) we can expand the exponential term in eq. (8.10) to obtain:

$$I_{\text{ob}}^{\text{R}}(\nu_0) \sim \frac{\Omega}{4\pi} k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \varepsilon(\nu_0)^2 \cdot A \cdot N_A \cdot K \cdot c \cdot r_{\text{max}} \quad (8.12)$$

In this case the back-scattered Raman intensity more strongly increases with analyte absorption since it is proportional to $\varepsilon(\nu_0)^2$. For low absorption r_{max} is not simply the sample thickness. It is also determined by the depth of focus of the collection lens that limits the volume from which the Raman intensity can be collected. Eq. (12) predicts that the observed intensity increases without limit as the sample thickness increases. The observed intensity now depends upon the analyte concentration.

8.2.2 Dependence of Raman Intensities on Self Absorption Where the Sample Contains Absorption Bands That Do Not Enhance The Analyte Raman Bands

In this case, self absorption decreases the observed Raman intensities compared to the case above where the single absorption band enhances the Raman intensities. This can be shown from eq. (8.13) which explicitly includes an absorption band not involved in resonance Raman enhancement.

$$I_{\text{obi}}^{\text{R}}(\nu_0) \sim \frac{k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \varepsilon_a(\nu_0)^2 \cdot c_a \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot (\varepsilon_a(\nu_0)c_a + \varepsilon_i(\nu_0)c_i)} \left(1 - e^{-4.606(\varepsilon_a(\nu_0)c_a + \varepsilon_i(\nu_0)c_i)r_{\text{max}}}\right) \quad (8.13)$$

where ε_a and c_a are the molar absorptivity and the concentration of the analyte that is resonance enhanced. ε_i and c_i are the molar absorptivity and the concentration of an impurity or absorbing chromophore that does not enhance the analyte Raman bands. This additional absorption could include for example, chromophores such as Cs^+ in solutions of CsNO_3 . The NO_3^- in aqueous solution shows the same Raman excitation profile as occurs in the presence of the transparent NH_4^+ and Na^+ counterions. However the Cs^+ absorption will attenuate the observed Raman intensities. The additional absorptions may also derive from impurities, or from other electronic transitions of a macromolecular analyte, or derive from photochemically generated impurities.

Eq. (8.13) points out that for an optically thick sample the observed Raman intensity will be attenuated by:

$$R = \frac{I_{\text{obi}}^{\text{R}}(\nu_0)}{I_{\text{ob}}^{\text{R}}(\nu_0)} = \frac{\varepsilon_a(\nu_0)c_a}{\varepsilon_a(\nu_0)c_a + \varepsilon_i(\nu_0)c_i} \quad (8.14)$$

It should be noted that the attenuation calculated is strictly for the case where the absorbing species are intimately intermixed with the analyte. It would not be valid for detection of surface analyte particles that are not subject to interferent absorptions.

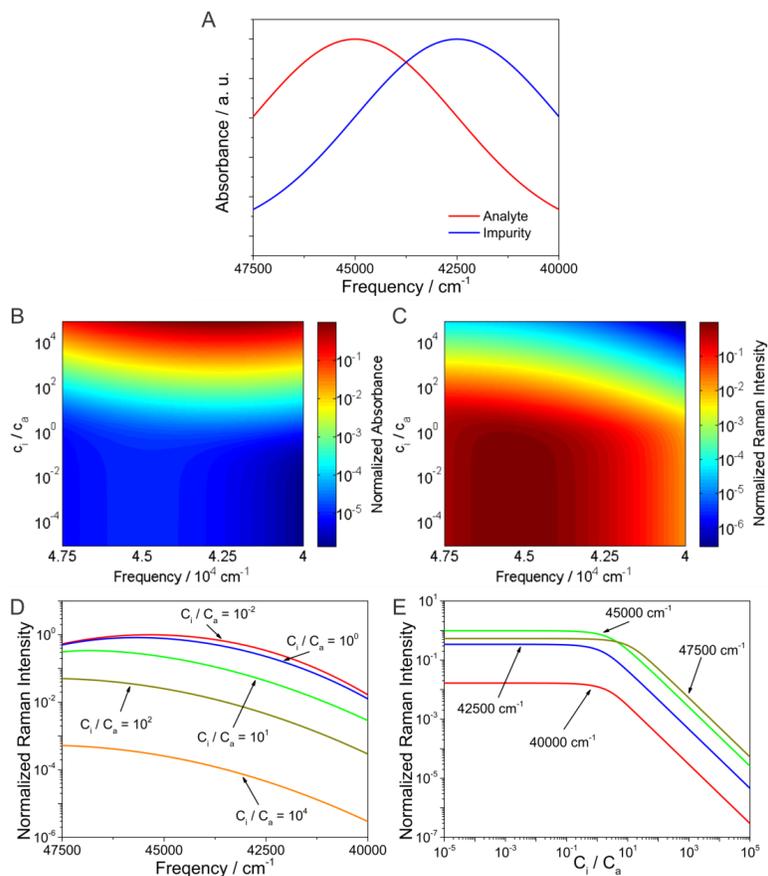


Figure 8. 2. Impact of impurity absorption on observed Raman intensity - case 1

(A) Analyte and impurity absorptions. The absorption bands are identical Gaussians, with $\sigma_a = \sigma_i = 2500 \text{ cm}^{-1}$, and with maxima at $\nu_{ac} = 45,000 \text{ cm}^{-1}$ and $\nu_{ic} = 42,500 \text{ cm}^{-1}$, respectively. (B) Contour plot of calculated frequency dependence of sample absorbance as a function of relative impurity concentration. The normalized absorbance is calculated by dividing by the maximum absorbance that occurs at the maximum impurity concentration of $c_i/c_a = 10^5$. (C) Observed contour plot of the frequency dependence of the relative resonance Raman intensities as a function of the relative impurity concentration calculated by using eq. (8.13). The resonance Raman intensities are normalized to the maximum Raman intensity that occurs for $c_i/c_a = 10^{-5}$. (D) Frequency dependence of normalized Raman intensities at different relative impurity concentrations. (E) Relative impurity concentration dependence of Raman intensities

Figure 8.2 shows the calculated sample absorbance and resonance Raman intensities as a function of frequency and relative impurity concentrations (c_i/c_a). At low impurity concentrations, the sample absorbance mainly derives from the analyte. As the impurity concentration increases, the absorbance becomes dominated by the impurity absorbance (Figure 8.2B).

For all impurity concentrations, the observed Raman intensities increase dramatically as the excitation frequencies move into resonance (Figure 8.2C). There is no case where impurity absorption causes a dip in the Raman intensities for this set of parameters. However, the Raman intensities do decrease as the impurity concentrations increase.

Figure 8.2D shows that the Raman intensity maxima slightly shift towards higher frequencies because the increased impurity absorbance (maximum at $42,500\text{ cm}^{-1}$) attenuates the observed Raman intensity on low frequency side. Figure 8.2E shows the dependence of Raman intensity on the relative impurity concentration (c_i/c_a) at different excitation frequencies. At low impurity concentrations, the Raman intensity is little affected. The Raman intensities decrease as the impurity concentrations increase.

The approximation that the Raman cross section increases with the square of the molar absorptivity, which is relevant for true resonance excitation, predicts that the maximum intensities occur for resonance excitation. In most cases the square of the analyte absorption grows faster with frequency than does the sample absorption. The observed resonance Raman intensity increases with excitation frequency. In the rare cases we examine below, attenuation of the observed Raman intensities caused by impurity absorption could exceed the resonance enhancement.

We should also mention that much more complex situations can occur for more complex samples. For example, analytes may have multiple absorption bands. In this case constructive and destructive interference can occur between different electronic transitions.¹³⁻¹⁴ This can complicate the behavior of the Raman excitation profiles, especially for excitation at frequencies between absorption bands.

8.2.3 Conditions for Sample Self Absorption That Causes Resonance Raman Intensity

Minima

Figure 8.2 demonstrates that the observed Raman intensities monotonically increase as the excitation frequency approaches resonance even in the presence of an absorbing impurity. For optically thick samples, the exponential term in eq. (8.13) vanishes.

$$I_{\text{obi}}^{\text{R}}(\nu_0) \sim \frac{k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \epsilon_a(\nu_0)^2 \cdot c_a \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot (\epsilon_a(\nu_0)c_a + \epsilon_i(\nu_0)c_i)} \quad (8.15)$$

In general, the observed resonance Raman intensities monotonically increase as the excitation frequency is tuned into resonance due to the dependence on the molar absorptivity squared in the numerator. However, it is possible to have situations where the impurity absorbs in such a way that the observed resonance Raman intensities decrease as the excitation frequency approaches resonance.

Below we determine the conditions for decreasing Raman intensities as the excitation frequency approaches resonance. For simplicity we assume Gaussian absorption bandshapes. This is also motivated by the fact that in the condensed phase, inhomogeneous broadening often dominates absorption bandshapes.

$$\varepsilon_a(\nu_0) = \varepsilon_{ac} e^{-\frac{(\nu_0 - \nu_{ac})^2}{2\sigma_a^2}} \quad (8.16)$$

$$\varepsilon_i(\nu_0) = \varepsilon_{ic} e^{-\frac{(\nu_0 - \nu_{ic})^2}{2\sigma_i^2}}$$

where ε_{ac} and ε_{ic} are the analyte and impurity maximum molar absorptivities, ν_{ac} and ν_{ic} are the analyte and impurity absorption frequency maxima, and σ_a and σ_i are standard deviations of the analyte and impurity absorption Gaussian bands, respectively.

We shall limit our discussion to the resonance excitation region in the case that the impurity absorption maximum occurs to lower frequency than the resonance absorption: $\nu_{ac} - 2\sigma_a \leq \nu_0 \leq \nu_{ic} \leq \nu_{ac}$. Obviously, the observed Raman intensity decrease for close to resonance excitation requires $\varepsilon_i(\nu_0)c_i \gg \varepsilon_a(\nu_0)c_a$. We can rewrite eq. (8.15) to concentrate on the absorption factors by substituting eq. (8.16) into eq. (8.15):

$$I_{obi}^R(\nu_0) \sim \frac{k \cdot I_0(\nu_0) \cdot \nu_0^4 \cdot \left(\varepsilon_{ac} e^{-\frac{(\nu_0 - \nu_{ac})^2}{2\sigma_a^2}} \right)^2 \cdot c_a \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot c_i \cdot \varepsilon_{ic} \cdot e^{-\frac{(\nu_0 - \nu_{ic})^2}{2\sigma_i^2}}} \quad (8.17)$$

In the UV region, $\nu_0 \gg \sigma_a$, therefore, the ν_0^4 factor changes little between $\nu_{ac} - 2\sigma_a$ and ν_{ac} , compared to the Gaussian factors. Thus, we conveniently ignore this variation and include the ν_0^4 factor in a constant K' which collects all constants and is independent on ν_0 .

$$I_{obi}^R(\nu_0) \sim \frac{K' c_a \left(\varepsilon_{ac} e^{-\frac{(\nu_0 - \nu_{ac})^2}{2\sigma_a^2}} \right)^2}{c_i \cdot \varepsilon_{ic} \cdot e^{-\frac{(\nu_0 - \nu_{ic})^2}{2\sigma_i^2}}} \quad (8.18)$$

As the excitation frequency increases, a decreased observed Raman intensity occurs if the derivative of $I_{obi}^R(\nu_0)$ with respect to ν_0 becomes negative:

$$\frac{dI_{\text{obi}}^{\text{R}}(v_0)}{dv_0} = I_{\text{obi}}^{\text{R}}(v_0) \frac{d \ln I_{\text{obi}}^{\text{R}}(v_0)}{dv_0} < 0 \quad (8.19)$$

Since $I_{\text{obi}}^{\text{R}}(v_0) > 0$, eq. (8.19) becomes,

$$\frac{d \ln I_{\text{obi}}^{\text{R}}(v_0)}{dv_0} = -\frac{2(v_0 - v_{\text{ac}})}{\sigma_{\text{a}}^2} + \frac{(v_0 - v_{\text{ic}})}{\sigma_{\text{i}}^2} < 0 \quad (8.20)$$

Collecting terms and rearranging,

$$(\sigma_{\text{a}}^2 - 2\sigma_{\text{i}}^2)v_0 < \sigma_{\text{a}}^2 v_{\text{ic}} - 2\sigma_{\text{i}}^2 v_{\text{ac}} \quad (8.21)$$

We denote

$$\beta \equiv \frac{\sigma_{\text{i}}}{\sigma_{\text{a}}}$$

$$\lambda \equiv \frac{v_{\text{ac}}}{\sigma_{\text{a}}} \quad (8.22)$$

$$\delta \equiv \frac{v_{\text{ac}} - v_{\text{ic}}}{\sigma_{\text{a}}} = \lambda - \frac{v_{\text{ic}}}{\sigma_{\text{a}}}$$

Obviously $\beta > 0$, $\lambda > 0$ and $\delta \geq 0$. Using eq. (8.22), we further replace σ_{i} , v_{ac} and v_{ic} with $\beta\sigma_{\text{a}}$, $\lambda\sigma_{\text{a}}$ and $(\lambda - \delta)\sigma_{\text{a}}$, respectively. Eq. (8.22) becomes,

$$(1 - 2\beta^2)\sigma_{\text{a}}^2 v_0 < \sigma_{\text{a}}^2((\lambda - \delta)\sigma_{\text{a}}) - 2(\beta\sigma_{\text{a}})^2 \lambda\sigma_{\text{a}} \quad (8.23)$$

Cancelling σ_{a}^2 on both sides and rearranging gives,

$$(1 - 2\beta^2)v_0 < (\lambda - \delta)\sigma_{\text{a}} - 2\lambda\beta^2\sigma_{\text{a}} \quad (8.24)$$

If $(1 - 2\beta^2) > 0$, $0 < \beta < 1/\sqrt{2}$, then

$$v_0 < \frac{(\lambda - \delta)\sigma_{\text{a}} - 2\lambda\beta^2\sigma_{\text{a}}}{(1 - 2\beta^2)} \quad (8.25)$$

Since $v_{\text{ac}} - 2\sigma_{\text{a}} \leq v_0$, the following must hold,

$$v_{\text{ac}} - 2\sigma_{\text{a}} = \lambda\sigma_{\text{a}} - 2\sigma_{\text{a}} < \frac{(\lambda - \delta)\sigma_{\text{a}} - 2\lambda\beta^2\sigma_{\text{a}}}{(1 - 2\beta^2)} \quad (8.26)$$

Multiplying both sides by $(1 - 2\beta^2)/\sigma_a$, and rearranging, we obtain $\delta < 2(1 - 2\beta^2)$. Since $\nu_0 \leq \nu_{ic}$ and $\nu_0 < \frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)}$ (eq. (25)), we want to compare the values of $\frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)}$ and ν_{ic} , and choose the smaller quantity as the upper limit for ν_0 . It can be shown that the following holds, because $\beta > 0$, $\delta \geq 0$.

$$\frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)} \leq (\lambda - \delta)\sigma_a = \nu_{ic} \quad (8.27)$$

Therefore, $\nu_0 < \frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)} \leq \nu_{ic}$.

In the case that $(1 - 2\beta^2) < 0$, where $\beta > 1/\sqrt{2}$, eq. (8.24) has a different solution,

$$\nu_0 > \frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)} \quad (8.28)$$

Since $\nu_0 \leq \nu_{ic}$, together with eq. (8.28) leads to

$$\frac{(\lambda - \delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1 - 2\beta^2)} < \nu_0 \leq \nu_{ic} = (\lambda - \delta)\sigma_a \quad (8.29)$$

Note $(1 - 2\beta^2) < 0$ in this case. Rearranging eq. (8.29),

$$2\beta^2\delta < 0 \quad (8.30)$$

Obviously, eq. (8.30) is not a relevant solution, because $\beta > 0$, $\delta \geq 0$.

Thus, we conclude that for the observed Raman intensity to show a self absorption decrease in the resonance excitation frequency range between $\nu_{ac} - 2\sigma_a$ and ν_{ac} requires:

- (1) The bandwidth of the impurity absorption must be less than the resonant analyte absorption: $\sigma_i/\sigma_a < 1/\sqrt{2}$;
- (2) The frequency difference between the analyte absorption maximum and the impurity absorbance maximum: $\nu_{ac} - \nu_{ic} \leq 2\sigma_a \left(1 - 2\left(\frac{\sigma_i}{\sigma_a}\right)^2\right)$.

When these two requirements are satisfied, the observed Raman intensity minimum is

expected at $\nu_0 = \frac{(\lambda-\delta)\sigma_a - 2\lambda\beta^2\sigma_a}{(1-2\beta^2)}$. Our neglect of the ν_0^4 dependence only causes small errors.

The above analysis considers only a single impurity absorption band. In the cases of multiple impurity absorptions, the overall impurity absorption bandshape could much deviate significantly from a Gaussian bandshape. In this case observation of a minimum still requires that the impurity absorption increases faster with frequency than does the Raman cross section.

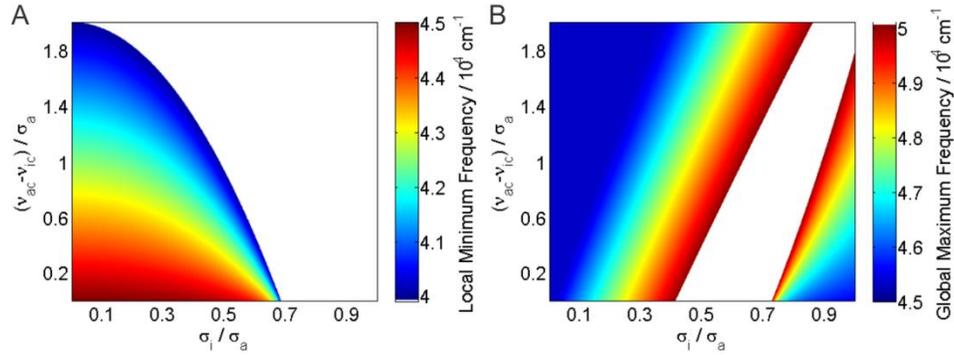


Figure 8.3. Observed Raman intensity excitation frequency local minimum and maximum

(A) Dependencies of the observed Raman intensity excitation frequency local minimum on relative absorption bandwidths (σ_i/σ_a) and absorption band maximum frequency offsets ($(\nu_{ac} - \nu_{ic})/\sigma_a$). (B) Maximal observed resonance Raman excitation frequency dependencies on relative absorption bandwidths (σ_i/σ_a) and absorption band maximum frequency offsets ($(\nu_{ac} - \nu_{ic})/\sigma_a$). The local minimum frequencies, local maximum frequencies, and resonance maximum frequencies are calculated from the observed Raman intensities of eq. (8.13) with $\nu_{ac} = 45,000 \text{ cm}^{-1}$, $\sigma_a = 2500 \text{ cm}^{-1}$, $c_i/c_a = 10^5$, and the corresponding δ and β . The blank regions indicate regions without minima or maxima in the domain of interest ($[\nu_{ac} - 2\sigma_a, \nu_{ac} + 2\sigma_a]$). The boundary in (A) is

approximate, due to neglecting ν_0^4

Figure 8.3 displays the frequency of the minima (if any) of the observed Raman intensities calculated from eq. (13) above. Figure 8.3A shows blank areas above the line $\frac{\nu_{ac} - \nu_{ic}}{\sigma_a} = 2 \left(1 - 2 \left(\frac{\sigma_i}{\sigma_a} \right)^2 \right)$ where no minima occur over the range where eq. (8.9) is valid. Below this boundary, the observed Raman intensity minima decrease in frequency as the impurity

absorption band broadens and as the impurity absorption band downshifts from the analyte absorption. Figure 8.3B shows the observed resonance Raman maximum excitation frequencies.

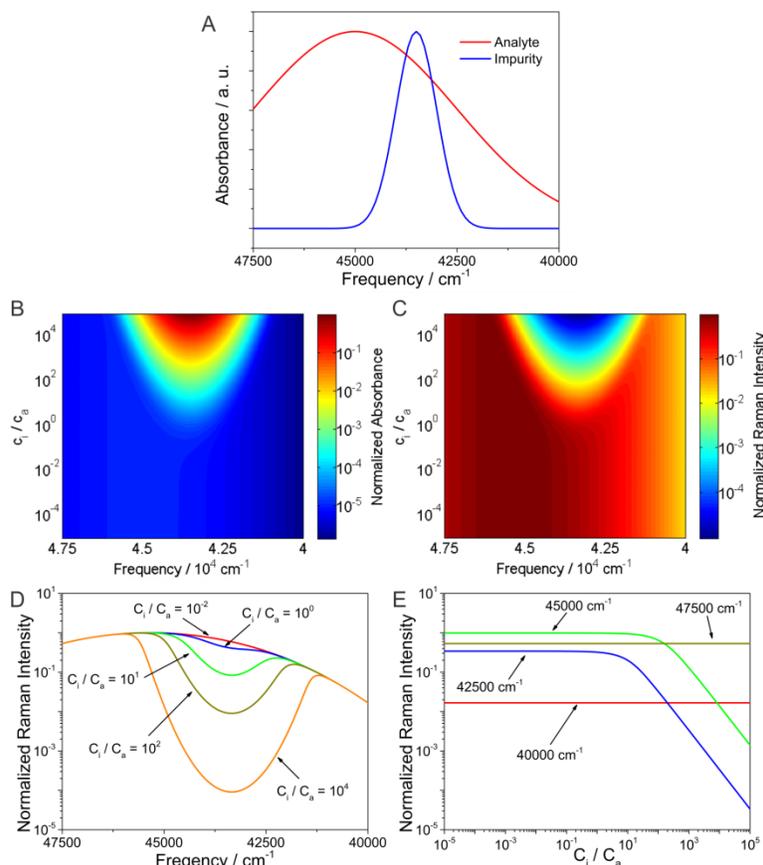


Figure 8. 4. Impact of impurity absorption on observed Raman intensity - case 2

(A) Analyte and impurity absorption spectra. The analyte absorption band maximum frequency is $\nu_{ac} = 45,000 \text{ cm}^{-1}$ with $\sigma_a = 2500 \text{ cm}^{-1}$. The impurity absorption band maximum frequency is $\nu_{ic} = 43,500 \text{ cm}^{-1}$ with $\sigma_i = 500 \text{ cm}^{-1}$. ($\sigma_i/\sigma_a = 0.2$, $(\nu_{ac} - \nu_{ic})/\sigma_a = 0.6$) (B) Contour plot of calculated frequency dependence of sample absorbance as a function of relative impurity concentrations. The normalized absorbance is calculated by dividing by the maximum absorbance that occurs at the maximum impurity concentration of $c_i/c_a = 10^5$. (C) Contour plot of the frequency dependence of the relative resonance Raman intensity as a function of the relative impurity concentration calculated from eq. (8.13). The resonance Raman intensities are normalized to the maximum Raman intensity that occurs for $c_i/c_a = 10^{-5}$. (D) Frequency dependence of normalized Raman intensities at different relative impurity concentrations. (E) Relative impurity concentration dependence of observed Raman intensities.

Figure 8.2 calculates the normalized Raman intensity at $\sigma_i/\sigma_a = 1$ and $(\nu_{ac} - \nu_{ic})/\sigma_a = 1$ and in a region where no local minima occur in the excitation frequency dependent Raman intensities. In contrast, Figure 8.4 calculates the normalized Raman intensities where minima exist for $\sigma_i/\sigma_a = 0.2$ and $(\nu_{ac} - \nu_{ic})/\sigma_a = 0.6$.

As expected, low impurity concentrations have little impact on the observed Raman intensities. However, impurity concentrations comparable to or greater than the analyte give rise to observed Raman intensity minima localized where the impurity has strong absorption (Figure 8.4C). Beyond this frequency region, the observed Raman intensities are independent of impurity concentration. (Figure 8.4E)

8.2.4 Dependence of Preresonance Raman Intensities on Self Absorption

In the cases above we assumed that the resonance Raman enhancement scaled with the square of the molar absorptivity. Thus, the Raman intensities will always be maximized with resonance excitation, unless the impurity absorbance increases faster than does square of the analyte absorbance.

If the resonance Raman intensities increase at a slowly rate with frequency, impurity absorption can dramatically impact the preresonance Raman intensities. This is the case where the preresonance Raman cross sections follow the preresonance Raman Albrecht A-term:⁹

$$\sigma^R(\nu_0) \sim k' \nu_0^4 \left[\frac{\nu_e^2 + \nu_0^2}{(\nu_e^2 - \nu_0^2)^2} \right]^2 \quad (8.31)$$

where ν_0 is the excitation frequency, and ν_e is the frequency of the electronic transition giving rise to preresonance Raman enhancement. The ν_0 excited observed Raman intensity for a sample containing impurity absorption is,

$$I_{\text{obi}}^{\text{R}}(\nu_0) \sim \frac{k' \cdot I_0(\nu_0) \cdot \nu_0^4 \left[\frac{\nu_e^2 + \nu_0^2}{(\nu_e^2 - \nu_0^2)^2} \right]^2 \cdot c_a \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot \varepsilon_i(\nu_0) c_i} (1 - e^{-4.606 \cdot \varepsilon_i(\nu_0) c_i r_{\text{max}}}) \quad (8.32)$$

In preresonance Raman the absorbance of the analyte is negligible. For optically thick samples, the exponential term vanishes and eq. (8.32) becomes

$$I_{\text{obi}}^{\text{R}}(\nu_0) \sim \frac{k' \cdot I_0(\nu_0) \cdot \nu_0^4 \left[\frac{\nu_e^2 + \nu_0^2}{(\nu_e^2 - \nu_0^2)^2} \right]^2 \cdot c_a \cdot A \cdot N_A \cdot \Omega \cdot K}{4.606 \cdot 4\pi \cdot \varepsilon_i(\nu_0) c_i} \quad (8.33)$$

If we assume that the impurity absorption band is Gaussian, we find cases where impurity absorption decreases the Raman intensities as the excitation frequency increases towards resonance.

Figure 8.5 shows the simulated sample absorbance and observed preresonance Raman intensities as function of frequency and relative impurity concentration (c_i/c_a). At very low impurity concentrations, the sample shows little absorption. At high impurity concentrations, it only shows impurity absorbance (Figure 8.5B). With low impurity concentrations, the observed preresonance Raman intensities monotonically increase with excitation frequency. The rate is slower than in the resonance case. In contrast, at high impurity concentrations, the observed preresonance Raman intensities show minima at $\sim 26,000 \text{ cm}^{-1}$ (Figure 8.5D). The observed preresonance Raman intensities decrease as the impurity concentration increases (Figure 8.5E).

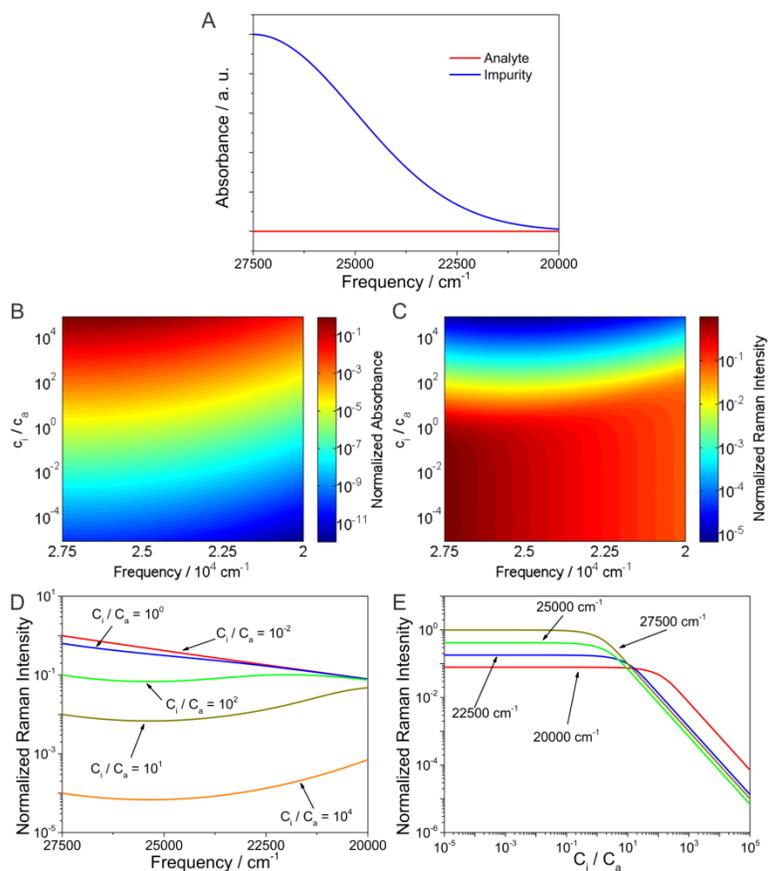


Figure 8. 5. Impact of impurity absorption on observed Raman intensity - case 3

(A) Analyte and impurity absorption spectra. The analyte and impurity absorption bands are identical Gaussians, with $\sigma_a = \sigma_i = 2500 \text{ cm}^{-1}$, and with maxima at $\nu_{ac} = 45,000 \text{ cm}^{-1}$ and $\nu_{ic} = 27,500 \text{ cm}^{-1}$, respectively. (B)

Contour plot of calculated frequency dependence of sample absorbance as a function of relative impurity concentrations. The normalized absorbance is calculated by dividing by the maximum absorbance that occurs at the maximum impurity concentration of $c_i/c_a = 10^5$. (C) Observed contour plot of the frequency dependence of the relative preresonance Raman intensity as a function of the relative impurity concentration calculated from eq. (8.32) The preresonance Raman intensities are normalized to the maximum Raman intensity that occurs for $c_i/c_a = 10^{-5}$.

(D) Frequency dependence of normalized preresonance Raman intensities at different relative impurity concentrations. (E) Relative impurity concentration dependence of observed preresonance Raman intensities.

8.3 CONCLUSIONS

We examined the impact of self absorption on the Raman intensity observed for a 180° backscattering geometry. In the absence of impurities, the observed Raman intensities are predicted to monotonically increase as resonance excitation is approached. In contrast, the observed resonance and preresonance Raman intensities can decrease as excitation approaches resonance if the sample contains absorbing impurities or chromophores that do not enhance the analyte Raman bands. For resonance excitation (defined to occur for excitation within two standard deviations of the analyte Gaussian absorption maximum), a minimum in the observed Raman intensities requires that: 1. the impurity absorption band be at least $\frac{\sqrt{2}}{2}$ narrower than the analyte absorption band; and, 2. that the frequency offset between analyte and impurity absorption maxima be $< 2\sigma_a \left(1 - 2\left(\frac{\sigma_i}{\sigma_a}\right)^2\right)$.

Thus, conditions for a decrease in the observed resonance Raman intensity with excitation closer to resonance will be rare. Spectral S/N maxima will typically occur with resonance excitation. Decreasing observed preresonance Raman intensities and normal Raman intensities can occur in the presence of impurity absorption.

8.4 ACKNOWLEDGEMENT

This work was supported by ONR contract N00014-12-1-0021.

8.5 BIBLIOGRAPHY

- (1) Long, D. A. *The Raman Effect: A unified treatment of the theory of raman scattering by molecules*; John Wiley & Sons Ltd.: West Sussex, England, 2002.
- (2) McCreery, R. L. *Raman Spectroscopy for Chemical Analysis*; John Wiley & Sons Ltd.: New York, 2000.
- (3) Asher, S. A. *Annu. Rev. Phys. Chem.* **1988**, *39*, 537.
- (4) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. *Chem. Rev.* **2012**, *112*, 2604.
- (5) Jones, C. M.; Asher, S. A. *J. Chem. Phys.* **1988**, *89*, 2649.
- (6) Shriver, D. F.; Dunn, J. B. R. *Appl. Spectrosc.* **1974**, *28*, 319.
- (7) Streckas, T. C.; Adams, D. H.; Packer, A.; Spiro, T. G. *Appl. Spectrosc.* **1974**, *28*, 324.
- (8) Liu, C.; Berg, R. W. *Appl. Spectrosc. Rev.* **2012**, *48*, 425.
- (9) Albrecht, A. C. *J. Chem. Phys.* **1961**, *34*, 1476.
- (10) Dudik, J. M.; Johnson, C. R.; Asher, S. A. *J. Chem. Phys.* **1985**, *82*, 1732.
- (11) Tuschel, D. D.; Mikhonin, A. V.; Lemoff, B. E.; Asher, S. A. *Appl. Spectrosc.* **2010**, *64*, 425.
- (12) Ghosh, M.; Wang, L.; Asher, S. A. *Appl. Spectrosc.* **2012**, *66*, 1013.
- (13) Friedman, J.; Hochstrasser, R. M. *Chem. Phys. Lett.* **1975**, *32*, 414.
- (14) Stein, P.; Miskowski, V.; Woodruff, W. H.; Griffin, J. P.; Werner, K. G.; Gaber, B. P.; Spiro, T. G. *J. Chem. Phys.* **1976**, *64*, 2159.

9.0 CONCLUDING SUMMARY

In this work, we discussed the methodology of using UVRR spectroscopy to study the protein secondary structures and their conformational transitions. The conformational transitions between α -helix-like conformations and PPII have been explored using UVRR spectroscopy. The impacts of salt bridge side chain interactions, macrodipole-terminal charge interactions, and surfactants on these transitions have also been investigated.

By studying the AEP peptide, we showed that the macrodipole-terminal charge interactions usually have a greater impact on α -helix stability than the salt bridge side chain interactions do. Besides a single α -helix segment, this short AEP peptide can also form α -helix-turn- α -helix conformation.

The UVRR spectrum of the Arg side chain guanidinium group has been studied. A vibrational band that is sensitive to the guanidinium environment has been identified as the spectral marker that reports on the hydration of guanidinium group. Using this spectral marker, we found that the Arg side chain is exposed to a more hydrophobic environment in the α -helix-like conformations than it is in the PPII conformation. The Arg side chain is less hydrated in AEP than in AP, presumably due to more α -helix-turn- α -helix conformation in AEP.

We also examined the XAO α -helix-like conformations induced by SDS. We demonstrated XAO that adopts predominately PPII conformation in aqueous solutions, shows significant α -helix-like conformations due the formation of XAO-SDS aggregation. XAO

nucleates the SDS aggregation at a concentration below the SDS *cmc*. We proposed a mechanism that four anionic SDS molecules neutralize four cationic XAO side chains, minimizing the electrostatic repulsion. Four XAO-SDS₄ complexes further aggregate. The SDS alkyl chains collapse, inducing XAO α -helix-like conformations that are more compact than the extended PPII conformation. One face of the XAO α -helix-like conformations partitions into the hydrophobic core and the other face is exposed to water. This mechanism may be relevant to some intrinsic disorder peptides (IDPs) and antimicrobial peptides (AMPs) that form α -helices upon association to membranes.

10.0 FUTURE WORK

With UVRR spectroscopy this powerful technique, it can be envisioned that many difficult problems can be tackled, especially those involving peptide conformational change upon association to a target, such as membranes, DNA and metal ions. It is also be important and necessary to further develop the UVRR spectroscopy methodology.

10.1 SEARCHING FOR NEW SPECTRAL MARKER(S) FOR PEPTIDE BACKBONE CONFORMATION

One advantage of UVRR spectroscopy over IR spectroscopy in studying protein secondary structure is that the former provides two conformation sensitive spectral markers, *viz.* the C $_{\alpha}$ -H bending and Am III bands, rather than the Am I band that IR spectroscopy solely relies on. Multiple spectral markers provide more accurate information about the peptide secondary structure. As discussed in Section 3.2.3, the C $_{\alpha}$ -H bending band intensity can be used to distinguish α -helix and PPII conformations, however, β -sheet structures also give rise to a strong C $_{\alpha}$ -H bending band intensity as PPII conformations do. The ψ angle dependent Am III band is insensitive to φ angle.¹ It is useful if we can identify additional UVRR spectral markers that are sensitive secondary structures, especially to φ angle.

It has been observed that an additional band appears at $\sim 1520\text{ cm}^{-1}$ (in some conditions, it partially overlaps with Am II band) in UVRR spectra of highly α -helical peptides. Examples include the spectra of low temperature AEP (Figure 5.2A), XAO with 20 mM SDS (Figure 7.1C) AP in acetonitrile,² polylysine with concentrated NaClO_4 ,³ and Ac-AP-NH₂ (Figure 10. 1).

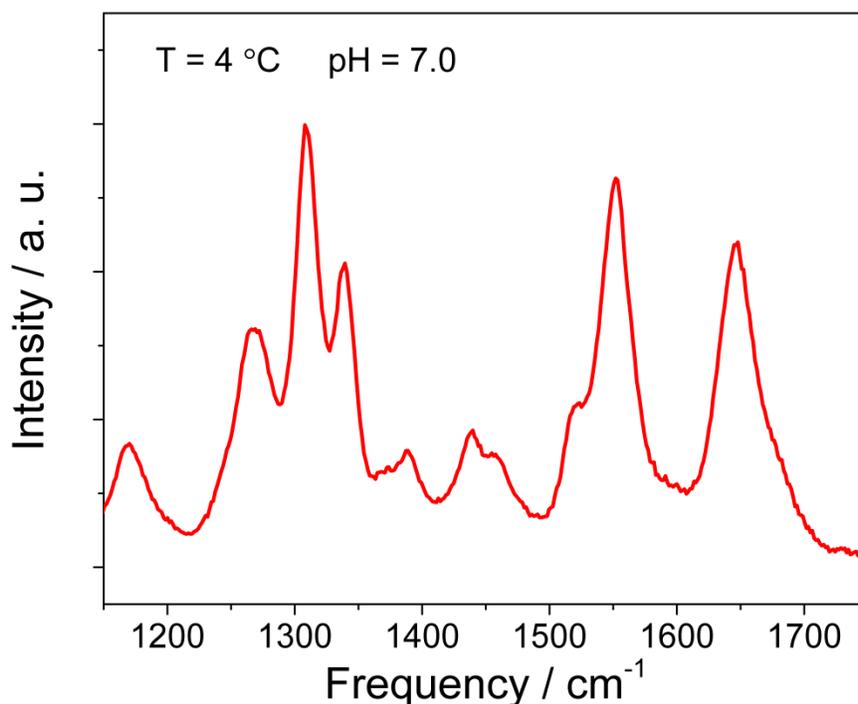


Figure 10. 1. 204 nm excited UVRR spectrum of Ac-AP-NH₂ at 4 °C

The origin of this band is unclear. One possibility is the overtone of amide group out-of-plane bending (opb). *Ab initio* calculation predicts that the NMA second (π - π^*) excited state equilibrium geometry (Figure 10. 2) significantly deviates from a planar structure. This non-planar amide structure indicates that the overtone of amide out-of-plane bending is significantly resonance enhanced, though the fundamental mode is silent due to the symmetry.

NMA gives an amide opb frequency of 650 cm^{-1} , however, DFT calculation on Ac-Ala-NHMe found the amide opb frequencies are in the range of 740 cm^{-1} to 780 cm^{-1} (Figure 10. 3). The overtone frequencies of Ac-Ala-NHMe match the observed $\sim 1520\text{ cm}^{-1}$. Figure 10. 3 shows

that the calculated Ac-Ala-NHMe amide opb frequency depends on both φ and ψ angles. In the negative φ angle range where most secondary structures locate, for a given ψ angle, the frequency shows a sine-like dependence on the φ angle. At a given φ angle, the frequency also shows dependence on ψ angle. It is difficult to separate the impacts of φ and ψ angles because the amide opb involves the distortions of both dihedral angles.

This band involves little N-H opb contribution, it thus depends little on the N-H H-bonding and does not significantly downshift upon N-H deuteration. Indeed, the UVRR spectrum of AP at 0 °C in D₂O does show this band is at a similar frequency,⁴ as expected.

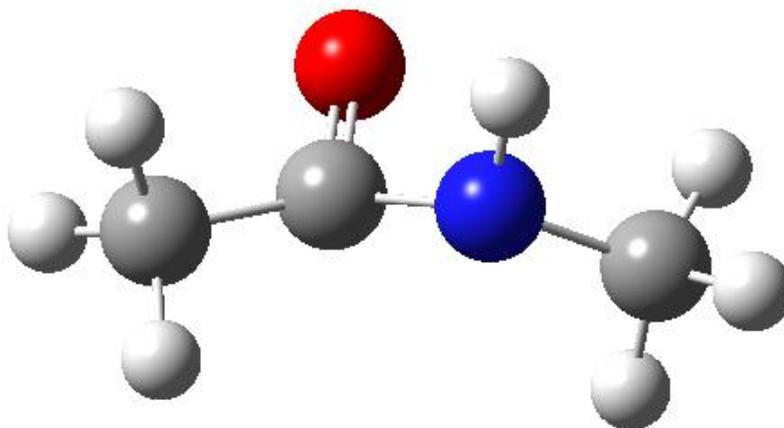


Figure 10. 2. NMA second (π - π^*) excited state equilibrium geometry optimized at CIS / 6-311++G** level of theory with Gaussian 09⁵

Experimentally, by measuring the IR and UVRR spectra of NMA and other model compounds, such as Ac-Ala-NMe, diglycine, etc, we can compare both the fundamental and overtone frequencies to confirm this assignment. Isotopic substitutions (¹³C and/or ¹⁵N and/or ¹⁸O) also help the assignment. Moreover, we can measure different small peptide crystals and find the correlation of the frequency with φ and ψ angles. Thus, we can calculate the ψ angle using the frequency of Am III band, and use the obtained ψ angle as an input to help calculate the φ angle by using the established correlation of this amide opb frequency. This correlation

between amide opb frequency and (φ, ψ) angles will provide an additional spectral marker that allows to calculate φ angle.

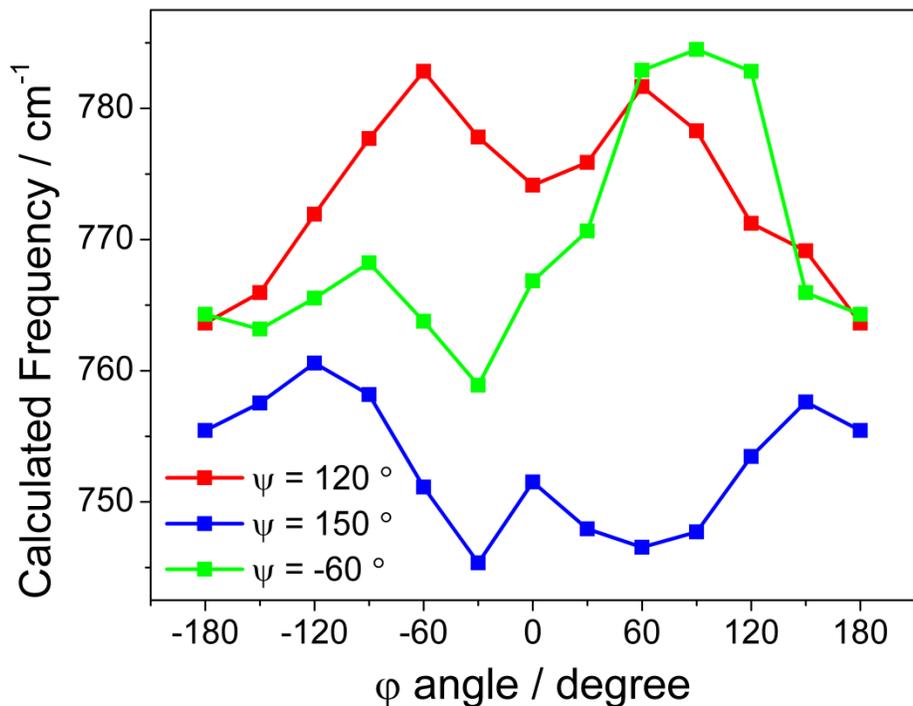


Figure 10. 3. φ angle dependence of Ac-Ala-NHMe amide out-of-plane bending frequency

Calculated in vacuum at B3LYP / 6-311++G** level of theory with Gaussian 09.⁵ The specific (φ, ψ) angles are fixed during the calculation

10.2 INVESTIGATING THE ROLE OF MEMBRANE ON THE AGGREGATION OF INTRINSICALLY DISORDERED PEPTIDES

The α -helices formed by some IDPs upon binding to membranes or membrane mimics, such as SDS micelles, have been proved to facilitate peptide aggregation and fibrillation.⁶⁻⁸ Therefore, membranes and membrane mimics have been proposed to promote the peptide fibrillation.⁹ However, excessive lipid¹⁰ or SDS¹¹ inhibits α -synuclein (α S) fibrillation and

disaggregates fibrils. The role of membrane or membrane mimics in the fibrillation of these IDPs remains unclear.

We observed that SDS induces XAO to aggregate and precipitate upon incubation for days. The precipitation indicates that the aggregates must be much larger than XAO₄-SDS₁₆ that forms in freshly prepared solution. Presumably, these large aggregates are β -structured XAO fibrils. Moreover, the precipitation only occurs at appropriate SDS concentrations. SDS concentrations that are lower than ~ 0.5 mM or higher than ~ 10 mM do not precipitate XAO (1 mM). Figure 9.1 shows three 1 mM XAO solutions with different SDS concentrations. After incubation in room temperature for five days, the XAO solution with 6.0 mM SDS shows significant precipitate. In contrast, other solution, in the absence of SDS or with 20 mM SDS, remains clear without any visible aggregates. These behaviors are similar to the α S-SDS system.

Therefore, the XAO-SDS system is an excellent model to study the role of membranes and membrane mimics in the aggregation and fibrillation of some IDPs, as well as peptide conformational change from α -helix to β -sheet. One hypothesis to be tested is that peptide association to the membranes or membrane mimics neutralizes charges in the peptide and increases the local peptide concentration, enabling peptide molecules to be close enough to each other for some short-ranged inter-molecular interactions,¹² such as inter-molecular H-bonding. These intermolecular interactions may cause conformational change, for example, the intermolecular peptide H-bonding formation breaks the existing intramolecular peptide H-bonds in α -helix, triggering a transition from α -helix to β -sheet. Eventually, these intermolecular interactions lead to peptide aggregation. However, in the presence of excessive membranes or membrane mimics, the local peptide concentration decreases, inhibiting peptide aggregation.

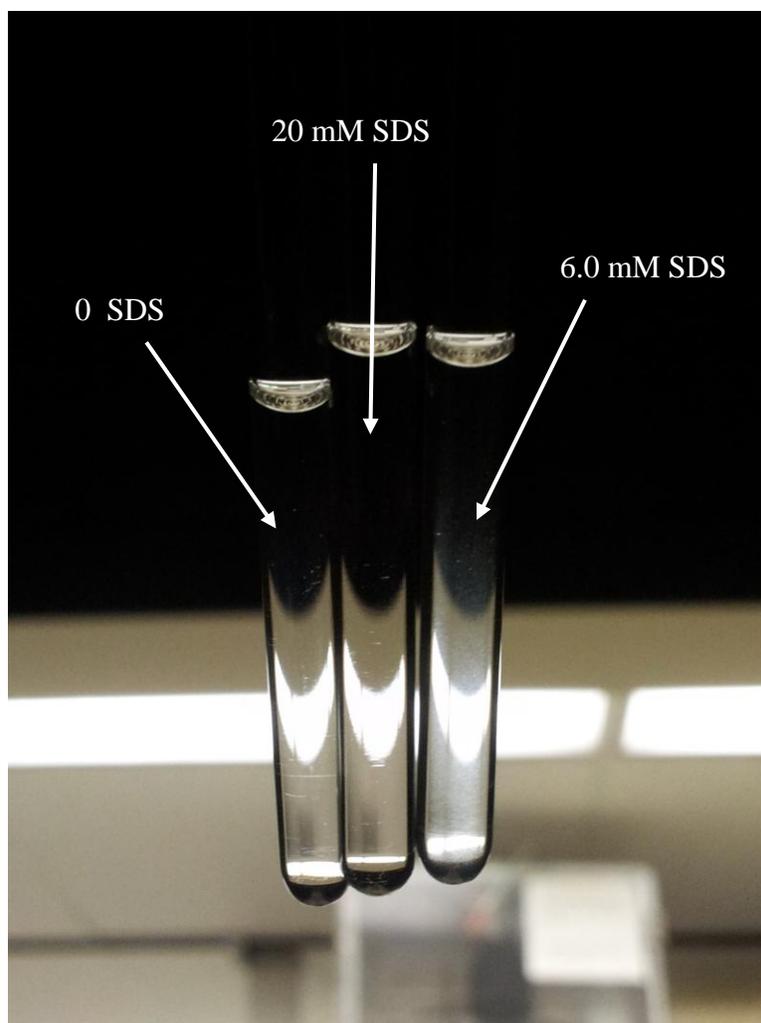


Figure 10. 4. Photograph of the 1 mM XAO in 0 , 6.0 mM and 20 mM SDS, incubated in room temperature for five days.

By monitoring XAO secondary structures change in SDS solution as a function of time, the slow conformational change can be extracted. This will provide information about XAO aggregation and fibrillation at the molecular level. By examining the fibrillation dynamics dependence on SDS concentrations (or $[SDS]/[XAO]$ ratio), the above hypothesis can be tested. This may provide a general picture for some IDPs aggregation upon association with membranes.

We can further expand this work to more biological relevant systems. Membranes are found to play a role⁸ in the aggregation and fibrillation of the Parkinson's disease related peptide,

α S, the Alzheimer's disease related peptide, amyloid β , and the type II diabetes related peptide, IAPP. It is of importance to examine the conformational change of these peptides that are induced by the membranes and the origin of the conformational change. Both the peptide/membrane ratio and the membrane composition affect the conformational change.⁸

10.3 BIBLIOGRAPHY

- (1) Ianoul, A.; Boyden, M. N.; Asher, S. A. *J. Am. Chem. Soc.* **2001**, *123*, 7433.
- (2) Bykov, S., personal communication.
- (3) Ma, L.; Ahmed, Z.; Mikhonin, A. V.; Asher, S. A. *J. Phys. Chem. B* **2007**, *111*, 7675.
- (4) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. *J. Phys. Chem. B* **2004**, *108*, 19020.
- (5) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2010.
- (6) Abedini, A.; Raleigh, D. P. *Protein Eng., Des. Sel.* **2009**, *22*, 453.
- (7) Abedini, A.; Raleigh, D. P. *Phys. Biol.* **2009**, *6*, 015005.
- (8) Butterfield, S. M.; Lashuel, H. A. *Angew. Chem. Int. Ed.* **2010**, *49*, 5628.
- (9) Lee, H.-J.; Choi, C.; Lee, S.-J. *J. Biol. Chem.* **2002**, *277*, 671.
- (10) Zhu, M.; Fink, A. L. *J. Biol. Chem.* **2003**.
- (11) Ahmad, M. F.; Ramakrishna, T.; Raman, B.; Rao, C. M. *J. Mol. Biol.* **2006**, *364*, 1061.
- (12) Byström, R.; Aisenbrey, C.; Borowik, T.; Bokvist, M.; Lindström, F.; Sani, M.-A.; Olofsson, A.; Gröbner, G. *Cell Biochem Biophys* **2008**, *52*, 175.