MOLECULAR STRUCTURAL INSIGHTS OF POLYGLUTAMINE-RICH AMYLOID-LIKE FIBRILS USING UV RESONANCE RAMAN SPECTROSCOPY

by

David Punihaole

B.S. in Molecular Biology, University of Pittsburgh, 2009

Submitted to the Graduate Faculty of the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016
This dissertation was presented

by

David Punihaole

It was defended on

June 24, 2016

and approved by

Dr. Sanford A. Asher, Distinguished Professor of Chemistry, Department of Chemistry

Dr. Linda Jen-Jacobson, Professor, Department of Biological Sciences

Dr. Ronald B. Wetzel, Professor, Department of Structural Biology

Dr. Sunil Saxena, Associate Professor, Department of Chemistry

Dr. Sean Garrett-Roe, Assistant Professor, Department of Chemistry

Dissertation Director: Dr. Sanford A. Asher, Distinguished Professor of Chemistry,

Department of Chemistry
MOLECULAR STRUCTURAL INSIGHTS OF POLYGLUTAMINE-RICH AMYLOID-LIKE FIBRILS USING UV RESONANCE RAMAN SPECTROSCOPY

David Punihaoale, PhD
University of Pittsburgh, 2016

There is currently little that is known about the structure of polyglutamine (polyQ) fibrils, which are involved in at least ten neurodegenerative diseases, including Huntington’s. Given the difficulty of studying these aggregates, new and incisive biophysical methods need to be developed in order to obtain high-resolution structural information of polyQ and other amyloid-like fibrils. Here, we present our recent advances in UV resonance Raman (UVRR) spectroscopy that enable the elucidation of molecular-level structural information of amyloid-like fibrils. We show, for example, how the primary amide UVRR bands report on the local hydrogen bonding and dielectric environment of glutamine side chains. We also discuss a newly discovered spectroscopic marker, the Amide III\(^{P}\) vibration, which sensitively reports on the OCCC dihedral angle of glutamine (Gln) and asparagine (Asn) side chains. These and other spectroscopic markers are used to gain insights into the peptide backbone and side chain conformations of polyQ peptides in solution-state and in fibrils. Finally, we demonstrate how the structural information obtained from UVRR can be utilized to guide Molecular Dynamics simulations in order to obtain experimentally validated structural models of polyQ fibrils.
# TABLE OF CONTENTS

**PREFACE** ........................................... xvii

1.0 POLYGLUTAMINE-RICH PROTEINS AND AGGREGATION ........................... 1

1.1 Clinicopathology of CAG Repeat Diseases ................................. 1

1.2 The Role of PolyQ-rich Aggregates in Pathology ....................... 3

1.2.1 Evidence Supporting the Toxic Role of Aggregates ................. 3

1.2.2 Evidence Supporting the Non-Toxic Role of Aggregates .......... 4

1.2.3 Reconciling Contradictory Results on Aggregate Toxicity ....... 4

1.3 Understanding the Aggregation of PolyQ-rich Peptides and Proteins ... 5

1.3.1 PolyQ Peptides are Disordered in Solution-state .................. 5

1.3.2 The Toxic Monomer Hypothesis ..................................... 6

1.3.3 Proposed Aggregate Structures ...................................... 7

1.3.4 Proposed Aggregation Mechanisms ................................... 9

1.4 The Need for New Biophysical Tools to Study Amyloid-like Fibrils ... 12

2.0 UV RESONANCE RAMAN SPECTROSCOPY ................................... 14

2.1 Classical Electrodynamics Theory of Rayleigh and Raman Scattering ... 14

2.2 Quantum Mechanical Theory of Rayleigh and Raman Scattering ........ 16

2.2.1 The Kramers-Heisenberg Dispersion relation ...................... 17

2.3 Non-Resonance vs. Resonance Scattering ................................ 20

2.3.1 General Vibronic Theory .......................................... 20

2.3.2 Non-Resonance Rayleigh and Raman Scattering .................... 23

2.3.3 Resonance Rayleigh and Raman Scattering ........................ 24

2.3.4 Resonance Raman and the Electronic Excited State Geometry ... 25
3.0 UV RESONANCE RAMAN INVESTIGATION OF THE AQUEOUS SOLVATION DEPENDENCE OF PRIMARY AMIDE VIBRATIONS

3.1 Introduction .......................................................... 29

3.2 Experimental Section ................................................. 30
  3.2.1 Materials .......................................................... 30
  3.2.2 Sample Preparation ............................................... 30
  3.2.3 Infrared and Non-Resonance Raman Spectroscopy ............ 31
  3.2.4 UVRR Spectroscopy ............................................. 31
  3.2.5 UV Absorption Measurements ................................. 32
  3.2.6 Raman Cross Section Calculations ............................ 32

3.3 Computational Section ............................................... 32

3.4 Results and Discussion .............................................. 33
  3.4.1 Normal Mode Analysis .......................................... 33
    3.4.1.1 Propanamide Molecular Structure ........................ 33
    3.4.1.2 2000-3500 cm\(^{-1}\) Region .............................. 38
    3.4.1.3 1500-1800 cm\(^{-1}\) Region .............................. 39
    3.4.1.4 1200-1500 cm\(^{-1}\) Region .............................. 40
    3.4.1.5 C–N Stretching Modes .................................... 40
    3.4.1.6 1000-1200 cm\(^{-1}\) Region .............................. 41
    3.4.1.7 <1000 cm\(^{-1}\) Region .................................... 42
  3.4.2 Solvation Dependence of UVRR Bands ......................... 42
    3.4.2.1 UVRR of Propanamide in Aqueous Solutions ............ 42
    3.4.2.2 Effect of Solvation on UVRR Spectra .................... 43

3.5 Conclusion .......................................................... 48

3.6 Acknowledgment ...................................................... 49

4.0 GLUTAMINE AND ASPARAGINE SIDE CHAIN HYPERCONJUGATION INDUCED STRUCTURALLY SENSITIVE VIBRATIONS

4.1 Introduction .......................................................... 51

4.2 Experimental Details ................................................ 53
  4.2.1 Materials .......................................................... 53
4.2.2 Sample Preparation .................................................. 54
4.2.3 X-ray Diffraction ....................................................... 54
4.2.4 Visible Raman Spectroscopy ........................................ 54
4.2.5 UVRR Spectroscopy ................................................... 55
4.3 Computational Details .................................................. 55
4.3.1 Density Functional Theory (DFT) Calculations ................. 55
4.4 Results and Discussion ................................................. 56
4.4.1 Assignment of L-Gln UVRR Bands in H$_2$O and D$_2$O ......... 56
4.4.2 Conformational Dependence of the Am III$^P$ Band ............. 61
4.4.3 Origin of the OCCC Dihedral Angle Dependence of the Am III$^P$ Vibration .................................................. 63
4.4.4 Experimental Dependence of Am III$^P$ Band Frequency on OCCC Dihedral Angle .................................................. 66
4.4.4.1 Dependence of Am III$^P$ Band Frequency in Crystals ..... 66
4.4.4.2 Dependence of Am III$^P$ Band Frequency for Fully Hydrated Primary Amides .................................................. 67
4.4.4.3 Dependence of Am III$^P$ Band Frequency for Low Dielectric Constant and Weak Hydrogen Bonding Environments ..... 67
4.4.4.4 Dependence of Am III$^P$ Band Frequency for Unknown Dielectric and Hydrogen Bonding Environments .......... 68
4.4.5 Predicting Side Chain $\chi_3$ and $\chi_2$ Dihedral Angles in Gln and Asn as a Function of Ramachandran ($\Phi$, $\Psi$) Angles ......................... 68
4.4.6 Experimentally Determined Gln PPII-like Structure Peptide $\chi_3$ Dihedral Angles .................................................. 72
4.4.6.1 UVRR Spectra of Gln Peptides in PPII-like Structures .... 72
4.4.6.2 $\chi_3$ Dihedral Angle Determination of Side Chains in Gln Peptides .................................................. 74
4.4.6.3 Determination of the Gibbs Free Energy Landscape for Gln and Gln peptides Along the $\chi_3$ Dihedral Angle Reaction Coordinate .................................................. 77
## 5.0 Polyglutamine Fibrils: New Insights into Antiparallel β-Sheet Conformational Preference and Side Chain Structure

### 5.1 Introduction

### 5.2 Experimental Section

| 5.2.1 Materials | 86 |
| 5.2.2 Sample Preparation | 86 |
| 5.2.3 Transmission Electron Microscopy (TEM) | 87 |
| 5.2.4 X-ray Diffraction of Fibril Films | 87 |
| 5.2.5 UVRR spectroscopy | 88 |

### 5.3 Computational Section

### 5.4 Results and Discussion

| 5.4.1 Q10 Forms Amyloid-like Fibril Aggregates | 90 |
| 5.4.2 UVRR of Polyglutamine Fibrils in H₂O | 92 |
| 5.4.2.1 Assignment of Gln Side Chain Bands | 92 |
| 5.4.2.2 Assignment of Peptide Backbone Bands | 94 |
| 5.4.3 UVRR of Polyglutamine Fibrils in D₂O | 96 |
| 5.4.4 Hydrogen-Deuterium (HX) Exchange of Polyglutamine Fibrils | 97 |
| 5.4.4.1 UVRR Bands of Partially Deuterated Primary Amides | 97 |
| 5.4.4.2 HX of NDQ10 and DQ10 fibrils | 98 |
| 5.4.5 Ramachandran Ψ Angle Distributions | 100 |
| 5.4.6 Developing a Molecular-level Structural Model | 103 |
| 5.4.7 Antiparallel β-sheets are more Favorable than Parallel β-sheets in polyQ fibrils | 106 |
| 5.4.8 Structure of NDQ10 and DQ10 Gln side chains | 107 |
| 5.4.8.1 Determination of Gln Side Chain χ₃ Angle Distributions | 107 |
| 5.4.8.2 Comparisons of Side Chain Structures with Other Models | 110 |
5.4.9 Dependence of Fibril Structure on Deposition of Different NDQ10 and DQ10 Solution Conformations .................................................. 111
5.5 Conclusion .................................................................................. 111
5.6 Note Added in Revision ................................................................. 113
5.7 Acknowledgments ........................................................................ 114

APPENDIX A. DERIVATIONS ............................................................... 117
A.1 Derivation of the Transition Scattering Rate .................................. 117
A.1.1 Second Order Perturbation Theory ........................................... 117
A.1.2 Fermi’s Golden Rule and the Transition Scattering Rate ............ 121
A.2 Derivation of the Albrecht $A$, $B$, and $C$ Terms .......................... 123

APPENDIX B. SUPPORTING INFORMATION FOR CHAPTER 3.0 . 126
B.1 X-ray Crystal Structure of Propanamide ...................................... 126
B.2 Deconvolution of UVRR Spectra ................................................ 127

APPENDIX C. SUPPORTING INFORMATION FOR CHAPTER 4.0 . 133
C.1 Dependence of the AmIII$^P$ Band Frequency on the OCCC Dihedral Angle of Butyramide .............................................................. 133
C.2 X-ray Diffraction of Glutamine and Derivatives ............................ 133
C.3 Raman Band Assignments of Crystalline Glutamine and Derivatives . 135
C.3.1 Spectral Deconvolution ............................................................ 135
C.3.2 Band Assignments ................................................................. 135
  C.3.2.1 L-glutamine and D-glutamine ............................................. 135
  C.3.2.2 N-Acetyl-L-glutamine ....................................................... 136
  C.3.2.3 L-glutamine t-butyl ester HCl .............................................. 136
  C.3.2.4 Glycl-L-glutamine ............................................................ 137
  C.3.2.5 L-seryl-L-asparagine ....................................................... 137
C.4 Dependence of the AmIII$^P$ Frequency on Hydrogen Bonding and Dielectric Environment .......................................................... 137
C.5 Circular Dichroism (CD) of the Gln$_3$ peptide ................................ 138

APPENDIX D. SUPPORTING INFORMATION FOR CHAPTER 5.0 . 157
D.1 Experimental Section ................................................................. 157
D.1.1 UVRR spectroscopy of fibril films ........................................ 157
D.1.2 UVRR Spectral Processing .................................................. 157
D.1.3 UVRR Spectral Peak Fitting .............................................. 158
D.1.4 Calculation of the $\Psi$ and $\chi_3$ Angle Distributions ............ 159
    D.1.4.1 Correlating the AmIII$^3$ Frequencies to $\Psi$ Angles ........ 159
    D.1.4.2 Correlating the AmIII$^p$ Frequencies to $\chi_3$ Angles ........ 160

D.2 Computational Section ...................................................... 160
    D.2.1 Density Functional Theory (DFT) Calculations .................. 160
    D.2.2 RMSD Metric ........................................................... 161

D.3 Results and Discussion ..................................................... 162
    D.3.1 UVRR of NDQ10 and DQ10 Fibril Films ............................. 162
    D.3.2 Band Assignments of Mono-deuterated Primary Amides .......... 162
        D.3.2.1 Assignment of Amide Vibrations ............................. 163
    D.3.3 Bennett Acceptance Ratio Method .................................. 164
    D.3.4 Hydrogen Bonding Analysis ....................................... 164

REFERENCES ............................................................................ 174
# LIST OF TABLES

1.1 Summary of Trinucleotide CAG Repeat Diseases ........................................ 2  
3.1 Frequencies of Crystalline CH$_3$CH$_2$CONH$_2$ and CH$_3$CH$_2$COND$_2$ ........... 35  
3.2 Frequencies and Assignments of Infrared and Raman Bands for Crystalline CH$_3$CH$_2$CONH$_2$ ........................................................................ 36  
3.3 Frequencies and Assignments of Infrared and Raman Bands for Crystalline CH$_3$CH$_2$COND$_2$ ........................................................................ 37  
3.4 Measured Frequencies and Cross Sections of UVRR bands in CH$_3$CH$_2$CONH$_2$ 43  
4.1 UVRR Frequencies and Assignments of L-glutamine in H$_2$O ......................... 59  
4.2 UVRR Frequencies and Assignments of L-glutamine in D$_2$O ......................... 60  
4.3 Predicted Amide III$^P$ Frequencies and OCCC Dihedral Angles for Glutamine and Asparagine Residues with Different Ramachandran Angles ................. 71  
B1 Summary of Crystallographic Data for CH$_3$CH$_2$CONH$_2$ ............................ 127  
B2 Atomic Coordinates and Equivalent Isotropic Displacement Parameters .......... 128  
B3 Anisotropic Atomic Displacement Parameters ............................................... 128  
B4 Comparison of Bond Lengths Between the DFT-optimized and X-ray Crystal Structure ................................................................. 128  
B5 Comparison of Torsion Angles Between the DFT-optimized and X-ray Crystal Structure ................................................................. 129  
B6 Comparison of Bond Angles Between the DFT-optimized and X-ray Crystal Structure ................................................................. 129  
C1 Summary of Crystallographic Data of Glutamine and Derivatives .................... 139  
C2 Summary of Crystallographic Data for L-glutamine t-butyl ester HCl ................ 140
C3 Data Collection and Structure Refinement for L-glutamine t-butyl ester HCl
C4 Atomic Coordinates and Equivalent Isotropic Atomic Displacement Parameters for L-glutamine t-butyl ester HCl
C5 Anisotropic Atomic Displacement Parameters for L-glutamine t-butyl ester HCl
C6 Hydrogen Atomic Coordinates and Isotropic Atomic Displacement Parameters for L-glutamine t-butyl ester HCl
C7 Bond Lengths and Angles for L-glutamine t-butyl ester HCl
C8 Raman Frequencies and Assignments of Crystalline L-glutamine
C9 Raman Frequencies and Assignments of Crystalline D-glutamine
C10 Raman Frequencies and Assignments of Crystalline N-acetyl-L-glutamine
C11 Raman Frequencies and Assignments of Crystalline L-glutamine t-butyl ester HCl
C12 Raman Frequencies and Assignments of Crystalline L-glycyl-L-glutamine
C13 Raman Frequencies and Assignments of Crystalline L-seryl-L-asparagine
C14 UVRR Frequencies and Assignments of Butyramide in Acetonitrile and Water
D1 DFT Calculated Frequencies and Assignments of cis-Glutamine-N$_2$HD
D2 DFT Calculated Frequencies and Assignments of trans-Glutamine-N$_2$HD
D3 UVRR Band Frequencies and Assignments of trans- and cis-Glutamine-N$_2$HD
D4 Average Number of Hydrogen Bonds for Antiparallel and Parallel $\beta$-sheet Fibril Models
LIST OF FIGURES

1.1 Proposed Structures for PolyQ Fibrils ........................................ 8
1.2 Proposed Aggregation Mechanisms for PolyQ Fibril Formation ............ 10
2.1 Classical Physics Picture of Rayleigh and Raman Scattering ............... 15
2.2 Selectivity Advantage of Resonance Raman Spectroscopy .................. 16
2.3 Energy Diagrams Showing Various Two Photon Optical Processes .......... 17
3.1 Structure of Propanamide Showing Atomic Numbering Scheme Used for Normal Mode Analysis .................................................. 34
3.2 IR, Non-resonance Raman, and UVRR spectra of Crystalline Propanamide . 38
3.3 UVRR Spectra of Propanamide in H₂O and D₂O ............................... 42
3.4 UVRR Spectra of Propanamide in Different Acetonitrile and Water Mixtures 45
3.5 Dependence of UVRR Propanamide Band Cross Sections and Frequencies on the Mole Fraction of Water ........................................ 46
3.6 UV Absorption Spectra of NMA and Propanamide in Acetonitrile and Water 47
4.1 Geometry of Optimized Structure and Atomic Labeling Scheme of L-glutamine Used in DFT Calculations and Band Assignments ............... 56
4.2 UVRR Spectra Excited at 204 nm of L-glutamine in H₂O and D₂O .......... 58
4.3 Calculated Amide IIÎ P Frequency and Bond Length Dependence on the χ₃ Dihedral Angle of the Glutamine Side Chain ............................ 62
4.4 Hyperconjugation Results in the Cβ–Cγ Bond Length Sensitivity to the χ₃ Dihedral Angle ................................................................. 64
4.5 NBO Charge of Cβ in L-glutamine as a Function of the χ₃ Dihedral Angle .. 65
4.6 Experimental Correlation of the Amide IIÎ P Frequency to the χ₃ Dihedral Angle 66
4.7 Glutamine and Asparagine Side Chain $\chi_3$ and $\chi_2$ Dihedral Angle Dependence on Secondary Structure ................................................................. 69
4.8 Deconvolution of the UVRR Spectra of $Q_3$ and $D_2Q_{10}K_2$ ................................. 73
4.9 $\chi_3$ Dihedral Angle Histograms Calculated by Decomposing Amide III P Bands into a Sum of Lorentzians for Gln, $Q_3$, and $D_2Q_{10}K_2$ in Water ............. 76
4.10 Comparison of $\chi_3$ Dihedral Angle Distributions Between Gln, $Q_3$ in a Predominately PPII-like Conformation, and $D_2Q_{10}K_2$ in a PPII/2.51-helix Equilibrium 78
4.11 Gibbs Free Energy Landscapes of Gln, $Q_3$, and $D_2Q_{10}K_2$ ...................... 79
5.1 MD Simulated Structures of Model Q10 Fibril Systems ................................. 85
5.2 Transmission Electron Micrographs and X-ray Diffraction Patterns of NDQ10 and DQ10 Fibril Aggregates ......................................................... 91
5.3 UVRR Spectra of NDQ10 and DQ10 Fibrils Prepared in H$_2$O .................. 93
5.4 UVRR Spectra of NDQ10 and DQ10 Fibrils Prepared in D$_2$O .................. 96
5.5 UVRR Spectrum of Glutamine Measured in a 50%H$_2$O/50%D$_2$O Mixture .. 97
5.6 HX-UVRR Spectra of NDQ10 and DQ10 Fibrils ........................................ 99
5.7 UVRR Determined $\Psi$ Angle Distributions for NDQ10 and DQ10 Fibrils .... 102
5.8 Time Evolution of Fibril Models in MD Simulations ...................... 103
5.9 Comparison of $\Psi$ Angle Distributions from Simulations and Experiments ... 105
5.10 Peak Fitting of the 197 nm – 204 nm UVRR Difference Spectra of NDQ10 and DQ10 Fibrils ................................................................. 108
5.11 $\chi_3$ Dihedral Angle Distributions of Glutamine Side Chains .................... 115
5.12 Comparison of Side Chain Geometries ................................................. 116
B1 Crystal Structure of Propanamide ...................................................... 130
B2 Spectral Deconvolution of the 204 nm UVRR Spectrum of Propanamide in Water and Acetonitrile in the Region from 1200–1800 cm$^{-1}$ .......................... 131
B3 Spectral Deconvolution of the 204 nm UVRR Spectrum of Propanamide in Water and Acetonitrile in the Region from 800–1200 cm$^{-1}$ .......................... 132
C1 Amide III P Frequency Dependence on the OCCC Dihedral Angle of Butyramide from DFT Calculations ...................................................... 150
C2 Crystal Structures of Glutamine and Glutamine Derivatives Examined .... 151

xiv
C3 OTREP Diagram of GlnTBE with Atomic Labeling Scheme Used in Table C4–Table C7 ........................................... 152
C4 Visible Raman Spectra of Compounds Examined ..................................................... 153
C5 UVRR spectra of compounds examined ................................................................. 154
C6 UVRR Spectra of Buytramide in Water and Acetonitrile ........................................... 155
C7 Temperature Dependent Circular Dichroism Spectra of Q₃ ........................................ 156
D1 Atomic Labeling Scheme for Glutamine used in DFT Calculations ............................... 169
D2 Plot of the Evolution of Backbone RMSD Metric for Three Fibril Models with Respect to their Initial Structure ................................................................. 170
D3 UVRR Difference Spectra of Fibril Films Prepared from NDQ10 and DQ10 ............... 171
D4 Potential Energy distributions for Simulated Fibril Models ....................................... 172
D5 Histograms of the Number of Hydrogen Bonds for Fibril Models .............................. 173
LIST OF SCHEMES

3.1 Resonance Structures of Propanamide ........................................... 47

5.1 Geometric Isomers of the Mono-deuterated Primary Amide Group ....... 98

5.2 Ramachandran $\Psi$ and $\Phi$ Dihedral Angles in Polyglutamine Peptides .. 101
I would like to thank my advisor, Prof. Sanford Asher, for his mentorship and inspiration during my graduate school years. Most PhD students are “handed” a project in graduate school. Under Sandy’s guidance, however, I was allowed to craft my own research program. I am extremely grateful for this rare opportunity, as it afforded me not only complete ownership of my work, but it also rigorously trained me to think critically as a scientist. I deeply admire and respect Sandy’s prowess as a scientist and his infectious enthusiasm towards research. I could not have worked for a better PI.

Over the years, I have collaborated extensively with several people. I would like to thank Dr. Zhenmin Hong and Prof. Nataliya Myshakina for performing DFT calculations. I would also like to thank Riley Workman and Prof. Jeffry Madura for conducting Molecular Dynamics simulations. I am also grateful to Ryan Jakubek and Elizabeth Dahlburg for their assistance in conducting experiments and critically discussing the data.

My lab mates and colleagues in the Department of Chemistry are a source of inspiration, friendship, and support. I am especially grateful to Jonathan Wert, Yangguang Ou, and Chino Cabalteja. I would also like to thank Sharon Mansfield, our group’s administrative assistant, for all her kindness, help, and support.

I would like to acknowledge my thesis committee, particularly Prof. Linda Jen-Jacobson. As an undergraduate student, I took a course taught by Linda, which inspired me to switch career paths and to pursue science. Her course instilled in me a deep appreciation for applying biophysical methods to gain fundamental insights into biological phenomena. It was probably the single most important and influential course that I took at the University of Pittsburgh.

Finally, I would like to express gratitude to my close friends and family, particularly
my mom and grandma. Their hard work, support, and sacrifices over the years enabled me
to become, not only the first person in my family to earn a bachelor’s degree, but also a
doctorate.
1.0 POLYGLUTAMINE-RICH PROTEINS AND AGGREGATION

1.1 CLINICOPATHOLOGY OF CAG REPEAT DISEASES

The abnormal expansion of DNA trinucleotide repeats that encode homopolymeric tracts of amino acids in proteins has been implicated in at least 16 diseases [1]. One of the most prevalent and devastating classes of trinucleotide diseases involve genomic expansions of CAG repeats that encode for polyQ tracts in proteins. At present, expansions in CAG repeats have been linked to at least 10 diseases [1–4] (Table 1.1), including several sinocerebellar ataxis and spinobulbar muscular atrophy. However, the most notable and prevalent CAG repeat disorder is Huntington’s disease.

The predominant clinical symptoms of CAG repeat diseases are neurodegenerative in nature (Table 1.1). The characteristic symptoms of Huntington’s disease, for example, are involuntary, jerky movements (chorea), cognitive impairments, mood swings, and behavioral changes that progressively worsen over time. In the case of spinobulbar muscular atrophy, additional non-neurological symptoms also manifest, including gynecomastia and sterility.

A common feature of these diseases is that the severity of the clinicopathological symptoms correlates with the length of the polyQ repeat expansion. Generally speaking, the larger the repeat length, the greater the disease morbidity and mortality rate. There is strong epidemiological evidence that indicates successive generations of families afflicted by CAG repeat diseases experience an earlier age of onset of disease symptoms [1, 5]. In all of these diseases, symptoms manifest when affected proteins possess polyQ tracts that exceed a threshold repeat length. For example, in Huntington’s disease [4, 6], repeats between 17–30 Gln residues in the huntingtin are generally considered benign, whereas repeat lengths that exceed 36 Gln residues typically result in disease symptoms.
Table 1.1: Summary of Trinucleotide CAG Repeat Diseases\textsuperscript{a}

<table>
<thead>
<tr>
<th>Disease\textsuperscript{b}</th>
<th>Gene Product</th>
<th>Normal Repeat Length</th>
<th>Pathological Repeat Length</th>
<th>Clinical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>Huntingtin</td>
<td>6–34</td>
<td>36–121</td>
<td>Chorea, dystonia, cognitive impairment, depression, psychiatric problems</td>
</tr>
<tr>
<td>SCA1</td>
<td>Ataxin1</td>
<td>6–44</td>
<td>39–82</td>
<td>Ataxia, cognitive impairments, slurred speech</td>
</tr>
<tr>
<td>SCA2</td>
<td>Ataxin2</td>
<td>15–24</td>
<td>32–200</td>
<td>Ataxia, decreased reflexes, infant variant with retinopathy</td>
</tr>
<tr>
<td>SCA3</td>
<td>Ataxin3</td>
<td>13–36</td>
<td>61–84</td>
<td>Ataxia, parkinsonism</td>
</tr>
<tr>
<td>SCA6</td>
<td>CACNA1\textsubscript{A}</td>
<td>10–33</td>
<td>36–121</td>
<td>Ataxia, dysarthria, tremors</td>
</tr>
<tr>
<td>SCA7</td>
<td>Ataxin7</td>
<td>4–35</td>
<td>37–306</td>
<td>Ataxia, blindness, infant variant with cardiac failure</td>
</tr>
<tr>
<td>SCA17</td>
<td>TBP</td>
<td>25–42</td>
<td>47–63</td>
<td>Ataxia, cognitive decline, seizures, psychiatric problems</td>
</tr>
<tr>
<td>SBMA</td>
<td>Androgen receptor</td>
<td>38–62</td>
<td>36–121</td>
<td>Gynecomastia, decreased fertility, motor weakness</td>
</tr>
<tr>
<td>DRPLA</td>
<td>Atrophin</td>
<td>7–34</td>
<td>49–88</td>
<td>Ataxia, seizures, dementia, choreoathetosis</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Table adapted from references [1, 3]. \textsuperscript{b}HD: Huntington’s disease; SCA: Spinocerebellar ataxia; SBMA: Spinobulbar muscular atrophy (Kennedy disease); DRPLA: Dentatorubral-pallidoluysian atrophy.
1.2 THE ROLE OF POLYQ-RICH AGGREGATES IN PATHOLOGY

The pathological hallmarks of all CAG repeat diseases is the accumulation of insoluble polyQ-rich protein aggregates in cells [7]. An important aspect in the pathophysiology of CAG repeat diseases is understanding the role that polyQ-rich aggregates play in neurotoxicity. Although these aggregates are pathological hallmarks, the question remains: Are they the toxic agents or merely by-products of the diseases? Numerous studies have addressed this question in recent years; however, there is little consensus, and the emerging picture is complex. Some studies link polyQ-rich aggregates with cellular toxicity and neurodegeneration, while others suggest that they potentially play a neuroprotective role.

1.2.1 Evidence Supporting the Toxic Role of Aggregates

Evidence that supports the toxic aggregate hypothesis stems from several observations. Studies on Huntington’s disease, for example, show that transcription factors [8–12], molecular chaperones [13], and proteasomal proteins [14] colocalize with nuclear and cytoplasmic inclusions. These results suggest that polyQ aggregates may disrupt normal homeostasis by sequestering transcription factors and other proteins, thereby depleting their cellular concentrations and disrupting their normal functions [15, 16].

More direct evidence supporting aggregate toxicity derives from cell-based assays. Exogenously introduced polyQ aggregates are highly cytotoxic when localized to the nucleus [17]. In addition, expression of atrophin-1 with long tracts of polyQ repeats (129 Gln residues) in mice results in significantly more neuronal inclusions and greater brain atrophy than protein constructs with smaller repeat lengths [18]. These results mirror some postmortem studies [19, 20] of presymptomatic Huntington’s patients, where the formation of polyQ-rich aggregates correlate to morphological changes in brain tissue.

The cytotoxic effects of nuclear inclusions can be mitigated by rescuing cells with inhibitors and molecular chaperones that can prevent aggregation. For example, Cummings et al. [21] showed that the overexpression of molecular chaperone proteins significantly decreases the accumulation of ataxin-1 aggregates. The most compelling evidence, however,
stems from Thakur et al. [22], who showed that treating cells with small proline-containing peptides that inhibit aggregation results in significantly reduced cell death.

1.2.2 Evidence Supporting the Non-Toxic Role of Aggregates

Despite mounting evidence, there are several studies that have strongly questioned whether polyQ-rich aggregates are the toxic agents. An early study that questioned the toxic aggregate hypothesis was published by Greenberg and coworkers [23]. Although their results showed that preventing huntingtin aggregation in the nucleus stopped apoptosis, they did not find a strong correlation between the formation of nuclear inclusions and cell death. Greenberg and coworkers noted, for example, that the formation of huntingtin aggregates accompanied neuronal cell death in the striatum, but not the hippocampus. Their findings have been corroborated by more recent studies [24, 25] that also fail to find a robust link between polyQ aggregates and cell death or neurodegenerative disease symptoms.

Other studies suggest that polyQ-rich aggregates may actually serve a neuroprotective role! Evidence supporting this hypothesis derives primarily from mice studies. For example, Cummings et al. [26] showed that reducing ataxin-1 aggregation actually correlated to an increase in the disease pathology spinocerebellar ataxia-1. Their findings are supported by several more recent studies [27–29], which show that the most apoptotically vulnerable neuronal cells do not show the formation of aggregate inclusions until the most advanced stages of disease. In light of these studies, it has been proposed [1] that monomeric (and not aggregated) polyQ-rich proteins are the cytotoxic agents (vide infra).

1.2.3 Reconciling Contradictory Results on Aggregate Toxicity

It is difficult to reconcile the seemingly contradictory observations regarding the cellular toxicity of polyQ aggregates. There are many potential reasons for the conflicting reports, including significant differences in experimental designs and controls. It is conceivable, for instance, that the different cell lines and animal models used by various research groups exhibit differential susceptibilities to polyQ aggregate toxicity. In addition, the structural nature and location of polyQ aggregates may also be important. For example, aggregates
located in the nucleus are significantly more cytotoxic than those located in the cytoplasm [17]. Similarly, some studies even suggest that various aggregate polymorphs exhibit different cytotoxic effects on cells [30]. Finally, it may also be the case that the pathophysiologies of different CAG repeat diseases vary so that in some disorders aggregates play a more cytotoxic role. However, it is difficult to imagine that the formation of large neuronal polyQ-rich aggregates do not at least play some significant role in neuronal cell death.

1.3 UNDERSTANDING THE AGGREGATION OF POLYQ-RICH PEPTIDES AND PROTEINS

Understanding the underlying aggregation mechanisms of polyQ peptides and proteins is important in developing fundamental insights into the etiologies of CAG repeat diseases, as well as formulating potential therapies. To this end, biophysical studies have focused largely on understanding the structures of polyQ-rich peptides in both solution-state and in aggregates. Given the only apparent commonality of CAG repeat diseases are the polyQ repeats, most studies have focused on investigating synthetic homopolymeric model peptide systems, with sequences such as D2QnK2 or K2QnK2 (where n is the number of Gln repeats).

1.3.1 PolyQ Peptides are Disordered in Solution-state

Most biophysical studies indicate that polyQ peptides are structurally disordered in solution-state [31–35]. One of the most surprising findings is that there is no obvious difference in the solution-state secondary structures between “pathologic” and “non-pathologic” polyQ repeat lengths [34]. These findings are also supported by computational studies [36–39], which show that polyQ peptides, regardless of repeat length, are largely disordered, with only transient elements of regular secondary structures such as α-helices, β-sheets, and turns.

Although polyQ peptides are structurally disordered, they do not behave as true random coil polymers [4]. Several experimental and computational studies [36, 37, 40–44] indicate
that the end-to-end distances of polyQ peptides deviate significantly from random coils, since they adopt relatively collapsed, globule-like structures. In addition, a detailed analysis of circular dichroism (CD) and nuclear magnetic resonance (NMR) spectra [33] shows that polyQ peptides show a high propensity towards adopting residual polyproline II-like (PPII-like) secondary structures, although not necessarily in long, continuous tracts.

These studies indicate that water is a poor solvent for polyQ peptides that contain long repeat lengths [43], despite the apparent hydrophilic nature of the Gln side chains. The collapsed, globule-like structures adopted by polyQ peptides are the result of the extensive hydrogen bonding interactions that can occur between the secondary amides of the peptide backbone and the primary amides of the Gln side chains. These inter-amide interactions are presumably stronger than amide-water hydrogen bonding interactions.

1.3.2 The Toxic Monomer Hypothesis

Although polyQ monomers are predominately disordered in solution-state, there are some studies that suggest these peptides adopt low concentrations of non-disordered conformations. These non-disordered structures are believed to be the putative cytotoxic agents that contribute to neurodegeneration [35]. According to this hypothesis, these non-disordered states increase significantly in concentration for pathological polyQ repeat lengths (i.e. 35–45 Gln residues in the case of Huntington’s disease). The appeal of this so-called “toxic monomer hypothesis” is that it provides an explanation for the pathological gain-of-function that arises in mutated proteins with expanded polyQ repeat domains, while also accounting for observations that polyQ aggregates do not robustly correlate with neuronal cell death.

The supporting data for this hypothesis derives primarily from antibody binding studies. Trottier et al. [45], for example, observed with Western blot analysis that monoclonal anti-polyQ antibodies have a higher binding affinity for proteins with longer polyQ repeat domains than shorter, non-pathologically relevant repeat lengths. These data led to the speculation that the antibodies preferentially recognize stretches of polyQ repeats that adopt specific secondary structures. Similar results have been observed more recently with a different monoclonal antibody, called 3B5H10 [46]. X-ray crystallographic and small-angle X-ray
scattering data [47] suggests that 3B5H10 antibodies recognize \( \beta \)-hairpins that are formed in polyQ peptides with pathologically long repeat lengths.

Although it is difficult to unequivocally rule out the existence of low concentrations of polyQ monomers that adopt stable, ordered structures, there are alternative explanations that account for the results of these antibody binding studies. One explanation, for example, is that antibodies can trap kinetically accessible polyQ conformations that are not normally populated. Another explanation is the so-called “linear lattice” effect [48, 49], where the preferential binding is due to there being a significant increase in the number of epitopes in polyQ peptides of pathological repeat lengths compared to non-pathological repeat lengths.

1.3.3 Proposed Aggregate Structures

PolyQ aggregates prepared from synthetic peptides in vitro bear many of the hallmark features of amyloid-like fibrils, including displaying filamentous morphologies, exhibiting \( \beta \)-sheet-rich structures, and binding Thioflavin-T [32, 35]. Despite this, polyQ aggregates do not exhibit all the characteristics of amyloids as exemplified, for example, by A\( \beta \) fibrils. Some amyloid features that polyQ aggregates do not display include exhibiting birefringence upon Congo Red binding and a classical cross-\( \beta \) X-ray fiber diffraction pattern [4, 32, 35, 50]. There are currently no atomic-resolution structures of polyQ amyloid-like fibril aggregates. However, numerous structures (Figure 1.1) have been proposed for polyQ fibrils on the basis of a variety of classical biophysical methods.

The first detailed structural model of polyQ fibrils was proposed by Perutz and coworkers [52], who utilized X-ray fiber diffraction. Their results showed that \( \text{D}_2\text{Q}_{15}\text{K}_2 \) aggregates produce diffraction patterns with prominent reflections located at 4.2 Å, 4.8 Å, and 8.4 Å. They attributed the diffraction pattern to a cross-\( \beta \) structure and assigned the 4.8 Å reflection to the inter-strand distance of the fibril \( \beta \)-sheets. However, the origin of the 8.4 Å reflection was mysterious since typical amyloids show inter-\( \beta \)-sheet distances between 10–12 Å. As a result, Perutz et al. proposed that polyQ fibril aggregates were composed of “polar-zipper” structures (Figure 1.1a), where the neighboring \( \beta \)-sheets are stacked \( \sim \)17 Å apart and are hydrogen bonded to each other via side chain amides.
The lack of the characteristic 10–12 Å reflection in polyQ fibril diffraction patterns compelled Perutz and coworkers to reanalyze their data in a subsequent study [53]. In their revised model, they proposed that polyQ fibrils consist of water-filled nanotubes that are formed by cylindrical $\beta$-helices (Figure 1.1b). However, other studies provide alternative interpretations of the polyQ diffraction data collected by Perutz et al. [52, 53]. For example, Sikorski and Atkins [54] analyzed the original X-ray data of Perutz [52] and argued that the $\beta$-helical model was incorrect. Instead, they proposed that polyQ fibrils are composed of a cross-$\beta$ structure, but, because of extensive side chain inter-amide hydrogen bonding interactions, the $\beta$-sheets are stacked closer together ($\sim$8–9 Å) than typical amyloids ($\sim$10–12 Å).

The Sikorski and Atkins interpretation of the Perutz data have been further substantiated by Sharma et al.’s [55] work, who also propose that polyQ fibrils are composed of cross-$\beta$ structures. The Sharma et al. model posits that polyQ fibril aggregates are composed of antiparallel $\beta$-sheets that contain reverse hairpin turns. Their model, however, differs significantly from the Sikorski and Atkins [55] structure. For example, despite obtaining similar diffraction patterns as those analyzed by Sikorski and Atkins, Sharma et al. assign different unit cells and thus their proposed fibril structure contains significant structural differences.

More recent studies have investigated polyQ fibril structure using solid state NMR. Schneider et al. [56] studied a series of peptides with polyQ repeats ranging from 15–54
residues. They proposed that the basic structural motif of $D_2Q_{15}K_2$ fibrils is an extended $\beta$-strand structure, which assembles to form antiparallel $\beta$-sheets. Based on their interpretation of solid state NMR spectra of $GK_2Q_{38}K_2$ and $GK_2Q_{54}K_2$ aggregates, Schneider et al. propose that fibrils prepared from larger polyQ peptides adopt $\beta$-arc structures (Figure 1.1c), similar to those observed in $A\beta$ [57].

The $\beta$-arc model has been challenged by other solid state NMR studies. Van der Wel and coworkers have measured solid state NMR spectra of polyQ fibrils prepared from both model and more complex peptides that contain flanking sequences found in the huntingtin protein [58–61]. Their fibril spectra are very similar to those reported by Schneider et al. [56]; however, van der Wel and coworkers believe the data are inconsistent with a $\beta$-arc structure. Instead, they hypothesize that polyQ fibrils are composed of canonical antiparallel $\beta$-sheet structures that contain reverse-hairpin turns [61] (Figure 1.1d).

There have been a number of computational studies that have examined the structure of polyQ fibrils. Early simulation studies suggested that $\beta$-helical nanotubes are potentially stable fibril structures [62–65]. However, these studies did not rigorously validate their simulation results against experiments. More recent and rigorous computational studies [51, 66] that employ multiple force fields clearly show that $\beta$-nanotubes are highly unstable. Instead, these studies indicate that the most stable fibril architectures are $\beta$-arc and $\beta$-sheet structures.

### 1.3.4 Proposed Aggregation Mechanisms

Mechanistic studies on polyQ fibril aggregation have been pioneered by Wetzel and coworkers [59, 67–71]. Using sedimentation assays [72], the Wetzel group [35] proposes that simple polyQ peptide fibril aggregation proceeds via a classical nucleated growth-polymerization model. Interestingly, their systematic studies examining the aggregation concentration dependence of $K_2Q_nK_2$ peptides indicate that polyQ fibrils are homogeneously nucleated and that the critical nucleus size is a monomer for peptides containing repeats of more than 25 Gln residues [32, 59].

Traditionally, a monomeric nucleus is considered unusual in polymer theory [32]. To
Figure 1.2: Proposed aggregation mechanisms for polyQ fibril formation. The region shaded in gray represents the aggregation pathway proposed by Wetzel and coworkers. In this pathway, nucleation occurs through the energetically unfavorable structural conversion of a structurally disordered monomeric polyQ peptide to a $\beta$-sheet-rich structure that can initiate fibril formation. The alternative pathway is proposed by Pappu and coworkers. In this mechanism, polyQ peptides oligomerize into globule-like aggregates. These aggregates contain interfaces that promote the formation of $\beta$-sheet structures that eventually lead to fibrils. This figure was reprinted and adapted from [38], copyright © (2009), with permission from Elsevier.

To rationalize their observations, Wetzel and coworkers [32] proposed that fibril nucleation is initiated by the energetically unfavorable conversion of a structurally disordered polyQ peptide into a $\beta$-sheet-rich peptide through the formation of a hairpin turn (Figure 1.2). Wetzel and coworkers have published a number of elegant mutagenesis studies [59, 69] that provide evidence supporting this model. Their studies show, for example, that polyQ peptides containing $\beta$-hairpin enhancing motifs (e.g. tryptophan zippers, Pro-Gly insertions, Cysteine disulfide bonds, salt bridge forming residues) result in significantly increased aggregation kinetics compared to unmutated peptide sequences.

The classical nucleated growth-polymerization model has been challenged in recent years...
by a series of computational studies conducted primarily by Pappu and coworkers [36–38, 41, 44, 62]. Their simulations corroborate [36, 41] experimental evidence that monomeric polyQ peptides are structurally disordered in solution-state. However, in contrast to the Wetzel model [73], Pappu and coworkers predict that β-sheet formation of monomeric polyQ peptides becomes more energetically unfavorable with increasing Gln repeat lengths. Their computational data [37, 38] indicates that the spontaneous formation of dimers and higher order oligomers increases dramatically as a function of polyQ repeat length due to the formation of non-specific hydrogen bonding interactions between main chain and side chain amides.

In light of their results, Pappu and coworkers argue that polyQ fibrils are not nucleated via a coil to β-sheet structural transition. Instead, they hypothesize that monomeric polyQ peptides adopt disordered “globule” structures that non-specifically aggregate into high molecular weight oligomers [38, 44] (Figure 1.2). According to this model, these aggregates contain interfaces that promote the formation of β-sheet-rich structures that then structurally convert into amyloid-like fibrils [38]. The Pappu model has prompted several researchers to re-evaluate the interpretation of experimental aggregation kinetics data. For example, both the Pappu [44] and Murphy groups [74] argue that the Wetzel et al. sedimentation assay data are insufficient to differentiate between a nucleated growth-polymerization and more complex aggregation models.

Reliable experimental data supporting the Pappu model remains scarce. Currently, only one study, by Lee et al. [75], claims to observe the soluble oligomeric aggregates predicted by the Pappu model. However, the presence of these oligomer aggregates may be due to an artifact that occurred during sample preparation [35]. It is well known that many synthetic preparations of amyloidogenic peptides contain micro-aggregates that can heterogeneously seed fibrils [76]. To guard against this, “disaggregation” protocols [72, 76, 77] have been developed to remove these aggregates.

These protocols typically treat synthetic peptides with volatile, fluorinated solvents such as trifluoroacetic acid (TFA) and hexafluoroisopropanol (HFIP) that disrupt hydrogen bonding interactions and dissolve aggregates. As shown by Kar et al. [70], rigorous disaggregation of polyQ peptides do not result in the formation of high molecular weight amorphous aggre-
gates during fibril formation. Furthermore, studies [78, 79] show that HFIP can drive peptides towards oligomeric aggregate formation. Thus, the experimental evidence of oligomeric aggregates observed by Lee et al. [75] may be due to the failure to completely remove HFIP during disaggregation.

1.4 THE NEED FOR NEW BIOPHYSICAL TOOLS TO STUDY AMYLOID-LIKE FIBRILS

The lack of consensus regarding the structure and aggregation mechanism(s) of polyQ fibrils motivates the need to develop incisive biophysical tools that can quantitatively discriminate between different proposed models. The fact, for example, that three independent groups [52–55] analyzed similar X-ray fiber diffraction data and proposed four distinctly different structures highlights the difficulties associated with using conventional biophysical techniques to study polyQ amyloid-like fibrils.

Traditional methods such as X-ray diffraction and solution-state NMR typically cannot be utilized to study amyloid-like fibrils due to their insoluble and non-crystalline nature. Currently, the gold standard biophysical method to study amyloid-like fibrils is solid state NMR. Solid state NMR has been used with great success to solve several fibril structures [80–83]. An impressive array of sophisticated pulse sequences and methodologies have been developed to measure dihedral angles and distance constraints in fibrils [84–87]. In recent years, for example, the development of high-field dynamic nuclear polarization experiments have enabled the measurement of very high signal-to-noise fibril spectra [87, 88].

These sophisticated measurements, however, remain highly challenging and expensive. In order to obtain high signal-to-noise spectra, typical solid state NMR studies require extensive and costly isotopic $^{13}$C and $^{15}$N labeling, which can be very challenging to incorporate into proteins. Furthermore, spectral data collection typically requires long acquisition times that can take days and consume large quantities of precious samples. Finally, since spectral dispersion is often poor, many solid state NMR studies provide only qualitative information on the secondary structure of amyloid fibrils.
Ultraviolet resonance Raman (UVRR) spectroscopy is a powerful biophysical tool that can provide detailed molecular-level insights into polyQ solution-state and fibril structure. An advantage of UVRR is that, compared to NMR methods, structural information can be obtained quickly and under dilute concentrations [89]. Lednev and coworkers [90–95] have shown that UVRR is an excellent tool for probing the cross-β structure of fibrils, as well as monitoring the conformational changes that occur in amyloidogenic peptides during aggregation. In recent years, Asher and coworkers have identified several UVRR spectroscopic markers that are highly sensitive to the structure and hydrogen bonding environment of the peptide backbone [96–98], as well as amino acid side chains [99–103]. This has enabled new and deep insights into protein folding [104–107]. The primary goal of this dissertation work is to discover new spectroscopic markers that can be utilized to quantitatively investigate the structure of polyQ-rich peptides in solution-state and fibril aggregates.
2.0 UV RESONANCE RAMAN SPECTROSCOPY

2.1 CLASSICAL ELECTRODYNAMICS THEORY OF RAYLEIGH AND RAMAN SCATTERING

Figure 2.1 illustrates Rayleigh and Raman scattering from a classical electrodynamics perspective. When electromagnetic radiation excites a molecule, its electron cloud oscillates at the frequency of the incident light, and results in a displacement of charge that induces a change in the dipole moment. Because electrons are being accelerated, energy is radiated in the form of light [108]. Most of the radiating light occurs at the same frequency ($\omega_I$) as the incident electromagnetic radiation and is said to be elastically scattered. Assuming the wavelength of the exciting electromagnetic radiation is much larger than the dimensions of the molecule, this phenomenon is called Rayleigh scattering.

As shown in Figure 2.1, not all the light is elastically scattered. The motions of the oscillating electrons can couple to slower moving nuclear vibrational motions in the molecule. This coupling results in Raman (inelastic) scattering, where the electrons oscillate at a beat frequency. The Raman scattered light can be shifted to higher ($\omega_I + \omega_S$) or lower ($\omega_I - \omega_S$) frequencies with respect to the incident electromagnetic radiation. Scattered light that occurs at lower frequencies than the exciting incident radiation is said to be Stokes-shifted. In contrast, scattered light that occurs at higher frequencies than the incident light is anti-Stokes shifted.

Resonance Raman scattering occurs when the exciting radiation falls within the “natural” frequencies of the electron oscillators. These natural frequencies correspond to electronic transitions that, in biological molecules, typically lie in the deep ultraviolet (UV). Resonance Raman results in a tremendous enhancement in the scattering signal because the exciting
radiation is tuned into the natural frequencies of the electron oscillators. Another advantage is that only those vibrations that couple most efficiently to the electronic transition being excited are observed in resonance Raman spectra.

This second advantage confers the unique selectivity of resonance Raman spectroscopy. The vibrations of different chromophores can be selectively excited by judiciously tuning the wavelength of the excitation light so that it lies within a specific absorption band \cite{89}. This dramatically simplifies resonance Raman spectra (Figure 2.2) and consequently relieves spectral congestion that plagues non-resonance Raman and FTIR spectra of complex biological macromolecules such as proteins. For example, in the case of myoglobin, vibrations of the heme group can be excited by tuning into the Soret band at $\sim 400$ nm. Excitation in the UV, at the $\sim 220$ nm, tunes into the $\pi \rightarrow \pi^*$ transition of aromatic amino acids, which enables side chain vibrations of tryptophan, tyrosine, and phenylalanine to be studied. Deeper excitation in the UV at $\sim 180$ nm–$210$ nm occurs within the $\pi \rightarrow \pi^*$ transitions of amides, which enables investigations of the peptide backbone and the amino acid side chains of Gln.
Figure 2.2: Selectivity advantage of resonance Raman spectroscopy. Excitation within different electronic absorption bands of myoglobin enable the selective enhancement of Raman bands that derive from vibrations of different chromophores. Adapted with permission from [89]. Copyright © (2012), American Chemical Society.

2.2 QUANTUM MECHANICAL THEORY OF RAYLEIGH AND RAMAN SCATTERING

The classical theory correctly predicts the existence of Rayleigh and Raman light scattering. However, it fails to provide accurate insights into the physical origins of these phenomena, as well as predict scattering intensities. The accurate theoretical treatment of Rayleigh and Raman scattering requires quantum mechanics. The traditional theoretical treatment of Rayleigh and Raman scattering begins with the dispersion equation originally derived by Kramers and Heisenberg [110]. The first quantum-mechanical derivation of the dispersion equation was given by Born, Heisenberg, and Jordan [111] using second order perturbation theory. Dirac [112] later expanded on this derivation by also quantizing the radiation field.

A key result of applying second order perturbation theory is that Rayleigh and Raman scattering belong to two photon processes, as shown in Figure 2.3. For these phenomen-
Figure 2.3: Energy diagrams showing various two photon optical processes. (a) Rayleigh scattering: a molecule absorbs a photon of frequency $\omega_I$, which excites it into a virtual state. A photon of frequency $\omega_S = \omega_I$ is emitted, destroying the virtual state, so that the molecule’s final state is the same as its initial state. (b) Non-resonance Raman scattering: a molecule absorbs a photon of frequency $\omega_I$, which excites it into a virtual state. The virtual state is destroyed following the emission of a scattered photon of frequency $\omega_S \neq \omega_I$. (c) Resonance Raman scattering: the frequency, $\omega_I$, of an incident photon corresponds to an electronic transition from $S_0 \rightarrow S_1$. A scattered photon of frequency $\omega_S$ is emitted so that the molecule’s final state is different from the initial state. (d) Two photon absorption: a molecule absorbs two photons of frequency $\omega_I$ and $\omega'_I$. (e) Two photon emission: a molecule emits two photons of frequency $\omega_S$ and $\omega'_S$.


2.2.1 The Kramers-Heisenberg Dispersion relation

The central equation to understand Rayleigh and Raman scattering is the Kramers-Heisenberg dispersion relation, which we will derive here. To derive the dispersion relation,
we begin with the transition rate (see APPENDIX A, Derivation of the Transition Scattering Rate for a detailed derivation and discussion) associated with two photon scattering \[114, 115\]:

\[
W_{fi} = \frac{\pi E_I^2 E_S^2}{8h^3} \sum_r \left| \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | i \rangle}{\omega_{ri} + \omega_S} \right|^2 \rho_f \tag{2.1}
\]

where \(E, \epsilon, \omega\) are the amplitudes, polarization directions, and frequencies, respectively, of the incident (denoted by the subscript \(I\)) and scattered (denoted by the subscript \(S\)) electromagnetic fields. The term, \(\omega_{ri}\), is equal to the frequency difference between an arbitrary eigenstate, \(|r\rangle\), and the initial state of the system, \(|i\rangle\) (i.e. \(\omega_r - \omega_i\)). The transition dipole operators, \(\hat{\mu}\), couple the molecule to the incident electromagnetic field perturbing the system, and the subscripts, \(\sigma\) and \(\rho\), denote the directions of the transition moments. The density of \(|f\rangle\) states is denoted by \(\rho_f\).

The differential scattering cross section, \(d\sigma/d\Omega_f\), can be derived from the scattering rate by dividing eq. 2.1 by the incident photon flux, \(\Phi\) (eq. 2.2), and substituting \(E_S\) and \(\rho_f\) with eqs. 2.3 and 2.4 \[115\] (see eq. A.15):

\[
\Phi = \frac{E_I^2 c}{8\pi h \omega_I} \tag{2.2}
\]

\[
E_S = \frac{8\pi h \omega_S}{V} \tag{2.3}
\]

\[
\rho_f = \frac{\omega_S^2 V}{(2\pi c)^3 h} d\Omega_f \tag{2.4}
\]

\[
\frac{d\sigma}{d\Omega_f} = \frac{\omega_I \omega_S^3}{c^4 \hbar^2} \sum_r \left| \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | i \rangle}{\omega_{ri} + \omega_S} \right|^2 \tag{2.5}
\]

where \(c\) is the speed of light and \(d\Omega_f\) is the scattering solid angle.

Eq. 2.5 is the famous Kramers-Heisenberg dispersion relation. The scattering cross section, which is proportional to the scattering intensity, can be derived from this equation...
by averaging over all molecular orientations, integrating over the full $4\pi$ solid angle, and summing over all incident and scattered light polarizations [115]:

$$
\sigma = \frac{8\pi e^4 \omega_I^2 \omega_S^3}{9e^4} \frac{1}{\hbar^2} \sum_{\rho,\sigma} \sum_r \left| \frac{\langle f|\hat{r}_\rho|r\rangle \langle r|\hat{r}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{r}_\sigma|r\rangle \langle r|\hat{r}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 
$$

(2.6)

where we utilize the fact that $\hat{\mu} = e\hat{\tau}$ ($\hat{\tau}$ is the position operator and $e$ is the charge of an electron). The cross section equation can be further simplified by introducing the second rank tensor, $(\alpha_{\rho\sigma})_{fi}$, which is known as the molecular polarizability:

$$(\alpha_{\rho\sigma})_{fi} = \frac{1}{\hbar} \sum_r \left[ \frac{\langle f|\hat{r}_\rho|r\rangle \langle r|\hat{r}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{r}_\sigma|r\rangle \langle r|\hat{r}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right] 
$$

(2.7)

so that eq. 2.6 now becomes:

$$
\sigma = \frac{8\pi e^4 \omega_I^2 \omega_S^3}{9e^4} \sum_{\rho,\sigma} |(\alpha_{\rho\sigma})_{fi}|^2
$$

(2.8)

The quantum mechanical origins of Rayleigh and Raman scattering, as discussed earlier in regards to Figure 2.3, is revealed in eq. 2.7 and 2.8. These expressions indicate that the virtual state is actually a linear combination of all the eigenstates in the system. Since the molecule does not possess an eigenstate that is resonant with this virtual state, energy conservation appears to be violated. However, this is not the case since the absorption and emission events occur nearly simultaneously. Given the short lifetime of the intermediate state, Heisenberg’s uncertainty principle ($\Delta E \Delta t \geq \hbar$) states that the energy uncertainty ($\Delta E$) of the system is very large over the short time interval ($\Delta t$) during which the virtual state exists. Thus, any $|r\rangle$ eigenstate of the system can be momentarily produced, thereby maintaining energy conservation [113, 114].

For Rayleigh scattering, there is no frequency shift in the scattered light relative to the incident light ($\omega_I = \omega_S$), so that the initial and final states of the excited molecule are the same. From eq. 2.8, it can be seen that, under this condition, the cross section scales as $\omega_I^4$, in accordance Rayleigh’s famous scattering law. This $\omega_I^4$ frequency dependence explains, for example, the color of the sky. Molecules in the atmosphere scatter shorter (blue) wavelength light more efficiently than longer (red) wavelength light. This is why the color of the sky is blue during the day.
In Raman scattering, there is an exchange of a quantum of energy so that the initial and final states of the molecule are different. As shown in Figure 2.3b, the frequency difference between the $\omega_I$ and $\omega_S$ corresponds to a vibrational transition from state $|i\rangle$ to $|f\rangle$. In this sense, Raman spectroscopy is complementary to infrared (IR) spectroscopy. However, because Raman is a two photon scattering process and IR is a one photon absorption process, the selection rules between these two spectroscopies are different.

### 2.3 Non-Resonance vs. Resonance Scattering

The traditional approach to understand resonance Raman scattering (Figure 2.3c) is to utilize the vibronic theory championed primarily by Albrecht and coworkers [115–118]. This theory treats Rayleigh and Raman scattering in the framework of vibronic spectroscopies, where, the resonance scattering intensities of the electronic ground state vibrations are intimately tied to transitions of excited vibronic states. In this section, we examine the vibronic theory by Albrecht in order to elucidate important physical aspects of resonance and non-resonance Raman scattering.

#### 2.3.1 General Vibronic Theory

The wave functions in eq. 2.7 depend on both the electronic and vibrational states. Albrecht [117] first assumed that the electronic and vibrational state wavefunctions are separable by invoking the Born-Oppenheimer approximation:

$$|\Psi_m\rangle = |\phi_e(q,Q)\chi_{em}(Q)\rangle = |\phi_e(q,Q)\rangle |\chi_{em}(Q)\rangle = |\phi_e\rangle |\chi_{em}\rangle$$ (2.9)

where $|\phi_e\rangle$ is the wave function for an electronic state, $e$, which depends on the complete set of internal coordinates of both the nuclei ($Q$) and the electrons ($q$). The state $|\chi_{em}\rangle$ is the vibrational wave function associated with the resonant electronic excited state, $e$, and the vibrational state, $m$. Inserting eq. 2.9 into eq. 2.7 gives the following equation:
\[
(\alpha_{\rho\sigma})_{g f',g i'} = \frac{1}{\hbar} \sum_e \sum_{r'} \left[ \frac{\langle \chi_{g f'} | \langle \phi_g | \hat{r}_\rho | \phi_e \rangle | \chi_{er'} \rangle \langle \chi_{er'} | \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle | \chi_{gi'} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I} + \frac{\langle \chi_{g f'} | \langle \phi_g | \hat{r}_\sigma | \phi_e \rangle | \chi_{er'} \rangle \langle \chi_{er'} | \langle \phi_e | \hat{r}_\rho | \phi_g \rangle | \chi_{gi'} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \right]
\]  
(2.10)

where \(\omega_{ri} = \omega_{er'} - \omega_{gi'}\). The electronic wave functions can be expanded using a Taylor’s series about the nuclear coordinates around the equilibrium ground state structure. To first order theory, the result of this Herzberg-Teller expansion [117] is:

\[
|\phi_g\rangle = |\phi_g^{(0)}\rangle + \frac{1}{\hbar} \sum_a \sum_{t\neq g} \frac{(h_a^{(0)})_{es} \cdot \hat{Q}_a}{\omega_g^{(0)} - \omega_t^{(0)}} |\phi_t^{(0)}\rangle
\]  
(2.11)

\[
|\phi_e\rangle = |\phi_e^{(0)}\rangle + \frac{1}{\hbar} \sum_a \sum_{s\neq e} \frac{(h_a^{(0)})_{et} \cdot \hat{Q}_a}{\omega_g^{(0)} - \omega_s^{(0)}} |\phi_s^{(0)}\rangle
\]  
(2.12)

where \(|\phi_g^{(0)}\rangle\), \(|\phi_e^{(0)}\rangle\), \(|\phi_s^{(0)}\rangle\), and \(|\phi_t^{(0)}\rangle\) are the unperturbed electronic wave functions for states \(g\), \(e\), \(s\), and \(t\), respectively. State \(g\) is the electronic ground state, while \(s\) and \(t\) are non-resonant electronic excited states. The operator, \(\hat{Q}_a\), is the displacement operator for the \(a^{th}\) normal mode. The Herzberg-Teller coupling integrals, \((h_a^{(0)})_{et}\) and \((h_a^{(0)})_{es}\), are defined as:

\[
(h_a^{(0)})_{et} = \langle \phi_g^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_t^{(0)} \rangle
\]  
(2.13)

\[
(h_a^{(0)})_{es} = \langle \phi_s^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle
\]  
(2.14)

Substituting eqs. 2.11 and 2.12 into eq. 2.10 gives the following equation [115] (see APPENDIX A, Derivation of the Albrecht A, B, and C Terms for details):

\[
(\alpha_{\rho\sigma})_{g f',g i'} \approx A + B + C
\]  
(2.15)

where,

\[
A = \frac{1}{\hbar} \sum_{e \neq g} \sum_{r'} \left[ \frac{\langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I} + \frac{\langle \phi_g | \hat{r}_\sigma | \phi_e \rangle \langle \phi_e | \hat{r}_\rho | \phi_g \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S} \right] \langle \chi_{g f'} | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle
\]  
(2.16)
versely proportional to the frequency differences, \( \omega \) terms due to Herzberg-Teller coupling are proportional to the displacement, \( Q \), and the resonant electronic excited state, \( \chi \). Typically, the energy difference between \( |s\rangle \) and \( |t\rangle \) electronic states are close in energy to the electronic ground state, \( |g\rangle \), and the resonant electronic excited state, \( |e\rangle \). For a particular vibration, \( a \), the scattering intensities that derive from the \( B \) and \( C \) terms due to Herzberg-Teller coupling are proportional to the displacement, \( Q_a \), and inversely proportional to the frequency differences, \( \omega_{er' - g} - \omega_s \). The frequency dependence of the \( B \) and \( C \) terms means that significant coupling will occur only when the \( |s\rangle \) and \( |t\rangle \) electronic states are close in energy to the electronic ground state, \( |g\rangle \), and the resonant electronic excited state, \( |e\rangle \). Typically, the energy difference between \( |g\rangle \) and \( |t\rangle \) is much larger than that between \( |e\rangle \) and \( |s\rangle \) \( (i.e., \omega_{g'} - \omega_s \gg \omega_{e} - \omega_s) \). This means that, in typical cases, the \( C \) term can be neglected, so that eq. 2.15 becomes [115]:

\[
(\alpha_{\rho a})_{g'g'} \approx A + B
\]  \hspace{1cm} (2.19)
2.3.2 Non-Resonance Rayleigh and Raman Scattering

Eq. 2.19 can be used to understand non-resonance and resonance scattering. We will first briefly consider the non-resonance case, wherein the frequency of the exciting electric field is far away from an electronic absorption band such that \( |\omega_{e} - \omega_{g} - \omega_{I}| \approx 0 \). In this situation, \( \omega_{e} - \omega_{g} - \omega_{I} \approx \omega_{e} - \omega_{g} + \omega_{S} \) and \( \omega_{e} - \omega_{g} \approx \omega_{g}^{(0)} - \omega_{g}^{(0)} \). Using these approximations, eq. 2.19 can be written as:

\[
A = \frac{1}{\hbar} \sum_{e \neq g} \left[ \frac{\langle \phi_{g} | \hat{r}_{p} | \phi_{e} \rangle}{\omega_{e}^{(0)} - \omega_{g}^{(0)} - \omega_{I}} \langle \phi_{e} | \hat{r}_{\sigma} | \phi_{g} \rangle + \frac{\langle \phi_{g} | \hat{r}_{\sigma} | \phi_{e} \rangle}{\omega_{e}^{(0)} - \omega_{g}^{(0)} + \omega_{I}} \langle \phi_{e} | \hat{r}_{p} | \phi_{g} \rangle \right] \langle \chi_{g} | r' \rangle \langle \chi_{g'} \rangle \tag{2.20}
\]

\[
B = \frac{1}{\hbar} \sum_{e \neq g} \sum_{a \neq e} \sum_{s \neq e} \left\{ \left[ \frac{2(\omega_{e}^{(0)} - \omega_{g}^{(0)})}{(\omega_{e}^{(0)} - \omega_{g}^{(0)})^{2} - \omega_{I}^{2}} \frac{\langle \phi_{g} | \hat{r}_{p} | \phi_{e} \rangle\langle \phi_{e} | \hat{r}_{\sigma} | \phi_{g} \rangle}{\omega_{e}^{(0)} - \omega_{g}^{(0)} - \omega_{I}} \langle \chi_{g} | r' \rangle \langle \chi_{g'} \rangle \right] \right\} \right. \tag{2.21}
\]

To derive these equations, all \( r' \) vibrational states were summed over using the identity shown in A.9. Eqs. 2.20 and 2.21 can be further simplified by assuming that the wavefunctions are all real, so that the numerators belonging to the \( \omega_{e}^{(0)} - \omega_{g}^{(0)} - \omega_{I} \) and \( \omega_{e}^{(0)} - \omega_{g}^{(0)} + \omega_{I} \) terms are equal. Thus, using the following algebraic identity,

\[
\frac{1}{a - b} + \frac{1}{a + b} = \frac{2a}{a^{2} - b^{2}} \tag{2.22}
\]

eqs. 2.20 and 2.21 can be simplified to:

\[
A = \frac{1}{\hbar} \sum_{e \neq g} \left[ \frac{2(\omega_{e}^{(0)} - \omega_{g}^{(0)})}{(\omega_{e}^{(0)} - \omega_{g}^{(0)})^{2} - \omega_{I}^{2}} \frac{\langle \phi_{g} | \hat{r}_{p} | \phi_{e} \rangle\langle \phi_{e} | \hat{r}_{\sigma} | \phi_{g} \rangle}{\omega_{e}^{(0)} - \omega_{g}^{(0)} - \omega_{I}} \langle \chi_{g} | r' \rangle \langle \chi_{g'} \rangle \right] \tag{2.23}
\]

\[
B = \frac{1}{\hbar} \sum_{e \neq g} \sum_{a \neq e} \sum_{s \neq e} \left\{ \left[ \frac{2(\omega_{e}^{(0)} - \omega_{g}^{(0)})}{(\omega_{e}^{(0)} - \omega_{g}^{(0)})^{2} - \omega_{I}^{2}} \frac{\langle \chi_{g} | r' \rangle \langle \chi_{g'} \rangle}{\omega_{e}^{(0)} - \omega_{g}^{(0)} - \omega_{I}} \right] \right. \tag{2.24}
\]
It can be seen that the $A$ term is non-zero only when $|\chi_{gi'}\rangle = |\chi_{gf'}\rangle$. Therefore, the $A$ term is solely responsible for Rayleigh scattering in non-resonance cases. In contrast, $B$ term is non-zero on three conditions: $|\chi_{gf'}\rangle = |\chi_{gi'} \pm 1\rangle$; the Herzberg-Teller coupling integrals are non-zero; and there are no interference effects that occur when summing over the electronic states. Thus, according to Albrecht’s vibronic theory, the $B$ term is solely responsible for non-resonance Raman scattering intensities of fundamentals.

### 2.3.3 Resonance Rayleigh and Raman Scattering

We now consider the case of resonance, where the excitation frequency lies within a molecule’s electronic absorption band. As $\omega_I$ approaches $\omega_{er'} - \omega_{gi'}$, eq. 2.10 is no longer valid because the scattering amplitude that derives from the term containing $\omega_{er'} - \omega_{gi'} - \omega_I$ becomes infinite. In reality, radiation dampening occurs due to the finite lifetime of populating the eigenstates. To account for this, we introduce a phenomenological damping coefficient, $\Gamma$, so that eq. 2.10 can be re-written as:

$$
(\alpha_{\rho\sigma})_{gf',gi'} = \frac{1}{\hbar} \sum_{e} \sum_{r'} \left[ \frac{\langle \chi_{gf'} | \langle \phi_g | \hat{r}_\rho | \phi_e \rangle | \chi_{er'} \rangle \langle \chi_{er'} | \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle | \chi_{gi'} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} + \frac{\langle \chi_{gf'} | \langle \phi_g | \hat{r}_\sigma | \phi_e \rangle | \chi_{er'} \rangle \langle \chi_{er'} | \langle \phi_e | \hat{r}_\rho | \phi_g \rangle | \chi_{gi'} \rangle}{\omega_{er'} - \omega_{gi'} + \omega_S + i\Gamma} \right] \quad (2.25)
$$

Using this expression, eq. 2.19 can be modified in two different ways. First, the summation over $e$ disappears, since the molecule is being excited into a specific electronic excited state. Second, the “non-resonant” $\omega_{er'} - \omega_{gi'} + \omega_S$ term in eq. 2.25 can be neglected because it is significantly smaller than the $\omega_{er'} - \omega_{gi'} - \omega_I$ term. Thus, eqs. 2.16 and 2.17 can be re-written as:

$$
A = \frac{1}{\hbar} \langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \sum_{r'} \left[ \frac{\langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{gi'} \rangle}{\omega_{er'} - \omega_{gi'} - \omega_I + i\Gamma} \right] \quad (2.26)
$$

24
\[ B = \frac{1}{\hbar^2} \sum_{r'} \sum_{a} \sum_{s \neq c} \left\{ \frac{\langle \phi_y^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_v^{(0)} | (\partial \hat{H} / \partial Q_a)_0 | \phi_s^{(0)} \rangle \langle \phi_r | \hat{r}_\sigma | \phi_g^{(0)} \rangle}{\omega_{e,r'} - \omega_{g,i'} - \omega_I + i \Gamma} \right\} \left\{ \frac{\langle \chi_{gf'} | \hat{Q}_{e,r'} | \chi_{e,r} \rangle \langle \chi_{e,r} | \hat{Q}_{a} | \chi_{g,i'} \rangle}{\omega_{e}^{(0)} - \omega_{s}^{(0)}} \right\} \]

(2.27)

It is impossible to sum over all \(|r'|\) vibrational states, as was done in deriving eqs. 2.23 and 2.24. This is because in resonance, the frequency terms in the denominators of the summations are now sensitive to only a subset of \(|r'|\) states.

A comparison of these equations with 2.23 and 2.24 reveals interesting differences in the mechanism of Rayleigh and Raman scattering between the resonance and non-resonance cases. Unlike the non-resonance situation, the resonance \(A\) and \(B\) terms are both responsible for Rayleigh and Raman scattering processes. In the case of the \(A\) term, Albrecht [117] notes that this is because the vibrational wavefunctions of the electronic ground and resonant excited states are generally not solutions to the same Schrödinger equation. Therefore, the Franck-Condon factors \(\left\langle \chi_{gf'} | \chi_{e,r'} \right\rangle \left\langle \chi_{e,r'} | \chi_{g,i'} \right\rangle\) in the \(A\) term need not be zero.

It is also interesting to note that the contribution of the \(A\) term to the Raman scattering intensity of a vibrational mode does not depend on vibronic mixing between the resonant excited state, \(e\), with other electronic states. In contrast, \(B\) term scattering depends on vibronic mixing of the resonant state, \(e\), with other electronic states. The \(1/(\omega_e^{(0)} - \omega_s^{(0)})\) dependence of the terms containing the Herzberg-Teller integrals in eq. 2.27 means that the lowest lying \(e\) to \(s\) state electronic transitions will be responsible for most of the \(B\) term scattering intensity.

### 2.3.4 Resonance Raman and the Electronic Excited State Geometry

Eqs. 2.26 and 2.27 indicate that the resonance Raman intensities are inherently sensitive to details of the electronic excited state potential energy surface. This can be most easily seen in the Albrecht \(A\) term, where only one electronic transition contributes to the resonance Raman scattering intensity. In the \(A\) term, information about the excited state geometry is encoded in the Franck-Condon factors that are in the summation of eq. 2.26. The electronic
excited state must be displaced along a particular coordinate for at least one of the Franck-Condon factors to be non-zero.

The simplest approach to determine the Franck-Condon integrals is to assume that the electronic ground and excited states are both harmonic and differ only in their equilibrium positions [119]. In this limit, no Duschinsky rotation occurs [119], and the vibrational frequencies of the normal modes do not change between the ground and excited states. Each vibrational \( j^{\text{th}} \) mode can be treated as a pair of ground and excited state harmonic oscillators that are separated in energy by some frequency, \( \omega_j \), and displaced relative to each other along a particular coordinate by an origin shift, \( \Delta \). As a result, the multidimensional Franck-Condon factors can be written as the products of one dimensional overlap integrals [119]. \( i.e. \)

\[
\langle \chi_{gf'} | \chi_{er'} \rangle = \prod_{j=1}^{N} \langle \chi_{gf'_j} | \chi_{er'_j} \rangle
\]  

(2.28)

Using eq. 2.28, the resonance Raman polarizability corresponding to the transition from \( |\chi_{gf'} = 0_1 \rangle \) to \( |\chi_{gf'} = 1_1 \rangle \) for the \( j = 1 \) Raman mode (denoted by the subscript 1) is [119]:

\[
\alpha_{0\rightarrow 1} = \frac{1}{\hbar} \langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \times \sum_{m_1} \sum_{m_2} \cdots \sum_{m_{3N-6}} \left[ \frac{\langle 1_1 | m_1 \rangle \langle m_1 | 0_1 \rangle \prod_{j=2}^{3N-6} \langle j_1 | m_j \rangle \langle m_j | 0_j \rangle}{\omega_e^0 - \omega_g^0 + \sum_{j=1}^{3N-6} (m_j) \omega_j - \omega_I + i\Gamma} \right]
\]  

(2.29)

where \( |\chi_{er'_j} \rangle = |m_j \rangle \).

The one dimensional Franck-Condon factors can be re-written using a set of recursion equations derived by Manneback [120]. Using these recursion relations, the Frank-Condon factors can be re-written as [119]:

\[
\langle 0_j | m_j \rangle \langle m_j | 0_j \rangle = \frac{\Delta_j^2}{2^m m!} e^{-\Delta_j^2/2}
\]  

(2.30)

\[
\langle 1_j | m_j \rangle \langle m_j | 0_1 \rangle = \frac{\Delta_j}{\sqrt{2}} [(0_1 | m_1 \rangle \langle m_1 | 0_1 \rangle - \langle 0_1 | m_1 - 1 \rangle \langle m_1 - 1 | 0_1 \rangle]
\]  

(2.31)

in the limit that the vibrational frequencies for the electronic ground and the excited state normal modes do not change. Using these equations, 2.29 can be written as [119]:

\[
26
\[ \alpha_{0 \rightarrow 1} = \left( \frac{1}{\hbar} \right) \left( \frac{\Delta_1}{\sqrt{2}} \right) \langle \phi_g | \hat{r}_\rho | \phi_e \rangle \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \]
\[ \times \sum_{m_1} \sum_{m_2} \cdots \sum_{m_{3N-6}} \left[ \frac{\langle 0_1 | m_1 \rangle \langle m_1 | 0_1 \rangle - \langle 0_1 | m_1 - 1 \rangle \langle m_1 - 1 | 0_1 \rangle \Pi_{j=2}^{3N-6} \langle 0_j | m_j \rangle \langle m_j | 0_j \rangle}{\omega_e^0 - \omega_g^0 + \sum_{j=1}^{3N-6} (m_j) \omega_j - \omega_I + i\Gamma} \right] \]

Eq. 2.32 shows that the resonance Raman cross sections for fundamentals are proportional to \( \Delta^2 \), the square of the displacement between the electronic ground and excited states along a particular normal coordinate. Thus, eq. 2.32 shows that resonance Raman band intensities are sensitive to the geometry changes that occur between the electronic ground and excited states of molecules.
3.0 UV RESONANCE RAMAN INVESTIGATION OF THE AQUEOUS SOLVATION DEPENDENCE OF PRIMARY AMIDE VIBRATIONS


**Author Contributions:** D.P. acquired and analyzed the UVRR data with the assistance of R.S.J. and E.M.D. D.P. and S.G. acquired and analyzed X-ray diffraction data. Z.H. performed DFT calculations and normal mode analysis. D.P. and N.S. assisted in analyzing the results from the DFT calculations. The manuscript was prepared by D.P. and S.A.A. with the assistance of R.S.J. and E.M.D.

We investigated the normal mode composition and the aqueous solvation dependence of the primary amide vibrations of propanamide. Infrared, normal Raman, and UVRR spectroscopy were applied in conjunction with density functional theory (DFT) to assign the vibrations of crystalline propanamide. We examined the aqueous solvation dependence of the primary amide UVRR bands by measuring spectra in different acetonitrile/water mixtures. As previously observed in the UVRR spectra of N-methylacetamide, all of the resonance enhanced primary amide bands, except for the Amide I (Am I), show increased UVRR cross sections as the solvent becomes water-rich. These spectral trends are rationalized by a
model wherein the hydrogen bonding and the high dielectric constant of water stabilizes the
*ground state* dipolar $^\cdot$O$\cdots$C$=$NH$_2^+$ resonance structure over the neutral O$\cdots$C$=$NH$_2$ resonance structure. Thus, vibrations with large C–N stretching show increased UVRR cross sections because the C–N displacement between the electronic ground and excited state increases along the C–N bond. In contrast, vibrations dominated by C–O stretching, such as the Am I, show a decreased displacement between the electronic ground and excited state, which result in a decreased UVRR cross section upon aqueous solvation. The UVRR primary amide vibrations can be used as sensitive spectroscopic markers to study the local dielectric constant and hydrogen bonding environments of the primary amide side chains of Gln and Asn.

### 3.1 INTRODUCTION

The primary amide functional group is of significant biological interest since it is found in the side chains of Gln and Asn. These side chains may be of structural and functional significance to peptides and proteins since they can participate in both intra- and inter-molecular hydrogen bonding, which may be important to the formation and stabilization of prion and amyloid-like fibril aggregates that are involved in a number of protein diseases [53, 54, 56, 121, 122]. Given the importance of these primary amide groups, it is of great value to find spectroscopic markers that can be used to monitor the Asn and Gln hydrogen bonding and dielectric environments.

UVRR spectroscopy is a powerful tool for studying the conformations of proteins [89], as well as for determining hydrogen bonding, protonation states, and local dielectric environments of aromatic amino acids [101], arginine [103], and histidine [123, 124] side chains. Deep UV excitation (∼200 nm) selectively enhances the peptide bond secondary amide vibrations of the protein backbone [89] and the primary amide vibrations of the Asn and Gln side chains [125]. Investigations of the secondary amide vibrations [96, 97, 105, 106, 126–128] have developed a deep understanding of the spectral dependence of the peptide bond secondary structure and its hydrogen bonding. This understanding has enabled incisive investigations
of protein and peptide structure. The work presented here is developing a similar deep understanding of the structural, hydrogen bonding, and dielectric environmental dependence of primary amide vibrations.

We investigated propanamide, one of the simplest primary amide compounds with a structure similar to that of the Asn and Gln side chains. We assigned the vibrations observed in the infrared, non-resonance Raman, and UVRR spectra of crystalline propanamide with the aid of DFT. We then examined the solution behavior of the primary amide UVRR bands in mixtures of acetonitrile and water. The primary amide bands are very sensitive to their hydrogen bonding and dielectric environments. These bands will be useful as spectroscopic probes to monitor the side chain environment and structure of Gln and Asn.

### 3.2 EXPERIMENTAL SECTION

#### 3.2.1 Materials

Propanamide (CH$_3$CH$_2$CONH$_2$, 97% purity) and acetonitrile (HPLC, far-UV grade) were purchased from Acros Organics. N-methylacetamide (CH$_3$CH$_2$CONHCH$_3$, NMA, ≥99% purity) and sodium perchlorate (NaClO$_4$, ≥98% purity) were purchased from Sigma-Aldrich. D$_2$O (99.9% atom D purity) was purchased from Cambridge Isotope Laboratories, Inc.

#### 3.2.2 Sample Preparation

The propanamide solid samples consisted of a crystalline powder, which was used without further purification or re-crystallization. N-deuterated propanamide crystals were prepared by multiple re-crystallizations in D$_2$O. Propanamide solutions were prepared in H$_2$O at 10 mM concentration. Samples prepared in mixtures of acetonitrile and water were prepared at 30 mM concentrations. For these experiments, NaClO$_4$ (200 mM) was used as an internal standard. The N-deuterated propanamide solution was prepared at 30 mM concentration in pure D$_2$O.
3.2.3 Infrared and Non-Resonance Raman Spectroscopy

The mid-infrared spectrum of propanamide crystals was measured in the 600–4000 cm\(^{-1}\) region at 1 cm\(^{-1}\) resolution. The data were collected using a Perkin-Elmer model Spectrum 100 series FTIR equipped with a Universal diamond ATR. The propanamide crystals were lightly ground for \(~30\) s using a mortar and pestle in order to ensure good optical contact between the sample and the diamond crystal. The sample was placed on the diamond crystal and a force of \(~145\) N was applied using a pressure arm.

The visible excitation Raman spectra of propanamide crystals were measured using a Renishaw inVia Raman spectrometer equipped with a research-grade Leica microscope. Spectra were collected using a 5\(\times\) objective lens with a \(~2\) cm\(^{-1}\) resolution spectrometer. The 633 nm exciting line was generated by a HeNe laser. The 380 cm\(^{-1}\), 918 cm\(^{-1}\), 1376 cm\(^{-1}\), 2249 cm\(^{-1}\), and 2942 cm\(^{-1}\) bands of acetonitrile [129] were used for calibration.

3.2.4 UVRR Spectroscopy

The UVRR spectra of crystalline propanamide were measured using cw 229 nm light generated by an Innova 300 FreD frequency doubled Ar\(^+\) laser [130]. Solid samples were spun by using a cylindrical brass rotation cell to prevent thermal degradation or photodegradation. A SPEX Triplemate spectrograph, modified for use in the UV, was utilized to disperse the Raman scattered light. A Spec-10 system charge-coupled device (CCD) camera (Princeton Instruments, Model 735-0001) was employed to detect the scattered light.

UVRR solution-state measurements were made using \(~204\) nm excitation. The UV light was generated by Raman shifting the third harmonic of an Nd:YAG Infinity laser (Coherent, Inc.) in H\(_2\) gas (\(~30\) psi) and selecting the fifth anti-Stokes line. Solutions were circulated using a thermostatted (20°C) flow cell [104] to prevent accumulation of photodegradation products. The scattered light was dispersed and imaged using a double monochromator, modified for use in the UV in a subtractive configuration, [131] and detected with a Spec-10 CCD camera.
3.2.5 UV Absorption Measurements

Absorption spectra were taken of 30 mM propanamide and 10 mM N-methylacetamide solutions dissolved in acetonitrile and water. A Varian Cary 5000 spectrophotometer with a 0.2 mm path length quartz cuvette was used for all measurements.

3.2.6 Raman Cross Section Calculations

The UVRR scattering cross sections of the spectrally deconvoluted propanamide Raman bands (see APPENDIX B for details) were calculated using the following equation [125]:

\[ \sigma_i = \frac{I_i k_i C_r \sigma_r}{I_r k_i C_i} \left( \frac{\epsilon_i + \epsilon_{ex}}{\epsilon_r + \epsilon_{ex}} \right) \]  \hspace{1cm} (3.1)

where \( \sigma_i \) is the cross section of the \( i^{th} \) propanamide Raman band, and \( \sigma_r \) is the cross section of the 932 cm\(^{-1} \) ClO\(_4^{-} \) stretching band from our internal standard. \( C_r \) and \( C_i \) are the concentrations of NaClO\(_4\) and propanamide, respectively. The factors \( k_r \) and \( k_i \) are the spectrometer efficiencies at the 932 cm\(^{-1} \) and \( i^{th} \) propanamide Raman bands. Assuming only the analyte absorbs, the factors \( \epsilon_i, \epsilon_r, \) and \( \epsilon_{ex} \) are the molar absorptivities at the \( i^{th} \) propanamide Raman band, the 932 cm\(^{-1} \) band, and the excitation wavelength, respectively. The expression in the parenthesis corrects the cross section measurement for self-absorption [132, 133]. The Raman cross section of the 932 cm\(^{-1} \) reference band, \( \sigma_r \), at 204 nm excitation, was estimated to be \( \sim 1.18 \times 10^{-27} \text{ cm}^2\text{-molecule}^{-1}\text{-sr}^{-1} \) by extrapolating the Raman cross section measurements of Dudik et al. [134]

3.3 COMPUTATIONAL SECTION

The DFT calculations [135] were performed using the GAUSSIAN 09 package [136] with the M06-2X functional [137] and the 6-311++g** basis set. The calculated frequencies of the propanamide vibrations were calculated using a harmonic approximation and scaled linearly to the average band frequencies observed experimentally. The calculations simulated solvation implicitly by placing the propanamide molecule in an ellipsoidal cavity surrounded...
by a polarizable continuum dielectric modeled to simulate water. The potential energy distribution (PED) of each vibration was obtained from the GAUSSIAN output files by employing the GAR2PED program [138].

### 3.4 RESULTS AND DISCUSSION

#### 3.4.1 Normal Mode Analysis

The infrared and Raman spectra of propanamide were reported previously [139–144]. The most detailed normal mode analysis and assignments were carried out by Kuroda et al. [143] and Nandini and Sathyanarayana [144]. Kuroda et al. [143] utilized semi-empirical calculations that employed modified Urey-Bradley and valence force fields, while Nandini and Sathyanarayana [144] employed \textit{ab initio} Hartree-Fock calculations. The band assignments significantly differ between these studies, particularly in regard to the amide vibrations. In the work here, we performed a new normal mode analysis for propanamide using more accurate DFT calculations, and use these results to assign our infrared, normal Raman, and UVRR spectral bands.

#### 3.4.1.1 Propanamide Molecular Structure

We determined the crystal structure of our propanamide crystals (see APPENDIX B for details), and found it to be close to the structure reported previously [145]. The crystal unit cell is monoclinic (P2\textsubscript{1}/c space group), with four molecules per unit cell (Figure B1). The measured crystallographic axes and angles are a = 8.851(4), b = 5.750(2), c = 9.766(3), and $\beta = 114.780(15)^\circ$.

We optimized the propanamide geometry (Figure 3.1) for the DFT calculations by taking the crystal structure as an initial starting point and determining the minimum energy geometry. As in the crystal, the calculated ground state equilibrium geometry of propanamide shows $C_1$ symmetry since the NCCC dihedral angle deviates from 180°. Our results agree with the calculated structure of Nandini and Sathyanarayana [144]. However, our calculated structure differs from that of Kuroda et al. [143], who assumed a $C_s$ symmetry for
Figure 3.1: DFT-optimized structure of propanamide showing atomic numbering scheme used for normal mode analysis. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

Propanamide, since its crystal structure had not been determined at that time. Propanamide has 30 fundamental vibrations, which, for a C\textsubscript{1} point group are all both infrared and Raman active, and unlike in the C\textsubscript{s} point group, there is no differentiation between in-plane (a') and out-of-plane (a'') modes.

Table B4–Table B6 enable comparisons between the structural parameters of the energy minimized structure used in our calculations, and our measured crystal structure. The geometry and heavy atom bond lengths and angles of the minimized structure are close to the crystal structure, but not identical. This is most likely because the polarizable continuum model (PCM) employed in our calculations does not take into account crystal packing forces and hydrogen bonding.

**Vibrational Band Assignments**

Figure 3.2 shows the infrared, visible Raman, and UVRR spectra of CH\textsubscript{3}CH\textsubscript{2}CONH\textsubscript{2} and CH\textsubscript{3}CH\textsubscript{2}COND\textsubscript{2}. Table 3.1 compares the frequencies and relative intensities of the infrared and non-resonance Raman spectra. We utilized UVRR to help identify the primary amide Raman bands. The primary amide NV\textsubscript{1} electronic transition absorption band at \(~\)180 nm resonance enhances the 204 nm excited UVRR spectra and preresonance enhances the 229 nm excited UVRR spectra. We expect the most resonance enhanced bands will consist of vibrations with large contributions of C\textendash N stretching, since the excited state
Table 3.1: Frequencies (cm$^{-1}$) of Crystalline CH$_3$CH$_2$CONH$_2$ and CH$_3$CH$_2$COND$_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>CH$_3$CH$_2$CONH$_2$$^a$</th>
<th>CH$_3$CH$_2$COND$_2$$^{a,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrared</td>
<td>Raman</td>
</tr>
<tr>
<td>3356 vs, br</td>
<td>3356 s, br</td>
</tr>
<tr>
<td>3177 vs, br</td>
<td>3171 s, br</td>
</tr>
<tr>
<td>2943 m</td>
<td>2940 vs</td>
</tr>
<tr>
<td>2922 w</td>
<td>2909 s</td>
</tr>
<tr>
<td>2882 w</td>
<td>2882 sh</td>
</tr>
<tr>
<td>2811 m</td>
<td>2827 vw</td>
</tr>
<tr>
<td>2737 sh</td>
<td>2734 m</td>
</tr>
<tr>
<td>1643 sh, vs</td>
<td>1676 vw</td>
</tr>
<tr>
<td>1628 vs</td>
<td>1588 m</td>
</tr>
<tr>
<td>1464 m</td>
<td>1464 s</td>
</tr>
<tr>
<td>1418 vs</td>
<td>1420 s</td>
</tr>
<tr>
<td>1379 w</td>
<td>1381 w</td>
</tr>
<tr>
<td>1294 s</td>
<td>1302 w</td>
</tr>
<tr>
<td>1260 s</td>
<td>1261 m</td>
</tr>
<tr>
<td>1141 s</td>
<td>1148 s</td>
</tr>
<tr>
<td>1007 w</td>
<td>1009 m</td>
</tr>
<tr>
<td>822 m</td>
<td>822 s</td>
</tr>
<tr>
<td>811 sh</td>
<td>812 sh</td>
</tr>
<tr>
<td>648 s, br</td>
<td>632 sh, br</td>
</tr>
<tr>
<td>471 s</td>
<td>471 s</td>
</tr>
<tr>
<td>287 vw</td>
<td>287 vw</td>
</tr>
</tbody>
</table>

$^a$vs: very strong; s: strong; m: medium; w: weak; vw: very weak; br: broad; sh: shoulder. $^b$Bands that derive from the mono-deuterated amide group are not reported.
Table 3.2: Frequencies (cm\(^{-1}\)) and Assignments of Infrared and Raman Bands for Crystalline CH\(_3\)CH\(_2\)CONH\(_2\). Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>ν(_1)</th>
<th>Infrared</th>
<th>Raman</th>
<th>Calc.</th>
<th>δ(^a) (%)</th>
<th>PED(^b) (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3356</td>
<td>3356</td>
<td>3434</td>
<td>2.3</td>
<td>νCH (57), −νNH (43)</td>
<td></td>
</tr>
<tr>
<td>ν(_2)</td>
<td>3177</td>
<td>3171</td>
<td>3319</td>
<td>4.6</td>
<td>νNH (57), νNH (43)</td>
</tr>
<tr>
<td>ν(_3)</td>
<td>2979</td>
<td>2975</td>
<td>2906</td>
<td>2.4</td>
<td>νC1H (69), −νC1H (22)</td>
</tr>
<tr>
<td>ν(_4)</td>
<td>2979</td>
<td>2975</td>
<td>2894</td>
<td>2.8</td>
<td>νC1H (49), −νC1H (45)</td>
</tr>
<tr>
<td>2ν(_10)</td>
<td>2943</td>
<td>2940</td>
<td>–</td>
<td>–</td>
<td>νC2H (75), −νC2H (19)</td>
</tr>
<tr>
<td>ν(_5)</td>
<td>2922</td>
<td>2909</td>
<td>2866</td>
<td>1.7</td>
<td>νC2H (57), −νC2H (43)</td>
</tr>
<tr>
<td>ν(_6)</td>
<td>2882</td>
<td>2882</td>
<td>2830</td>
<td>1.8</td>
<td>νC1H (42), νC1H (32), νC2H (23)</td>
</tr>
<tr>
<td>ν(_7)</td>
<td>2811</td>
<td>2827</td>
<td>2825</td>
<td>0.2</td>
<td>νC2H (76), −νC2H (20)</td>
</tr>
<tr>
<td>2ν(_14)</td>
<td>2737</td>
<td>2734</td>
<td>–</td>
<td>–</td>
<td>νC2H (57), −νC2H (43)</td>
</tr>
<tr>
<td>ν(_8)</td>
<td>1643</td>
<td>1676</td>
<td>1667</td>
<td>0.5</td>
<td>νC=O (75), −νCN (7), δ(_s)NC(O)C (7)</td>
</tr>
<tr>
<td>ν(_9)</td>
<td>1628</td>
<td>1588</td>
<td>1562</td>
<td>2.9</td>
<td>σNH(_2) (86), νCN (10)</td>
</tr>
<tr>
<td>ν(_10)</td>
<td>1464</td>
<td>1464</td>
<td>1453</td>
<td>0.7</td>
<td>δ(_s)CH(_3) (43), −δ(_s)CH(_3) (38), ρ′CH(_3) (8)</td>
</tr>
<tr>
<td>ν(_11)</td>
<td>– 1450</td>
<td>1443</td>
<td>0.5</td>
<td>δ(_s)CH(_3) (51), δ(_as)CH(_3) (39), −ρ′CH(_3) (8)</td>
<td></td>
</tr>
<tr>
<td>ν(_12)</td>
<td>– 1430</td>
<td>1393</td>
<td>2.6</td>
<td>ωCH(_2) (30), νC2C6 (20), −νCN (19), −βC=O (10), −νC1C2 (7), δ(_as)CH(_3) (6)</td>
<td></td>
</tr>
<tr>
<td>ν(_13)</td>
<td>1418</td>
<td>1420</td>
<td>1422</td>
<td>0.2</td>
<td>σCH(_2) (89)</td>
</tr>
<tr>
<td>ν(_14)</td>
<td>1379</td>
<td>1381</td>
<td>1374</td>
<td>0.4</td>
<td>δ(_s)CH(_3) (89), νC1C2 (6)</td>
</tr>
<tr>
<td>ν(_15)</td>
<td>1294</td>
<td>1302</td>
<td>1278</td>
<td>1.6</td>
<td>ωCH(_2) (31), νCN (28), −τCH(_2) (12), βC=O (8), νNH(_2) (5)</td>
</tr>
<tr>
<td>ν(_16)</td>
<td>– 1260</td>
<td>1271</td>
<td>0.9</td>
<td>τCH(_2) (61), ωCH(_2) (13), −ρ′CH(_3) (7), −ρCH(_2) (6)</td>
<td></td>
</tr>
<tr>
<td>ν(_17)</td>
<td>1141</td>
<td>1148</td>
<td>1130</td>
<td>1.3</td>
<td>ρNH(_2) (30), −νC1C2 (19), ρ′CH(_3) (11), −ρCH(_3) (10), −νC=O (8), δ(_s)NC(O)C (7), δCCC (7)</td>
</tr>
<tr>
<td>ν(_18)</td>
<td>1068</td>
<td>1070</td>
<td>1108</td>
<td>3.6</td>
<td>ρCH(_2) (27), −ρCH(_3) (21), −τCH(_2) (17), −ρCH(_3) (16), −ΠC=O (11)</td>
</tr>
<tr>
<td>ν(_19)</td>
<td>1068</td>
<td>1070</td>
<td>1090</td>
<td>1.9</td>
<td>νC1C2 (38), ρNH(_2) (26), −νCN (16), ρCH(_3) (6)</td>
</tr>
<tr>
<td>ν(_20)</td>
<td>1007</td>
<td>1009</td>
<td>1028</td>
<td>2.0</td>
<td>ρ′CH(_3) (28), ωCH(_2) (19), νC1C2 (18), −ρCH(_3) (11), −νC2C6 (9), −ρCH(_2) (8)</td>
</tr>
<tr>
<td>ν(_21)</td>
<td>822</td>
<td>822</td>
<td>853</td>
<td>3.8</td>
<td>ρCH(_2) (23), ρ′CH(_3) (18), −ΠC=O (18), ρCH(_3) (15), τCH(_2) (10), νC2C6 (8)</td>
</tr>
<tr>
<td>ν(_22)</td>
<td>811</td>
<td>812</td>
<td>850</td>
<td>4.7</td>
<td>νC2C6 (43), −ρCH(_3) (14), ρNH(_2) (11), −ρCH(_2) (7), νC1C2 (6)</td>
</tr>
</tbody>
</table>

\(^a\)δ = |ν\(_{obs}\) − ν\(_{calc}\)|/ν\(_{obs}\) × 100%. \(^b\)ν: stretch; δ\(_s\): sym deformation; σ: scissoring; δ\(_as\): asym deformation; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting; Π: out-of-plane bending. \(^c\)Frequency obtained from 229 nm excitation UVRR data.
Table 3.3: Frequencies (cm$^{-1}$) and Assignments of Infrared and Raman Bands for Crystalline $\text{CH}_3\text{CH}_2\text{COND}_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>Infrared</th>
<th>Raman</th>
<th>Calc.</th>
<th>$\delta^a$ (%)</th>
<th>PED$^b$ (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_1$</td>
<td>2981</td>
<td>2977</td>
<td>2941</td>
<td>1.3</td>
<td>$\nu\text{C1H}$ (69), $-\nu\text{C1H}$ (22)</td>
</tr>
<tr>
<td>$\nu_2$</td>
<td>2981</td>
<td>2977</td>
<td>2929</td>
<td>1.7</td>
<td>$\nu\text{C1H}$ (49), $-\nu\text{C1H}$ (45)</td>
</tr>
<tr>
<td>$2\nu_1$</td>
<td>2943</td>
<td>2941</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\nu_3$</td>
<td>2923</td>
<td>2912</td>
<td>2900</td>
<td>0.6</td>
<td>$\nu\text{C2H}$ (75), $-\nu\text{C2H}$ (19)</td>
</tr>
<tr>
<td>$\nu_4$</td>
<td>2884</td>
<td>2886</td>
<td>2863</td>
<td>0.8</td>
<td>$\nu\text{C1H}$ (42), $\nu\text{C1H}$ (32), $\nu\text{C1H}$ (23)</td>
</tr>
<tr>
<td>$\nu_5$</td>
<td>–</td>
<td>2833</td>
<td>2858</td>
<td>0.9</td>
<td>$\nu\text{C2H}$ (76), $\nu\text{C2H}$ (20)</td>
</tr>
<tr>
<td>$2\nu_2$</td>
<td>–</td>
<td>2734</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\nu_6$</td>
<td>2527</td>
<td>2523</td>
<td>2607</td>
<td>3.3</td>
<td>$\nu\text{ND}$ (54), $-\nu\text{ND}$ (45)</td>
</tr>
<tr>
<td>$\nu_7$</td>
<td>2407</td>
<td>2393</td>
<td>2458</td>
<td>2.4</td>
<td>$\nu\text{ND}$ (54), $\nu\text{ND}$ (45)</td>
</tr>
<tr>
<td>$\nu_8$</td>
<td>1623</td>
<td>1610</td>
<td>1662</td>
<td>2.8</td>
<td>$\nu\text{C=O}$ (78), $\delta\text{NC(O)C}$ (7), $-\nu\text{CN}$ (7)</td>
</tr>
<tr>
<td>$\nu_9$</td>
<td>1466</td>
<td>1464</td>
<td>1447</td>
<td>1.2</td>
<td>$\delta\text{as}_\text{CH}<em>3$ (42), $-\delta\text{as}</em>\text{CH}_3$ (37), $\rho\text{CH}_3$ (8)</td>
</tr>
<tr>
<td>$\nu_{10}$</td>
<td>–</td>
<td>1450</td>
<td>1437</td>
<td>0.9</td>
<td>$\delta\text{as}_\text{CH}<em>3$ (52), $-\delta\text{as}</em>\text{CH}_3$ (38), $-\rho\text{CH}_3$ (8)</td>
</tr>
<tr>
<td>$\nu_{11}$</td>
<td>–</td>
<td>1434$^c$</td>
<td>1401</td>
<td>2.3</td>
<td>$\nu\text{CN}$ (33), $-\omega\text{CH}<em>2$ (19), $-\nu\text{C2C6}$ (19), $\beta\text{C=O}$ (10), $-\delta\text{as}</em>\text{CH}_3$ (6), $\delta\text{ND}_2$ (5)</td>
</tr>
<tr>
<td>$\nu_{12}$</td>
<td>1425</td>
<td>1422</td>
<td>1415</td>
<td>0.6</td>
<td>$\sigma\text{CH}_2$ (89)</td>
</tr>
<tr>
<td>$\nu_{13}$</td>
<td>1379</td>
<td>1378</td>
<td>1367</td>
<td>0.8</td>
<td>$\delta\text{CH}_3$ (89), $\nu\text{C1C2}$ (8)</td>
</tr>
<tr>
<td>$\nu_{14}$</td>
<td>1318</td>
<td>1317</td>
<td>1295</td>
<td>1.7</td>
<td>$\omega\text{CH}_2$ (45), $\nu\text{CN}$ (23), $\delta\text{ND}_2$ (15),</td>
</tr>
<tr>
<td>$\nu_{15}$</td>
<td>–</td>
<td>1261</td>
<td>1261</td>
<td>0.0</td>
<td>$\tau\text{CH}_2$ (75), $-\rho\text{CH}_3$ (9), $-\rho\text{CH}_3$ (5)</td>
</tr>
<tr>
<td>$\nu_{16}$</td>
<td>1168</td>
<td>1176</td>
<td>1134</td>
<td>3.2</td>
<td>$\delta\text{ND}_2$ (60), $\beta\text{C=O}$ (12), $\nu\text{C2C6}$ (8), $-\omega\text{CH}_2$ (5)</td>
</tr>
<tr>
<td>$\nu_{17}$</td>
<td>1079</td>
<td>1079</td>
<td>1093</td>
<td>1.3</td>
<td>$\nu\text{C1C2}$ (35), $\rho\text{CH}_3$ (19), $-\rho\text{CH}_3$ (11), $-\delta\text{CCC}$ (7), $-\delta\text{NC(O)C}$ (5)</td>
</tr>
<tr>
<td>$\nu_{18}$</td>
<td>1079</td>
<td>1079</td>
<td>1091</td>
<td>1.1</td>
<td>$\rho\text{CH}_2$ (25), $-\rho\text{CH}_3$ (19), $\rho\text{CH}_3$ (17), $-\tau\text{CH}_2$ (16), $-\Pi\text{C}=$O (11), $\nu\text{C1C2}$ (7)</td>
</tr>
<tr>
<td>$\nu_{19}$</td>
<td>1007</td>
<td>1006</td>
<td>1019</td>
<td>1.2</td>
<td>$\nu\text{C1C2}$ (33), $\rho\text{CH}_3$ (20), $\omega\text{CH}_2$ (14), $-\nu\text{C2C6}$ str (9), $-\rho\text{CH}_3$ (6), $\delta\text{ND}_2$ (5)</td>
</tr>
<tr>
<td>–</td>
<td>989</td>
<td>992</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\nu_{20}$</td>
<td>940</td>
<td>942</td>
<td>936</td>
<td>0.5</td>
<td>$\rho\text{ND}_2$ (34), $-\nu\text{CN}$ (15), $-\rho\text{CH}_3$ (11), $\delta\text{NC(O)C}$ (9), $\delta\text{ND}_2$ (8), $\rho\text{CH}_3$ (6)</td>
</tr>
<tr>
<td>–</td>
<td>819</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\nu_{21}$</td>
<td>807</td>
<td>806</td>
<td>828</td>
<td>2.7</td>
<td>$\rho\text{CH}_2$ (29), $\rho\text{CH}_3$ (24), $-\Pi\text{C}=$O (19), $\rho\text{CH}_3$ (13), $\tau\text{CH}_2$ (12)</td>
</tr>
<tr>
<td>$\nu_{22}$</td>
<td>–</td>
<td>770</td>
<td>776</td>
<td>0.7</td>
<td>$\nu\text{C2C6}$ (51), $\rho\text{ND}_2$ (25), $-\delta\text{CCC}$ (7)</td>
</tr>
</tbody>
</table>

$^a\delta = \left| \nu_{\text{obs}} - \nu_{\text{calc}} \right| / \nu_{\text{obs}} \times 100\%$. $^b\nu$: stretch; $\delta$: sym deformation; $\sigma$: scissoring; $\delta_{as}$: asym deformation; $\rho$: rocking; $\omega$: wagging; $\beta$: in-plane bending; $\tau$: twisting; II: out-of-plane bending. $^c$ Frequency obtained from 229 nm excitation UVRR data.
Figure 3.2: (a, b) Infrared, (c, d) non-resonance Raman, and (e, f) UVRR spectra of crystalline CH$_3$CH$_2$CONH$_2$ and CH$_3$CH$_2$COND$_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

is expected to be expanded along this coordinate [147]. Table 3.2 and Table 3.3 list the vibrational assignments, calculated, scaled frequencies, and PEDs for CH$_3$CH$_2$CONH$_2$ and CH$_3$CH$_2$COND$_2$, respectively.

3.4.1.2 2000-3500 cm$^{-1}$ Region  The high frequency region is dominated by N–H and C–H stretching bands. The N–H stretches show broad, strong peaks in both the infrared and Raman spectra that substantially downshift in frequency upon N-deuteration. The NH$_2$ asymmetric stretching band is located at $\sim$3355 cm$^{-1}$, but downshifts to $\sim$2525 cm$^{-1}$ upon N-deuteration. The NH$_2$ symmetric stretching band appears at $\sim$3175 cm$^{-1}$, while its ND$_2$ counterpart appears at $\sim$2400 cm$^{-1}$. The $\sim$3320 cm$^{-1}$ feature observed in the CH$_3$CH$_2$COND$_2$ spectra derives from an N–H stretching vibration for mono-N-deutered propanamide.

The bands located between 2700–3100 cm$^{-1}$ are insensitive to N-deuteration, which indicates that they are primarily C–H stretching modes. The $\sim$2975 cm$^{-1}$ and $\sim$2880 cm$^{-1}$
bands derive from CH$_3$ asymmetric and symmetric stretching modes, respectively. The CH$_2$ asymmetric stretching mode appears at $\sim 2910$ cm$^{-1}$ in the Raman, and at $\sim 2920$ cm$^{-1}$ in the infrared spectra. In contrast, the CH$_2$ symmetric stretching band appears at $\sim 2830$ cm$^{-1}$ and $\sim 2810$ cm$^{-1}$ in the Raman and infrared spectra, respectively.

The bands at $\sim 2735$ cm$^{-1}$ and $\sim 2940$ cm$^{-1}$ cannot be assigned to fundamentals. Neither band shifts upon N-deuteration, and according to Kuroda et al.’s [143] data, these bands also do not shift upon deuteration of the methylene group. This indicates that they do not derive from overtone or combination of amide or methylene bands. Therefore, we assign the $\sim 2735$ cm$^{-1}$ band to the first overtone of the the $\sim 1380$ cm$^{-1}$ CH$_3$ symmetric deformation vibration. The $\sim 2940$ cm$^{-1}$ band is strong in both the non-resonance Raman and UVRR spectra, but is of only moderate intensity in the infrared. We assign this band to a Fermi resonance between the first overtone of the $\sim 1460$ cm$^{-1}$ CH$_3$ asymmetric deformation and the CH$_3$ symmetric stretching fundamental based on the suggestions of Kuroda et al. [143] and Nolin and Jones [148]. This assignment disagrees with Nandini and Sathyanarayana [144], who attribute this band to a CH$_2$ asymmetric stretching mode. We disagree with the Nandini and Sathyanarayana [144] assignment since this band is upshifted beyond the typical frequency range ($\sim 2910$–$2930$ cm$^{-1}$) observed for CH$_2$ asymmetric stretching vibrations [149, 150].

### 3.4.1.3 1500-1800 cm$^{-1}$ Region
In this region there are two vibrations that involve the primary amide group, the Amide I (Am I) and Amide II (Am II) vibrations. The Am I at $\sim 1640$ cm$^{-1}$ and the Am II at $\sim 1620$ cm$^{-1}$ are strong and overlap in the infrared spectrum. In contrast, in the non-resonance Raman spectrum they are well-resolved, and show up as a moderately weak band located at $\sim 1675$ cm$^{-1}$ (Am I), and a stronger band at $\sim 1590$ cm$^{-1}$ (Am II). In the UVRRR spectrum, the Am I band shows a $\sim 1675$ cm$^{-1}$ peak followed by an overlapping $\sim 1640$ cm$^{-1}$ feature. The Am II band is strong and occurs at $\sim 1590$ cm$^{-1}$.

The large frequency differences between the infrared and Raman bands for the Am I and Am II vibrations presumably derive from the coupling of molecular vibrations within the crystal lattice into phonons. The spectral frequency differences for the infrared, normal Raman, and resonance Raman spectra derive from their differing selection rules for the
different phonon modes with different phasings of relative molecular motion.

Kuroda et al. [143] indicated that C=O stretching and NH\(_2\) scissoring motions are both important to the PEDs of the Am I and Am II vibrations. In contrast, Nandini and Sathyanarayana’s [144] normal mode analysis, as well as other studies on acetamide [151–153], indicate that the Am I vibration is mainly C=O stretching and the Am II is mainly NH\(_2\) bending. Our analysis concludes that the Am I mode consists mostly of C=O stretching (\(\sim75\%\)), with minor C–N stretching and NC(O)C in-plane bending (\(\sim7\%\) each) components, while the Am II mode is essentially pure NH\(_2\) scissoring (\(\sim86\%\)) with a small C–N stretching component (\(\sim10\%\)).

The spectral changes that are observed upon N-deuteration are consistent with our normal mode analysis of the Am I and Am II bands. The Am II band completely disappears, and a new band, which derives from ND\(_2\) scissoring, appears at \(\sim1170\,\text{cm}^{-1}\), supporting the notion that this mode is essentially pure NH\(_2\) scissoring. In contrast, N-deuteration results in the Am I mode downshifting to \(\sim1610\,\text{cm}^{-1}\) in the non-resonance Raman and UVRR spectra, and to \(\sim1620\,\text{cm}^{-1}\) in the infrared spectrum. This behavior is similar to the Am I band in secondary amides, and strongly supports the idea that this mode is predominately C=O stretching.

3.4.1.4 1200-1500 cm\(^{-1}\) Region Most bands in this region are easily assigned to CH\(_3\) or CH\(_2\) deformations and bending vibrations. The bands at \(\sim1464\,\text{cm}^{-1}\) and \(\sim1450\,\text{cm}^{-1}\) derive from CH\(_3\) asymmetric deformations, while the \(\sim1380\,\text{cm}^{-1}\) band is assigned to the CH\(_3\) symmetric deformation. In the non-resonance Raman and UVRR spectra, there is a weak band at \(\sim1260\,\text{cm}^{-1}\) that is assigned to a CH\(_2\) twisting mode.

3.4.1.5 C–N Stretching Modes Our assignments of the remaining bands observed in the 1200–1500 cm\(^{-1}\) region differ from Kuroda et al. [143] and Nandini and Sathyanarayana [144]. Kuroda et al. [143] previously assigned the \(\sim1420\,\text{cm}^{-1}\) band to two fundamentals, a CH\(_2\) scissoring mode and a C–N stretching mode. Nandini and Sathyanarayana [144] also assigned this band to two fundamentals, viz., a CH\(_2\) bending mode and a CH\(_3\) symmetric bending vibration. Nandini and Sathyanarayana [144] assigned a \(\sim1300\,\text{cm}^{-1}\) band to a
vibration that consists of \( \sim 30\% \) C–N stretching and CH\(_2\) wagging. They conflate this vibration with the C–N stretching mode that Kuroda et al. [143] assigned to the \( \sim 1420 \text{ cm}^{-1} \) band. In contrast, Kuroda et al. [143] assigns the \( \sim 1300 \text{ cm}^{-1} \) band to an almost pure CH\(_2\) wagging vibration.

Our normal mode analysis and UVRR data lead to very different assignments of these two vibrations. In the UVRR spectrum of CH\(_3\)CH\(_2\)CONH\(_2\), there is a very intense band at \( \sim 1430 \text{ cm}^{-1} \), which we assign to the \( \nu_{12} \) fundamental (Table 3.2) since it contains significant C–N stretching character. This band is not apparent in Kuroda et al.’s [143] or our non-resonance Raman spectra of CH\(_3\)CH\(_2\)CONH\(_2\). The assignment of this resonance enhanced \( \sim 1430 \text{ cm}^{-1} \) mode allows us to unambiguously assign the \( \sim 1420 \text{ cm}^{-1} \) band to a CH\(_2\) scissoring vibration. The \( \sim 1300 \text{ cm}^{-1} \) band, which appears weak in both the non-resonance Raman and UVRR spectra, is assigned to the \( \nu_{15} \) vibration, which appears to be related to the \( \nu_{12} \) mode.

Our normal mode analysis shows that the \( \nu_{12} \) and \( \nu_{15} \) amide vibrations contain significant C–N stretching. The largest \( \nu_{12} \) PED components are CH\(_2\) wagging (\( \sim 30\% \)), C–CH\(_2\) stretching (\( \sim 20\% \)), C–N stretching (\( \sim 19\% \)), and C=O in-plane bending (\( \sim 10\% \)). For \( \nu_{15} \), the major PED components are CH\(_2\) wagging and C–N stretching (\( \sim 30\% \) each), followed by CH\(_2\) twisting (\( \sim 12\% \)) and C=O in-plane bending (\( \sim 8\% \)).

### 3.4.1.6 1000-1200 cm\(^{-1}\) Region

In this spectral region, we expect NH\(_2\) rocking, CH\(_2\) rocking, CH\(_3\) rocking, and C–CH\(_3\) stretching vibrations. The \( \sim 1141 \text{ cm}^{-1} \) infrared and \( \sim 1148 \text{ cm}^{-1} \) Raman bands are easily assigned to the NH\(_2\) rocking vibration, since they downshift to \( \sim 940 \text{ cm}^{-1} \) upon N-deuteration. The C–CH\(_3\) (\( \nu_{19} \)) stretching vibration was previously assigned [143], but our normal mode calculation suggests that this mode is more complicated and contains significant contributions of NH\(_2\) rocking (\( \sim 26\% \)) and C–N stretching (\( \sim 16\% \)). The infrared and Raman spectra show only two bands located at \( \sim 1070 \text{ cm}^{-1} \) and \( \sim 1010 \text{ cm}^{-1} \). We assign the \( \sim 1070 \text{ cm}^{-1} \) band to both the CH\(_2\) rocking (\( \nu_{18} \)) and C–CH\(_3\) stretching (\( \nu_{19} \)) modes. In contrast, the \( \sim 1010 \text{ cm}^{-1} \) band shows a negligible change in frequency upon N-deuteration, and is thus assigned to the CH\(_3\) rocking mode.
3.4.1.7  <1000 cm\(^{-1}\) Region  The region below 1000 cm\(^{-1}\) is dominated mainly by torsional motions, as well as extensively coupled skeletal stretching and deformation modes. This region is difficult to unambiguously assign, especially below 700 cm\(^{-1}\). We assign the two bands at \(\sim 820\) cm\(^{-1}\) and 810 cm\(^{-1}\) to CH\(_2\) rocking and C–CH\(_3\) stretching fundamentals, respectively. The C–CH\(_3\) stretching mode shows a modest contribution of NH\(_2\) rocking, which likely accounts for its \(\sim 42\) cm\(^{-1}\) downshift to \(\sim 770\) cm\(^{-1}\) upon N-deuteration.

3.4.2  Solvation Dependence of UVRR Bands

Figure 3.3: UVRR (204 nm excitation) spectra of propanamide (a) in H\(_2\)O and (b) in D\(_2\)O. The contribution of solvent was subtracted from both spectra. For (a) 200 mM NaClO\(_4\) was used as an internal standard. The contribution of the ClO\(_4\)\(^-\) stretching band was also subtracted. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

3.4.2.1  UVRR of Propanamide in Aqueous Solutions  We measured the \(\sim 204\) nm excited UVRR spectra of propanamide in mixtures of acetonitrile and water in order to determine the effects of solvation on the primary amide vibrations. The spectra of propanamide in aqueous solutions (Figure 3.3) differ from that of crystalline propanamide. Compared to
Table 3.4: Measured Frequencies (cm$^{-1}$) and Cross Sections ($\sigma$, mbarns-molec.$^{-1}$sr$^{-1}$) of UVRR bands in CH3CH2CONH$_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>vibration</th>
<th>crystal$^a$</th>
<th>water$^b$</th>
<th>acetonitrile$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_8$ (Am I)</td>
<td>1676</td>
<td>1669±1</td>
<td>3.5±0.70</td>
</tr>
<tr>
<td>$\nu_9$ (Am II)</td>
<td>1588</td>
<td>1610±1</td>
<td>18±1.1</td>
</tr>
<tr>
<td>$\nu_{12}$</td>
<td>1430</td>
<td>1428±1</td>
<td>30±0.90</td>
</tr>
<tr>
<td>$\nu_{17}$</td>
<td>1148</td>
<td>1132±1</td>
<td>8.8±0.24</td>
</tr>
<tr>
<td>$\nu_{19}$</td>
<td>1070</td>
<td>1069±1</td>
<td>6.0±0.14</td>
</tr>
</tbody>
</table>

$^a$ UVRR cross sections were not calculated due to the lack of an internal standard. $^b$Values measured from n = 4 sample size. $^c$Values measured from n = 1 sample size.

crystalline spectra, propanamide in H$_2$O (Table 3.4) shows an Am I ($\nu_8$) band that downshifts $\sim$7 cm$^{-1}$ and appears as a shoulder, while the Am II ($\nu_9$) band upshifts $\sim$22 cm$^{-1}$. The $\nu_{12}$ and C–CH$_3$ stretching ($\nu_{19}$) bands do not change frequency, while the NH$_2$ rocking ($\nu_{17}$) mode downshifts $\sim$16 cm$^{-1}$.

The bands of propanamide in D$_2$O (Figure 3.3b) also show significant changes compared to their N-deuterated crystal spectra. The Am I'($\nu_8$) band is broad and is located at $\sim$1633 cm$^{-1}$. The Am II band disappears, and the ND$_2$ ($\nu_{16}$) scissoring band occurs at $\sim$1168 cm$^{-1}$, while the $\nu_{11}$ band is at $\sim$1443 cm$^{-1}$.

3.4.2.2 Effect of Solvation on UVRR Spectra Figure 3.4 shows the dramatic effect of solvation on the UVRR bands of primary amides. The spectra show that all bands, except for the Am I, increase their Raman cross sections as the mole fraction of H$_2$O increases (Table 3.4). The Am II and $\nu_{12}$ bands show the largest cross section increases ($\sim$3–4-fold), while the NH$_2$ rocking ($\nu_{17}$) and the C–CH$_3$ stretching ($\nu_{19}$) bands increase $\sim$2–3-fold. The Am I band is the only band whose Raman cross section decreases as the mole fraction of H$_2$O increases.

Figure 3.5 shows that there is a roughly linear cross section and frequency increase for most bands. The $\nu_{12}$ band shows a $\sim$29 cm$^{-1}$ per mole fraction H$_2$O frequency increase, the NH$_2$ rocking ($\nu_{17}$) band shows a $\sim$14 cm$^{-1}$ per mole fraction H$_2$O frequency increase, while
the C−CH₃ stretching (ν₁₉) vibration shows a ∼19 cm⁻¹ per mole fraction H₂O frequency increase. In contrast, the Am I band shows a ∼22 cm⁻¹ per mole fraction H₂O frequency decrease. The Am II band shows only a modest frequency decrease with increasing water mole fractions.

The dependence of the UVRR spectra of propanamide on the mole fraction of water shown in Figure 3.4 is very similar to that observed in valeramide [125] and NMA [98, 154–157]. To understand this behavior, we compared the UV absorption spectra of propanamide and NMA in acetonitrile and in H₂O (Figure 3.6). The molar absorptivities of the ∼200 nm NV₁ transitions increase for both propanamide and NMA as the solvent transfers from acetonitrile to H₂O. For NMA, the absorption peak maximum of the NV₁ transition redshifts going from acetonitrile to water. This is less clearly evident in the case of propanamide where the NV₁ absorption maximum lies deeper in the UV at ∼180 nm. This trend is expected from Nielsen and Schellman’s [158] results. They observe redshifts in the absorption maxima of several primary and secondary amides going from cyclohexane to water. This increase in molar absorptivity of the NV₁ transition, upon aqueous solvation, is in part responsible for increasing the UVRR cross sections due to the fact that the Raman scattering cross section is proportional to the square of the molar absorptivity.

However, most of the cross section increase results from changes in the ground state structure. The effect of aqueous solvation on the UVRR secondary amide band intensities and frequencies has been traditionally rationalized by considering the effects of the solvent dielectric and direct hydrogen bonding on the amide group resonance structures [98, 154–157]. We can invoke a similar argument for primary amides.

In low dielectric constant and hydrogen bonding environments, the O=C−NH₂ resonance form is typically dominant over the ‘O−C=NH⁺₂ structure in the propanamide electronic ground state (Scheme 3.1). The dipolar resonance structure becomes more favorable in water due to the high dielectric constant and the stabilizing hydrogen bonding to propanamide’s C=O and NH₂ groups. These two effects increase the C−N bond order and decrease the C−O bond order of the primary amide group in the electronic ground state.

It is also important to note that the electronic excited state can also be impacted by solvation effects. For example, in NMA, Hudson and Markham [156] argued that their
Figure 3.4: UVRR (204 nm excitation) spectra of propanamide (30 mM) in different acetonitrile and water mixtures. Arrows show frequency and intensity trends of UVRR bands as the fraction of H$_2$O increases. NaClO$_4^-$ (0.2 M) was used as an internal intensity and calibration standard in the solutions. Spectra were normalized to the integrated area of the 932 cm$^{-1}$ ClO$_4^-$ stretching band. The spectral contributions of acetonitrile, ClO$_4^-$, and water were subtracted. The asterisk indicates a spectral feature that is an artifact of subtracting out the $\sim$1376 cm$^{-1}$ acetonitrile band. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

$ab$ $initio$ post-Hartree Fock calculations indicate that the effects of hydrogen bonding due to solvation are greater in magnitude on the $\pi \rightarrow \pi^*$ electronic excited state equilibrium geometry than for the ground state geometry. They argue that the changes in the C=O and C–N bond lengths in the excited state due to solvation were also in the opposite direction of the excited state bond length changes of unsolvated NMA; i.e., the C–N bond length is larger and the C=O bond length is smaller in the excited state for a NMA(H$_2$O)$_3$ cluster compared to an isolated NMA molecule.

The changes in C–O and C–N bond orders (and bond lengths) of the amide group
going from acetonitrile to water profoundly effects the resonance Raman cross sections and the band frequencies. This is because resonance Raman cross sections scale with the square of the displacement along the enhanced vibrational normal coordinate between the equilibrium geometries of the electronic ground and excited states \[159\]. For example, the most resonance enhanced UVRR bands of NMA in water involve C–N stretching \[98, 154, 155, 160, 161\] because of the large expansion of the electronic excited state along the C–N bond \[156, 157, 161\]. In contrast, there is a relatively small enhancement of the Am I band for NMA in water because the excited state expansion along the C=O bond is much less.

In the case of propanamide, the dramatic spectral changes observed in Figure 3.4 can likewise be explained by changes in C=O and C–N bond lengths upon solvation. The elongation of the C=O bond results in a decrease of the C=O stretching force constant. This results in a downshift in the Am I band frequency. The Am I band UVRR cross section
Figure 3.6: UV absorption spectra of (a) NMA and (b) propanamide in acetonitrile and water. Arrows show trend of molar absorptivity as the fraction of water increases. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

Scheme 3.1: Resonance Structures of Propanamide. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.
also decreases because the magnitude of the displacement between the electronic ground and excited states along the C–O coordinate must decrease in water since the ground state C=O bond length elongates.

In contrast, for vibrations with significant C–N stretching, the contraction of the C–N bond length results in a vibrational frequency upshift due to the increase in the stretching force constant. This bond contraction also increases the magnitude of the displacement between the electronic ground and excited states along the C–N coordinate, which results in an increase in the Raman cross sections. It will be more difficult to explain the origin of the UVRR cross section increases for heavily coupled modes, such as the $\sim$1430 cm$^{-1} \nu_{12}$ and the $\sim$1300 cm$^{-1} \nu_{15}$ modes, which contain displacements of multiple atoms. The displacements of all of the primary amide atoms contribute to resonance enhancement. Phasing of this motion can be very important [161]. However, we conclude that the C–N stretching motion in these vibrations is predominantly responsible for their UVRR intensity enhancements.

3.5 CONCLUSION

We utilized DFT calculations, infrared, non-resonance Raman, and UVRR spectra to assign the vibrational bands of crystalline propanamide. Our study resolves previous inconsistencies in the vibrational assignments and normal mode compositions of primary amide bands. We also studied the effect of aqueous solvation on the primary amide UVRR bands by examining the $\sim$204 nm UVRR dependence as the solvent transfers from acetonitrile to water. The aqueous solvation dependence of primary amide UVRR bands can be rationalized by the stabilization between dipolar resonance structures of the ground electronic state of the amide group. Both hydrogen bonding interactions and the increased dielectric constant as the solvent transfers from acetonitrile to water contribute to the stabilization of the dipolar resonance structure in the ground state, which effectively increases the C–N bond order while decreasing the C–O bond order.

The resulting increased displacement between electronic ground and excited state geometries along the C–N coordinate increases the UVRR cross sections of vibrations that contain
significant C–N stretching. In contrast, the decreased displacement between the electronic ground and excited state geometries along the C–O coordinate results in a dramatic decrease in the Am I band UVRR cross section. These results indicate that the Am I, Am II, and $\nu_{12}$ band UVRR cross sections and their frequencies can be used as sensitive spectroscopic markers for hydrogen bonding and local dielectric environment of the side chains of Gln and Asn.

3.6 ACKNOWLEDGMENT

Funding for this work was provided by the University of Pittsburgh. E.M.D. gratefully acknowledges support through the NIH Molecular Biophysics and Structural Biology Training Grant (T32 GM88119-3). The computational work was supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources provided.
4.0 GLUTAMINE AND ASPARAGINE SIDE CHAIN HYPERCONJUGATION INDUCED STRUCTURALLY SENSITIVE VIBRATIONS


Author Contributions: D.P. acquired and analyzed the UVRR data with the assistance of R.S.J. and E.M.D. D.P. and S.G. acquired and analyzed X-ray diffraction data. Z.H. performed and analyzed the results of the DFT calculations with the assistance of D.P. The manuscript was prepared by D.P. and S.A.A. with the assistance of R.S.J. and E.M.D.

We identified vibrational spectral marker bands that sensitively report on the side chain structures of Gln and Asn. DFT calculations indicate that the Amide III$^P$ (Am III$^P$) vibrations of Gln and Asn depend cosinusoidally on their side chain OCCC dihedral angles (the $\chi_3$ and $\chi_2$ angles of Gln and Asn, respectively). We use UVRR and visible Raman spectroscopy to experimentally correlate the Am III$^P$ Raman band frequency to the primary amide OCCC dihedral angle. The Am III$^P$ structural sensitivity derives from the Gln (Asn) $C_\beta$–$C_\gamma$ ($C_\alpha$–$C_\beta$) stretching component of the vibration. The $C_\beta$–$C_\gamma$ ($C_\alpha$–$C_\beta$) bond length inversely correlates with the Am III$^P$ band frequency. As the $C_\beta$–$C_\gamma$ ($C_\alpha$–$C_\beta$) bond length decreases, its stretching force constant increases, which results in an upshift in the
Am III$^\text{P}$ frequency. The $C_{\beta}$–$C_{\gamma}$ ($C_\alpha$–$C_{\beta}$) bond length dependence on the $\chi_3$ ($\chi_2$) dihedral angle results from hyperconjugation between the $C_\delta$–$O_\epsilon$ ($C_\gamma$–$O_\delta$) $\pi^*$ and $C_{\beta}$–$C_{\gamma}$ ($C_\alpha$–$C_{\beta}$) $\sigma$ orbitals. Using a protein data bank library, we show that the $\chi_3$ and $\chi_2$ dihedral angles of Gln and Asn depend on the peptide backbone Ramachandran angles. We demonstrate that the inhomogeneously broadened Am III$^\text{P}$ band line shapes can be used to calculate the $\chi_3$ and $\chi_2$ angle distributions of peptides. The spectral correlations determined in this study enable important new insights into protein structure in solution, and in Gln- and Asn-rich amyloid-like fibrils and prions.

4.1 INTRODUCTION

Amyloid-like fibril protein aggregates and prion proteins often contain stretches of Gln and Asn residues. For example, polyQ-rich fibrils are the pathological hallmarks of several “CAG” codon repeat diseases [1, 3, 4, 35, 45, 162–165]. Similarly, Sup35p and Ure2p prions contain Gln- and Asn-rich regions that drive their aggregation and cause loss-of-function of these normally soluble proteins [166].

Because Gln and Asn side chains can hydrogen bond to water, the peptide backbone, or other side chains, they serve unique roles in protein structure and conformational transitions. Unfortunately, there is relatively little known about the mechanisms by which the primary amide groups of Gln and Asn interact with other protein constituents, or what role they play in the aggregation of prions and fibrils. Consequently, it is important to find spectroscopic markers that can be used to monitor the conformations and hydrogen bonding environments of Asn and Gln side chains in order to develop a deeper understanding of the roles that these residues play in protein aggregation.

There are few methods to quantitatively examine the conformations of Gln and Asn side chains in prion and fibril aggregates. Recent solid-state NMR studies [56, 58, 59] suggest that there are at least two different populations of Gln side chain conformers in polyQ fibrils. Sharma et al. [55] claim on the basis of low-resolution X-ray fiber and powder diffraction data that the side chains in polyQ fibrils adopt an unusual bent conformation; however,
these highly uncommon side chain structures have not been substantiated by other studies.

High resolution X-ray diffraction studies [167, 168] on small peptide microcrystals that contain amyloidogenic sequences have revealed important, atomic-resolution details regarding the steric zipper interactions that could occur in Gln- and Asn-rich prions and fibrils. These studies indicate, for example, that differences in the structures and hydrogen bonding interactions of amino acid side chains give rise to different fibril polymorphs. However, the conformations observed in the small peptide crystals may not reflect the side chain structures and hydrogen bonding interactions that occur in bona fide prion and fibril aggregates.

UVRR is a powerful, emerging tool for studying the conformations of proteins, as well as the structure, local hydrogen bonding, and dielectric environments of amino acid side chains [89, 101–103, 123, 124, 146, 169–173]. Deep UV excitation (∼200 nm) selectively resonance enhances secondary and primary amide vibrations [125, 134]. Previous investigations of secondary amide vibrations have developed a detailed understanding of the UVRR spectral dependence on the peptide bond structure and its hydrogen bonding [98, 155, 156, 174]. For example, Asher and coworkers [96, 97, 105] quantitatively correlated the Amide III$^{3}$ (Am$^{3}$I$^{3}$) frequency to the peptide bond Ramachandran Ψ dihedral angle. They determined that the structural sensitivity of the Am III$^{3}$ vibration derives from coupling between the peptide backbone amide N–H and C$^{\alpha}$–H bending motions [96]. This fundamental insight enabled incisive investigations that elucidated, in detail, the mechanism of α-helix (un)folding in a wide range of solution environments [106, 107, 175–178].

We seek to develop a similar deep understanding of the UVRR spectral dependence of primary amide vibrations on the structure of Gln and Asn side chains. In this work, we discover the structural sensitivity of the Amide III$^{P}$ (Am$^{P}$I$^{III}$) vibration on the primary amide OCCC dihedral angle (the $\chi_3$ and $\chi_2$ angles of the side chains Gln and Asn, respectively). The potential energy distribution (PED) of this vibration in Gln (Asn) contains significant contributions of C$^{\beta}$–C$^{\gamma}$ (C$^{\alpha}$–C$^{\beta}$) stretching, N$^{\epsilon}$H$_2$ (N$^{\delta}$H$_2$) rocking, and C$^{\delta}$–N$^{\epsilon}$ (C$^{\gamma}$–N$^{\delta}$) stretching motions. We find that the structural sensitivity of the Am III$^{P}$ mode originates mainly from the C$^{\beta}$–C$^{\gamma}$ (C$^{\alpha}$–C$^{\beta}$) bond length dependence on the $\chi_3$ ($\chi_2$) dihedral angle. We demonstrate that the C$^{\beta}$–C$_{\gamma}$ (C$^{\alpha}$–C$^{\beta}$) bond length correlation on the $\chi_3$ ($\chi_2$) dihedral angle derives from hyperconjugation between the C$^{\beta}$–C$_{\gamma}$ (C$^{\alpha}$–C$^{\beta}$) σ orbital and the
We compare our results with the Gln and Asn entries of the Shapovalov and Dunbrack side chain rotamer library [179] and examine the dependence of Gln (Asn) $\chi_3$ ($\chi_2$) dihedral angles on the peptide backbone Ramachandran ($\Phi$, $\Psi$) angles. We observe distinct $\chi_3$ and $\chi_2$ dihedral angle preferences for Gln and Asn residues that adopt PPII, $\beta$-sheet, and $\alpha$-helix Ramachandran angles. Applying this new insight, and the dependence of the $\chi_3$ dihedral angle on the Am III$^P$ vibrational frequency, we determine the $\chi_3$ angle distribution of Q$_3$ and D$_2$Q$_{10}$K$_2$ peptides in aqueous solution. We find that Q$_3$ and D$_2$Q$_{10}$K$_2$ favor $\chi_3$ dihedral angles similar to those of Gln in solution. This result is consistent with Q$_3$ and D$_2$Q$_{10}$K$_2$ containing side chains that are completely solvated.

Our work here develops a novel spectral marker for experimentally probing the structures of Asn and Gln side chains in fibrils and prion aggregates. Our methodology does not require extensive isotopic labeling or crystallization and allows us to monitor the side chain structural changes that occur during protein aggregation. This enables crucial, molecular-level insights into the role that Gln and Asn side chains play in stabilizing fibril and prion aggregates. We are developing new insights into why Gln- and Asn-rich sequences have strong propensities to aggregate into amyloid-like fibrils and prions.

4.2 EXPERIMENTAL DETAILS

4.2.1 Materials

L-glutamine (L-Gln, $\geq$99% purity), L-glutamine t-butyl ester hydrochloride (GlnTBE, $\geq$98% purity), and glycyl-L-glutamine (Gly-Gln, $\geq$97% purity) were purchased from Sigma-Aldrich. D-glutamine (D-Gln, 98% purity) was purchased from Acros Organics, and N-acetyl-L-glutamine (NAcGln, 97% purity) was purchased from Spectrum Chemical Mfg. Corp. L-seryl-L-asparagine (Ser-Asn, $\geq$99% purity) was purchased from Bachem. Optima-grade H$_2$O was purchased from Fisher Scientific, and D$_2$O (99.9% atom D purity) was purchased from Cambridge Isotope Laboratories, Inc. Gln$_3$ was purchased from Pierce Biotech-
nology at 95% purity.

4.2.2 Sample Preparation

Gly-Gln and Ser-Asn were obtained as crystalline powders and used without further purification or re-crystallization. D-Gln, NAcGln, L-Asn, and GlnTBE crystals were prepared by drying saturated solutions in water. L-Gln crystals were obtained by drying a saturated solution in the presence of 0.1 M NaCl. N-deuterated crystals were prepared via multiple rounds of re-crystallization in D$_2$O. Samples of Gln$_3$ were prepared at 0.5 mg mL$^{-1}$ in HPLC-grade water containing 0.05 mM sodium perchlorate (Sigma Aldrich, ≥98% purity). The sodium perchlorate was used as an internal intensity standard to allow us to subtract the contribution of water.

4.2.3 X-ray Diffraction

X-ray diffraction of crystals was performed using a Bruker X8 Prospector Ultra equipped with a copper micro-focus tube ($\lambda = 1.54178$ Å). The crystals were mounted and placed in a cold stream of N$_2$ gas (230 K) for data collection. The frames collected on each crystal specimen were integrated with the Bruker SAINT software package using the narrow-frame algorithm. APPENDIX C discusses, in detail, the methods used to determine the unit cells and crystal structures of the compounds examined.

4.2.4 Visible Raman Spectroscopy

Visible excitation Raman spectra of crystals were collected using a Renishaw inVia spectrometer equipped with a research-grade Leica microscope. Spectra were collected using the 633 nm excitation line from a HeNe laser and a 5× objective lens. The spectrometer resolution was $\sim$2 cm$^{-1}$. The 918 cm$^{-1}$ and 1376 cm$^{-1}$ bands of acetonitrile [129] were used to calibrate the spectral frequencies.
4.2.5 UVRR Spectroscopy

UVRR spectra of crystals were collected using CW 229 nm light generated by an Innova 300C FreD frequency doubled Ar$^+$ laser [130]. The crystalline specimens were spun in a cylindrical brass cell to prevent the accumulation of thermal or photodegradation products. A SPEX triplemate spectrograph, modified for use in the deep UV, was utilized to disperse the Raman scattered light. A Spec-10 CCD camera (Princeton Instruments, Model 735-0001) with a Lumagen-E coating was used to detect the Raman light. The power of UV light illuminating the sample ranged from $\sim$1.5–2 mW. The 801 cm$^{-1}$, 1028 cm$^{-1}$, 2852 cm$^{-1}$, and the 2938 cm$^{-1}$ bands of cyclohexane were used to calibrate the 229 nm excitation UVRR spectral frequencies.

UVRR solution-state measurements were made using $\sim$204 nm nm excitation. The UV light was generated by Raman shifting the third harmonic of a Nd:YAG laser (Coherent, Inc.) with H$_2$ gas ($\sim$30 psi) and selecting the fifth anti-Stokes line. A thermostatted ($20^\circ$C) flow cell was employed to circulate solutions in order to prevent the contribution of photodegradation products. The scattered light was dispersed and imaged using a double monochromator, modified for use in the UV in a subtractive configuration [131], and detected with a Spec-10 CCD camera.

4.3 COMPUTATIONAL DETAILS

4.3.1 Density Functional Theory (DFT) Calculations

The DFT calculations [135] were carried out using the GAUSSIAN 09 package [136]. The geometry optimizations and frequency calculations were performed using the M06-2X functional [137] and the 6-311++g** basis set. The presence of water was simulated implicitly by employing a polarizable continuum dielectric model (PCM). Vibrational frequencies were calculated using the harmonic approximation. The calculated frequencies were not scaled. The potential energy distribution (PED) of each vibration was obtained from the GAUSSIAN output files by employing a MATLAB program that we wrote. Figure 4.1 shows
the DFT calculated minimum energy structure of L-Gln and the atomic labeling scheme. In order to study the conformational dependence of the Raman bands, we fixed the $\chi_3$ dihedral angle of L-Gln, re-optimized the geometry, and calculated the harmonic vibrational frequencies for a series of conformers with $\chi_3$ angles of $-16^\circ$, $0^\circ$, $4^\circ$, $\pm 30^\circ$, $\pm 60^\circ$, $\pm 90^\circ$, $\pm 120^\circ$, $\pm 150^\circ$, and $\pm 180^\circ$.

![Figure 4.1: Geometry of optimized structure and atomic labeling scheme of L-Gln used in DFT calculations and band assignments. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.](image)

Figure 4.1: Geometry of optimized structure and atomic labeling scheme of L-Gln used in DFT calculations and band assignments. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

### 4.4 RESULTS AND DISCUSSION

#### 4.4.1 Assignment of L-Gln UVRR Bands in H$_2$O and D$_2$O

Figure 4.2 shows the band-resolved $\sim$204 nm excitation UVRR spectra of L-Gln in H$_2$O and D$_2$O. Visible Raman and infrared spectra band assignments of L-Gln were reported previously by Ramirez and coworkers for both solid-state crystalline samples [181] and in the solution-state [182]. Ramirez and coworkers made assignments without performing normal mode calculations in their study of crystalline L-Gln. In their solution-state study, they employed DFT calculations to aid in band assignments. We found that their reported frequencies did not match those in our solution-state UVRR spectra, and that their band assignments were inconsistent with our intensity expectations of the resonance enhanced
bands.

In the work here, we perform a new normal mode analysis of L-Gln in order to assign our UVRR spectra. We employ DFT calculations that use a more modern functional (M06-2X) than that of Ramirez and coworkers. These assignments build off of our previous, detailed assignment of propanamide [146], a model for the side chains of Gln and Asn. Our assignments of L-Gln in H$_2$O and D$_2$O are shown in Table 4.1 and Table 4.2, respectively.

The UVRR spectra are dominated by bands that derive from vibrations of the primary amide group. This is because these resonance enhanced vibrations couple to the strong $\sim 180$ nm NV$_1$ electronic transition. These resonance enhanced amide bands contain significant contributions of C$_{\delta}$–N$_{\epsilon}$ stretching, since the electronic excited state is expanded along this coordinate [147].

The spectral region between 1600 cm$^{-1}$ and 1700 cm$^{-1}$ is dominated by two primary amide vibrations, the Amide I$^p$ (Am I$^p$) and Amide II$^p$ (Am II$^p$) bands. The superscript $^p$ denotes the primary amide to distinguish these vibrations from the widely known vibrations of secondary amides found in proteins. The Am I$^p$ band is located at $\sim 1680$ cm$^{-1}$ and derives mainly from C$_{\delta}$=O$_{\epsilon}$ stretching. In D$_2$O, the Am I$^p$ band (called the AmI$^{Pp}$) downshifts to $\sim 1650$ cm$^{-1}$. The Am II$^p$ band at $\sim 1620$ cm$^{-1}$ derives from a vibration whose PED contains mostly N$_{\epsilon}$H$_2$ scissoring ($\sim 86\%$) and C$_{\delta}$–N$_{\epsilon}$ stretching ($\sim 10\%$). Upon N-deuteration, the C$_{\delta}$–N$_{\epsilon}$ stretching and ND$_2$ scissoring motions decouple. This causes the Am II$^p$ band to disappear, and a new band, which derives from N$_{\epsilon}$D$_2$ scissoring, appears at $\sim 1160$ cm$^{-1}$.

The most intense features of the Gln spectra in Figure 4.2 occur in the region between 1400 cm$^{-1}$ to 1500 cm$^{-1}$. Most of the bands found in this region derive from CH$_2$ scissoring or wagging modes. However, we assign the most intense band, located at $\sim 1430$ cm$^{-1}$, to a vibration that contains significant contributions of CH$_2$ wagging, C$_{\gamma}$–C$_{\delta}$ stretching, CH$_2$ scissoring, and C$_{\delta}$–N$_{\epsilon}$ stretching in its PED. This assignment is based on our previous work with propanamide [146], which shows a similar intense band at $\sim 1430$ cm$^{-1}$.

The region between 1200 cm$^{-1}$ to 1400 cm$^{-1}$ contains bands that derive mostly from C$_{\alpha}$H rocking, CH$_2$ wagging, and CH$_2$ twisting modes. We assign the $\sim 1365$ cm$^{-1}$ and $\sim 1350$ cm$^{-1}$ bands in the Figure 4.2a spectrum to C$_{\alpha}$H rocking modes. We assign the strong bands located at $\sim 1330$ cm$^{-1}$, $\sim 1290$ cm$^{-1}$, and the very weak $\sim 1205$ cm$^{-1}$ bands to CH$_2$ twisting
Figure 4.2: UVRR spectra excited at \(~\sim\)204 nm of L-Gln in (a) H\(_2\)O and (b) D\(_2\)O. The spectral contributions of the solvents have been subtracted. The reduced \(\chi^2\) (\(\chi^2_{\text{red}}\)) statistics for the spectral fits shown in (a) and (b) are 0.55 and 1.8, respectively. Adapted with permission from [180]. Copyright \(\copyright\) (2015), American Chemical Society.

modes. The \(~\sim\)1265 cm\(^{-1}\) feature is assigned to a CH\(_2\) wagging vibration. Only two bands, at \(~\sim\)1370 cm\(^{-1}\) and \(~\sim\)1345 cm\(^{-1}\), appear in D\(_2\)O. We assign the \(~\sim\)1370 cm\(^{-1}\) band to a CH\(_2\) wagging mode and the \(~\sim\)1345 cm\(^{-1}\) band to a C\(_\alpha\)H rocking mode. We conclude that these vibrations appear strongly in the UVRR spectrum in Figure 4.2b because they contain significant C\(_\delta\)–N\(_\epsilon\) stretching.

The region between 1000 cm\(^{-1}\) and 1200 cm\(^{-1}\) contains bands that derive from vibrations with large C–C stretching, N\(_\epsilon\)H\(_2\) rocking, or NH\(_3\) rocking contributions. Most of the vibrations in this region are complex. We assign the \(~\sim\)1160 cm\(^{-1}\) band to a coupled C\(_\alpha\)H
Table 4.1: UVRR Frequencies (cm\(^{-1}\)) and Assignments of L-Gln in H\(_2\)O. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>Potential Energy Distribution(^a) (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1679</td>
<td>1745</td>
<td>(\nu)C(_3)O (75), (\nu)C(_3)N (8), (\nu)N(_2)C(_3)C(_7) (7)</td>
</tr>
<tr>
<td>1652</td>
<td>1666</td>
<td>(\delta)(_{as})NH(<em>3) (48), (\delta)(</em>{as})NH(_3) (44), (\rho)NH(_3) (5)</td>
</tr>
<tr>
<td>1620</td>
<td>1623</td>
<td>(\sigma)N(_2)H(_2) (86), (\nu)C(_3)N (10)</td>
</tr>
<tr>
<td></td>
<td>1621</td>
<td>(\delta)(_{as})NH(<em>3) (40), (\delta)(</em>{as})NH(_3) (48), (\delta)NH(_3) (10)</td>
</tr>
<tr>
<td>1585</td>
<td>1715</td>
<td>(\nu)CO (53), (\nu)CO (33), (\rho)C(_\alpha)C (7)</td>
</tr>
<tr>
<td>1464</td>
<td>1494</td>
<td>(\sigma)C(_\gamma)H(_2) (88)</td>
</tr>
<tr>
<td></td>
<td>1466</td>
<td>(\delta)(_{as})NH(<em>3) (33), (\sigma)C(</em>\gamma)H(<em>2) (30), (\omega)C(</em>\gamma)H(_2) (7), (\omega)C(_3)H(_2) (5), (\nu)C(<em>3)C(</em>\gamma) (5)</td>
</tr>
<tr>
<td>1447</td>
<td>1460</td>
<td>(\delta)(_{as})NH(_3) (36)</td>
</tr>
<tr>
<td>1427</td>
<td>1445</td>
<td>(\omega)C(<em>\gamma)H(<em>2) (17), (\nu)C(</em>\gamma)C(</em>\delta) (13), (\sigma)C(_\gamma)H(_2) (13), (\omega)C(_3)H(<em>2) (13), (\nu)C(<em>3)N(</em>\varepsilon) (9), (\nu)C(<em>3)C(</em>\gamma) (5), (\delta)(</em>{as})NH(<em>3) (5), (\rho)C(</em>\delta)O (5)</td>
</tr>
<tr>
<td>1411</td>
<td>1419</td>
<td>(\nu)CO (26), (\nu)C(<em>\alpha)H (12), (\nu)C(</em>\alpha)C (11), (\omega)C(_\gamma)H(<em>2) (8), (\beta)COO (7), (\nu)C(<em>3)N(</em>\varepsilon) (7), (\delta)(</em>{as})NH(<em>3) (5), (\rho)C(</em>\delta)O (6)</td>
</tr>
<tr>
<td>1365</td>
<td>1388</td>
<td>(\rho)C(<em>\alpha)H (25), (\omega)C(<em>3)H(<em>2) (13), (\nu)CO (10), (\rho)C(</em>\alpha)H (9), (\nu)C(</em>\alpha)C(</em>\beta) (8), (\nu)C(<em>\alpha)H (9), (\nu)C(</em>\alpha)C(<em>\beta) (8), (\nu)C(</em>\alpha)C (5)</td>
</tr>
<tr>
<td>1351</td>
<td>1358</td>
<td>(\rho)C(_\alpha)H (30), (\omega)C(_3)H(_2) (16), (\tau)C(_3)H(_2) (15), (\tau)C(_3)H(_2) (9), (\nu)C(<em>3)N(</em>\varepsilon) (6)</td>
</tr>
<tr>
<td>1328</td>
<td>1348</td>
<td>(\rho)C(_\alpha)H (26), (\omega)C(_3)H(<em>2) (18), (\rho)C(</em>\alpha)H (13), (\nu)C(<em>3)N(</em>\varepsilon) (11)</td>
</tr>
<tr>
<td>1293</td>
<td>1309</td>
<td>(\tau)C(_3)H(<em>2) (35), (\nu)C(</em>\alpha)H (26), (\rho)C(_3)H(_2) (8), (\nu)C(_3)N (6)</td>
</tr>
<tr>
<td>1264</td>
<td>1272</td>
<td>(\omega)C(_3)H(_2) (43), (\omega)C(_3)H(<em>2) (20), (\nu)C(</em>\alpha)H (13)</td>
</tr>
<tr>
<td>1206</td>
<td>1215</td>
<td>(\tau)C(_3)H(_2) (21), (\tau)C(<em>3)C(</em>\beta) (18), (\tau)C(_3)H(_2) (16), (\tau)C(<em>3)H(<em>2) (9), (\omega)C(<em>3)C(</em>\varepsilon)C(</em>\alpha)C(</em>\varepsilon) (5)</td>
</tr>
<tr>
<td>1158</td>
<td>1153</td>
<td>(\rho)C(_\alpha)H (20), (\tau)C(_3)H(_2) (17), (\nu)C(<em>3)N(</em>\varepsilon) (13), (\tau)C(_3)H(_2) (12), (\nu)C(<em>3)C(</em>\beta) (8), (\rho)C(_3)H(_2) (6)</td>
</tr>
<tr>
<td>1130</td>
<td>1122</td>
<td>(\nu)C(<em>3)C(</em>\varepsilon) (34), (\rho)N(<em>2)H(<em>2) (17), (\nu)C(</em>\alpha)C(</em>\varepsilon) (7), (\nu)C(<em>\alpha)N (6), (\beta)N(<em>2)C(</em>\varepsilon)C(</em>\varepsilon) (5)</td>
</tr>
<tr>
<td>1109</td>
<td>1100</td>
<td>(\rho)N(<em>3) (27), (\rho)C(</em>\alpha)H (10), (\rho)N(<em>3)H(<em>2) (10), (\rho)C(</em>\alpha)H (9), (\delta)(</em>{as})N(<em>3)C(</em>\alpha) (5), (\rho)C(_\alpha)H (7), (\nu)C(_3)H(_2) (7)</td>
</tr>
<tr>
<td>1110</td>
<td>1097</td>
<td>(\nu)C(<em>3)C(</em>\varepsilon) (26), (\rho)N(_3)H(<em>2) (26), (\nu)C(</em>\alpha)N (13), (\rho)N(_3)H(_2) (8)</td>
</tr>
<tr>
<td>1038</td>
<td>1035</td>
<td>(\nu)C(_3)N (36), (\nu)C(<em>3)C(</em>\varepsilon) (9), (\rho)C(_3)H(_2) (8), (\rho)N(_3)H(_2) (6), (\rho)N(_3)H(<em>2) (5), (\rho)C(</em>\alpha)H (5)</td>
</tr>
<tr>
<td>1078</td>
<td>1003</td>
<td>(\rho)C(<em>\alpha)N (25), (\rho)C(</em>\alpha)N (19), (\rho)C(_3)H(_2) (14), (\rho)C(<em>3)H(<em>2) (14), (\nu)C(</em>\alpha)C(</em>\beta) (7), (\tau)C(_3)H(_2) (5)</td>
</tr>
<tr>
<td>1006</td>
<td>974</td>
<td>(\rho)N(<em>3)H(<em>2) (38), (\nu)C(<em>3)C(</em>\varepsilon) (25), (\tau)C(<em>3)C(</em>\varepsilon) (8), (\sigma)C(</em>\alpha)C(</em>\varepsilon) (7), (\rho)N(_3)H(_2) (6)</td>
</tr>
</tbody>
</table>

\(^{a}\)\(\nu\): stretch; \(\delta_{as}\): asymmetric deformation; \(\delta\): symmetric deformation; \(\delta\): deformation; \(\sigma\): scissoring; \(\rho\): rocking; \(\omega\): wagging; \(\beta\): in-plane bending; \(\tau\): twisting.
Table 4.2: UVRR Frequencies (cm\(^{-1}\)) and Assignments of L-Gln in D\(_2\)O. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>Potential Energy Distribution(^a) (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1650</td>
<td>1739</td>
<td>(\nu C_\delta O_\epsilon) (78), (-\nu C_\delta N_\epsilon) (7), (\beta N_\epsilon C_\delta C_\gamma) (7)</td>
</tr>
<tr>
<td>1637</td>
<td>1708</td>
<td>(\nu CO) (55), (-\nu CO) (34), (-\rho C_\alpha C) (7)</td>
</tr>
<tr>
<td>1465</td>
<td>1494</td>
<td>(\sigma C_\beta H_2) (88)</td>
</tr>
<tr>
<td>1453</td>
<td>1465</td>
<td>(-\sigma C_\beta H_2) (48), (-\omega C_\beta H_2) (15), (-\nu C_\beta C_\gamma) (8), (-\nu C_\delta N_\epsilon) (7), (\omega C_\beta H_2) (6)</td>
</tr>
<tr>
<td>1440</td>
<td>1454</td>
<td>(-\sigma C_\beta H_2) (35), (\nu C_\delta N_\epsilon) (19), (-\nu C_\alpha C_\delta) (14), (\omega C_\beta H_2) (9), (-\rho C_\delta O_\epsilon) (5)</td>
</tr>
<tr>
<td>1412</td>
<td>1424</td>
<td>(-\nu CO) (29), (\nu C_\alpha C) (12), (-\nu CO) (10), (\omega C_\beta H_2) (10), (\beta COO) (9), (\nu C_\delta N_\epsilon) (9), (\rho C_\alpha H) (8)</td>
</tr>
<tr>
<td>1368</td>
<td>1393</td>
<td>(\omega C_\beta H_2) (21), (\rho C_\alpha H) (15), (\nu CO) (11), (\nu C_\alpha C_\beta) (8), (-\rho C_\alpha H) (6), (\nu CO) (6), (-\nu C_\alpha C) (5), (\nu C_\delta N_\epsilon) (5)</td>
</tr>
<tr>
<td>1344</td>
<td>1362</td>
<td>(-\rho C_\alpha H) (37), (-\omega C_\beta H_2) (13), (\nu C_\delta N_\epsilon) (12), (\omega C_\beta H_2) (11), (\rho C_\alpha H) (7), (\sigma N_\delta D_2) (5)</td>
</tr>
<tr>
<td>1349</td>
<td>1349</td>
<td>(-\tau C_\beta H_2) (44), (\tau C_\gamma H_2) (20), (-\rho C_\alpha H) (15), (-\rho C_\alpha H) (6)</td>
</tr>
<tr>
<td>1304</td>
<td>1304</td>
<td>(-\rho C_\alpha H) (32), (-\tau C_\beta H_2) (29), (\rho C_\beta H_2) (8), (-\rho C_\alpha H) (7)</td>
</tr>
<tr>
<td>1282</td>
<td>1282</td>
<td>(-\omega C_\beta H_2) (37), (-\omega C_\beta H_2) (33), (\sigma N_\delta D_2) (8), (\rho C_\alpha H) (7), (\nu C_\delta N_\epsilon) (6)</td>
</tr>
<tr>
<td>1202</td>
<td>1202</td>
<td>(-\delta_\alpha N\delta D_3) (23), (\delta_\alpha \alpha N\delta D_3) (20), (\tau C_\beta H_2) (17), (\tau C_\gamma H_2) (12), (\nu C_\alpha C_\beta) (8)</td>
</tr>
<tr>
<td>1198</td>
<td>1198</td>
<td>(\delta_\alpha \alpha N\delta D_3) (27), (\tau C_\beta H_2) (12), (\tau C_\gamma H_2) (11), (-\rho C_\alpha H) (9), (\nu C_\alpha C_\beta) (8), (\delta_\alpha \alpha N\delta D_3) (7), (-\delta_\alpha \alpha N\delta D_3) (7)</td>
</tr>
<tr>
<td>1186</td>
<td>1186</td>
<td>(-\delta_\alpha \alpha N\delta D_3) (52), (-\delta_\alpha \alpha N\delta D_3) (27), (-\rho C_\alpha N) (8), (-\delta_\alpha \alpha \alpha N\delta D_3) (6)</td>
</tr>
<tr>
<td>1161</td>
<td>1150</td>
<td>(\sigma N\delta D_2) (56), (-\rho C_\delta O_\epsilon) (11), (\nu C_\gamma C_\delta) (9), (\omega C_\beta H_2) (8)</td>
</tr>
<tr>
<td>1131</td>
<td>1131</td>
<td>(\nu C_\alpha C_\beta) (21), (-\delta_\alpha N\delta D_3) (16), (-\nu C_\alpha N) (14), (\delta_\alpha \alpha N\delta D_3) (13), (-\tau C_\gamma H_2) (5)</td>
</tr>
<tr>
<td>1111</td>
<td>1111</td>
<td>(-\nu C_\beta C_\gamma) (17), (\delta_\alpha \alpha N\delta D_3) (10), (\delta N_\alpha C_\beta C_\delta) (9), (-\nu C_\alpha C_\beta) (8), (-\delta_\alpha \alpha N\delta D_3) (8), (\delta C_\alpha C_\beta) (7), (-\nu C_\alpha N) (6), (-\delta_\alpha \alpha N\delta D_3) (5)</td>
</tr>
<tr>
<td>1105</td>
<td>1105</td>
<td>(\nu C_\gamma C_\gamma) (44), (-\delta_\alpha N\delta D_3) (15), (\delta_\alpha \alpha N\delta D_3) (8), (-\delta_\alpha \alpha N\delta D_3) (6), (-\nu C_\alpha C_\beta) (5)</td>
</tr>
<tr>
<td>1058</td>
<td>1058</td>
<td>(\rho C_\beta H_2) (21), (\rho C_\beta H_2) (16), (\rho C_\beta H_2) (12), (\delta N_\alpha C_\beta C_\delta) (9), (-\nu C_\alpha C) (6), (\delta C_\alpha C_\beta) (6), (\nu C_\alpha C_\beta) (5), (-\rho N\delta D_3) (5)</td>
</tr>
<tr>
<td>992</td>
<td>998</td>
<td>(\nu C_\alpha N) (29), (\delta_\alpha N\delta D_3) (15), (-\delta N_\alpha C_\beta C_\delta) (10), (-\rho C_\alpha C) (7), (-\nu C_\beta C_\gamma) (5), (-\sigma C_\alpha C_\beta) (5)</td>
</tr>
<tr>
<td>960</td>
<td>959</td>
<td>(\rho N\delta D_2) (22), (\nu C_\delta N_\epsilon) (14), (\nu C_\delta N_\epsilon) (10), (-\sigma N_\delta D_2) (9), (-\nu C_\alpha N) (9), (-\beta N_\epsilon C_\delta C_\gamma) (8), (\nu C_\alpha C) (6), (\nu C_\delta O_\epsilon) (6)</td>
</tr>
<tr>
<td>929</td>
<td>920</td>
<td>(\nu C_\alpha C) (17), (\rho C_\beta H_2) (16), (\rho N\delta D_3) (16), (\tau C_\beta H_2) (8), (\beta COO) (8), (-\delta N_\alpha C_\beta C_\delta) (7), (\Pi N_\epsilon C_\delta C_\gamma) (6)</td>
</tr>
</tbody>
</table>

\(^a\): \(\nu\): stretch; \(\delta_\alpha\): asymmetric deformation; \(\sigma\): symmetric deformation; \(\delta\): deformation; \(\sigma\): scissoring; \(\rho\): rocking; \(\omega\): wagging; \(\beta\): in-plane bending; \(\tau\): twisting; \(\Pi\): out-of-plane deformation.
rocking/CH₂ twisting mode. The PED of this vibration contains a significant contribution of NH₃ rocking, which likely accounts for the disappearance of this band upon N-deuteration. We assign the ~1080 cm⁻¹ and ~1005 cm⁻¹ bands to NH₃ rocking vibrations.

The remaining two bands in the 1000 cm⁻¹ to 1200 cm⁻¹ region are located at ~1130 cm⁻¹ and ~1110 cm⁻¹. The observed frequency difference between these two vibrations is ~20 cm⁻¹, which is close to the calculated ~25 cm⁻¹ difference of our DFT calculations. We assign the 1130 cm⁻¹ band to a vibration that is mainly an in-phase combination of C β–C γ stretching and N εH₂ rocking. The ~1110 cm⁻¹ band is assigned to a vibration that consists of an out-of-phase combination of C β–C γ stretching and N εH₂ rocking. This vibration also contains a significant C δ–N ε stretching component (~13%), which is in-phase with N εH₂ rocking.

The in-phase combination of N εH₂ rocking and C δ–N ε stretching of the ~1110 cm⁻¹ vibration is reminiscent of the AmIII mode of secondary amides. While complex, the AmIII vibration contains significant contributions of in-phase C–N stretching and N–H in-plane bending motions of the secondary amide group. We propose to call the ~1110 cm⁻¹ mode the Amide III P (Am III P) since the eigenvector composition of this vibration is analogous to that of the canonical AmIII of secondary amides. As discussed in detail below, the Am III P vibration is sensitive to the χ₃ and χ₂ dihedral angles of Gln and Asn.

4.4.2 Conformational Dependence of the Am III P Band

We performed DFT calculations on L-Gln molecules with χ₃ dihedral angles fixed at different values (see Computational Section for details) in order to identify spectroscopic markers that are diagnostic of the side chain χ₃ and χ₂ dihedral angles of Gln and Asn, respectively. We examined the frequency dependence of different primary amide vibrations and found that the Am III P vibrational frequency and normal mode depends strongly on the OCCC dihedral angle.

Figure 4.3a shows the calculated cosinusoidal dependence of the Am III P vibrational frequency on the χ₃ dihedral angle. The maximum frequency of the vibration occurs at χ₃ ~0°, while minima occur near χ₃ ±90°. The Gln Am III P band frequency dependence on
Figure 4.3: Calculated Am III$^P$ frequency and bond length dependence on the $\chi_3$ dihedral angle of the Gln side chain. (a) Am III$^P$ frequency dependence; (b) $C_\beta$–$C_\gamma$ bond length; (c) $C_\delta$–$N_\epsilon$ bond length; (d) $C_\alpha$–$C_\beta$ bond length; and (e) shows the dependence of the Am III$^P$ frequency on the $C_\beta$–$C_\gamma$ bond length. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
the $\chi_3$ angle follows a cosinusoidal relationship:

$$\nu(\chi_3) = \nu_0 + A \cos(2\chi_3) + B \cos(\chi_3 - C)$$  \hspace{1cm} (4.1)$$

where $\nu_0 = 1084$ cm$^{-1}$, $A = 10$ cm$^{-1}$, $B = 3$ cm$^{-1}$, and $C = -31^\circ$. These parameters were calculated from a least-squares fit of eq. 4.1 to the frequency dependence on the $\chi_3$ angle in Figure 4.3a.

Figure 4.3a shows that the Am III$^P$ frequency dependence on the $\chi_3$ dihedral angle is asymmetric about $\chi_3 \sim 0^\circ$. This asymmetry is due to the chirality of L-Gln and L-Asn and leads to the requirement of two cosine terms to express the $\chi_3$ frequency dependence of eq. 4.1. This is evident when we compare the L-Gln $\chi_3$ dependence on the Am III$^P$ frequency with that of butyramide (shown in Figure C1). In the case of butyramide, which is achiral, there is no asymmetry about $0^\circ$. As a result, the Am III$^P$ frequency dependence on the OCCC dihedral angle of butyramide can be satisfactorily modeled with just one cosine term (eq. C.1).

4.4.3 Origin of the OCCC Dihedral Angle Dependence of the Am III$^P$ Vibration

Understanding the conformational dependence of the Am III$^P$ frequency on the primary amide OCCC dihedral angle requires a detailed knowledge of the atomic motions that give rise to the vibration. On the basis of our normal mode calculations of Gln, butyramide (Table C14), and propanamide [146], we conclude that $N_\epsilon H_2$ rocking, $C_\delta-N_\epsilon$ stretching, and $C_\beta-C_\gamma$ stretching define the Am III$^P$ vibration. However, depending on the OCCC dihedral angle, other motions such as $C_\beta H_2$ twisting and $C_\alpha-C_\beta$ stretching can contribute to this vibration.

Therefore, we examined how the Gln $C_\delta-N_\epsilon$, $C_\beta-C_\gamma$, and $C_\alpha-C_\beta$ bond lengths change as a function of the $\chi_3$ dihedral angle in order to understand the origin of the conformational sensitivity of the Am III$^P$ vibration. Changes in these bond lengths impact the Am III$^P$ frequency by affecting the vibrational mode bond force constants. As seen in Figure 4.3b-d, all the bond lengths show a dependence on the $\chi_3$ dihedral angle. However, as seen in Figure 4.3b, the $C_\beta-C_\gamma$ bond length shows the largest dependence on the $\chi_3$ dihedral angle.
The Am III$^\text{P}$ vibrational frequency has a strong correlation with the $C_\beta$–$C_\gamma$ bond length, as shown in Figure 4.3e. The Am III$^\text{P}$ vibrational frequency increases as the $C_\beta$–$C_\gamma$ bond length decreases and *vice versa*.

The $C_\beta$–$C_\gamma$ bond length dependence on the $\chi_3$ dihedral angle appears to be due to hyperconjugation between the $C_\beta$–$C_\gamma$ $\sigma$ and the $C_\delta$=$O_\epsilon$ $\pi^*$ orbitals (Figure 4.4). This interaction is strongest when these orbitals maximally overlap, in the absence of significant phase cancellation due to the $\pi^*$ orbital antisymmetry. When hyperconjugation occurs, the $\sigma$ orbital donates electron density to the $\pi^*$ orbital, which decreases the $C_\beta$–$C_\gamma$ bond order and increases its bond length. This decreases the $C_\beta$–$C_\gamma$ stretching force constant, which downshifts the Am III$^\text{P}$ frequency.

![Figure 4.4](image)

Figure 4.4: Hyperconjugation results in the $C_\beta$–$C_\gamma$ bond length sensitivity to the $\chi_3$ dihedral angle. Overlap of $C_\beta$–$C_\gamma$ $\sigma$ and $C_\delta$=$O_\epsilon$ $\pi^*$ NBO molecular orbitals when the $\chi_3$ dihedral angle is: (a) $0^\circ$; (b) $+90^\circ$; and (c) $\pm 180^\circ$. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

We tested this hypothesis with natural bond orbital (NBO) analysis, which allows the DFT calculated electron densities to be displayed in terms of approximate $\sigma$ and $\pi^*$ molecular orbitals. According to our hypothesis, the $C_\beta$–$C_\gamma$ bond length should be largest when
hyperconjugation is maximized and smallest when there is no hyperconjugation. Indeed, as seen in Figure 4.4b, there is significant overlap of the C_β–C_γ σ and C_δ=O_ϵ π* NBO molecular orbitals at ±90°, where the C_β–C_γ bond length is largest. In contrast, at χ_3 ~0°, where the C_β–C_γ bond length is shortest, the orbital overlap cancels due to the antisymmetry of the π* orbital. Figure 4.5 shows the NBO charge on the C_β atom. As expected from our hyperconjugation hypothesis, the NBO C_β atom charge is less negative at χ_3 ~ ±90° compared to χ_3 ~ 0°. The NBO C_β atom charge becomes even more negative at χ_3 ~ ±150° and χ_3 ~ ±180°, even without additional hyperconjugation of the C_β–C_γ σ and C_δ=O_ϵ π* orbitals. This result is likely an artifact because these extreme χ_3 dihedral angles are associated with physically impossible high energy structures that will be subject to other electron density alterations.

Our model accounts for the Am III^p frequency downshift as the dihedral angles approach χ_3 ~ ±90°, where hyperconjugation is strongest. This behavior is the reverse of the Bohlmann effect [183–186], where a “negative” hyperconjugation transfers electron density from a lone pair orbital to an optimally positioned C–H σ* orbital. This decreases the C–H bond order and substantially downshifts the C–H stretching frequencies.
4.4.4 Experimental Dependence of Am III$^p$ Band Frequency on OCCC Dihedral Angle

We experimentally examined the dependence of the Am III$^p$ band frequency on the primary amide OCCC dihedral angle by measuring the UVRR and visible Raman spectra of different Gln and Asn derivatives in the solid-state. We determined the structures of each of the different Gln and Asn derivative crystals with X-ray diffraction, and assigned the Am III$^p$ band by performing DFT calculations and examining band shifts upon N-deuteration. Our X-ray diffraction methods and the band assignments of the crystals are discussed, in detail, in APPENDIX C.

Figure 4.6: Experimental correlation of the Am III$^p$ frequency to the $\chi_3$ dihedral angle. The average frequency (from the 633 nm and 229 nm Raman spectra) of the Am III$^p$ band was plotted as a function of the OCCC dihedral angle. 1 = L-Gln; 2 = Gly-Gln; 3 = D-Gln; 4 = GlnTBE; 5 = NAcGln; and 6 = Ser-Asn. The data were fit with eq. 4.2 (black line, $r^2_{adj} = 0.83$). The blue curve corresponds to eq. 4.3. The red curve corresponds to eq. 4.4. The yellow curve corresponds to eq. 4.5 and is an average of the red and blue curves. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

4.4.4.1 Dependence of Am III$^p$ Band Frequency in Crystals Figure 4.6 shows the Am III$^p$ frequency dependence on the experimentally determined primary amide OCCC dihedral angles. We fit the experimental data to a function of the same form as eq. 4.1,
obtaining the following relationship:

\[ \nu(\chi_3) = 1066 \text{ cm}^{-1} + 29 \text{ cm}^{-1} \cos(2\chi_3) + 9 \text{ cm}^{-1} \cos(\chi_3 + 99^\circ) \] (4.2)

which is shown in the Figure 4.6 black curve. To obtain the eq. 4.2 parameters, we fixed the A/B ratio to \( \sim 3 \) as found in eq. 4.1, and performed a least squares minimization of the experimental data. Eq. 4.2 provides an excellent fit of the experimental data and captures the chiral asymmetry that occurs near \( \chi_3 \sim \pm 90^\circ \).

### 4.4.4.2 Dependence of Am III\(^{P}\) Band Frequency for Fully Hydrated Primary Amides

The Am III\(^{P}\) band frequency also depends on the local hydrogen bonding and dielectric environment of the primary amide group [146]. In water, the Am III\(^{P}\) band of L-Gln is located at \( \sim 1110 \text{ cm}^{-1} \), as compared with \( \sim 1097 \text{ cm}^{-1} \) in the solid-state. Based on Rhys et al.’s neutron diffraction study [187], the solution-state equilibrium structure of L-Gln in water does not appear to differ significantly from the single known L-Gln crystal structure [188]. From their solution-state structure, we determine that the equilibrium \( \chi_3 \) dihedral angle of L-Gln in water is \( \sim -12.8^\circ \). This differs by less than a degree \( (-13.54^\circ) \) from the L-Gln crystal examined in this study. Thus, by setting the Am III\(^{P}\) frequency to \( 1110 \text{ cm}^{-1} \), \( \chi_3 \) to \( -13.54^\circ \), and solving for \( \nu_0 \), we obtain eq. 4.3

\[ \nu(\chi_3) = 1083 \text{ cm}^{-1} + 29 \text{ cm}^{-1} \cos(2\chi_3) + 9 \text{ cm}^{-1} \cos(\chi_3 + 99^\circ) \] (4.3)

which is shown by the Figure 4.6 blue curve. This equation correlates the Am III\(^{P}\) band frequency to OCCC dihedral angles for situations in which the primary amide group is fully exposed to water, such as in PPII-like structures, 2.5\(_1\)-helices [189], and extended \( \beta \)-strand-like peptide conformations dissolved in water.

### 4.4.4.3 Dependence of Am III\(^{P}\) Band Frequency for Low Dielectric Constant and Weak Hydrogen Bonding Environments

The Am III\(^{P}\) frequency downshifts \( \sim 15 \text{ cm}^{-1} \) in the low dielectric and hydrogen bonding environment of acetonitrile compared to that in water (see APPENDIX C and Figure C6). This downshift derives from the different water versus acetonitrile stabilizations of the ground state \( O_e=\text{C}_\delta N_e \text{H}_2 \) and \( O_e\text{C}_\delta=N_e \text{H}_2^+ \).
resonance structures of the primary amide group [146]. In both solvents, the $O\equiv C\delta N, H_2$ resonance structure dominates; however, in acetonitrile the $O\equiv C\delta = N, H_2^+$ resonance structure contributes less than in water. Thus, the $C\delta - N$ bond length is larger in acetonitrile compared to water due to the lesser favorability of the $O\equiv C\delta = N, H_2^+$ resonance structure. Consequently, there is a smaller $C\delta - N$ stretching force constant in acetonitrile compared to water, which results in a downshift of the Am III$^P$ frequency.

Eq. 4.3 can be modified in order to account for situations where the primary amide group is not engaged in significant hydrogen bonding interactions or when located in a low dielectric environment. We apply a $15 \text{ cm}^{-1}$ downshift in $\nu_0$ from eq. 4.3 to determine eq. 4.4:

$$\nu(\chi_3) = 1068 \text{ (cm}^{-1}\text{)} + 29 \text{ (cm}^{-1}\text{)} \cos(2\chi_3) + 9 \text{ (cm}^{-1}\text{)} \cos(\chi_3 + 99^\circ)$$

which is shown in red in Figure 4.6.

### 4.4.4.4 Dependence of Am III$^P$ Band Frequency for Unknown Dielectric and Hydrogen Bonding Environments

We suggest the use of eq. 4.5, which is the average of eqs. 4.3 and 4.4, for cases where the hydrogen bonding and dielectric environment of the primary amide group is unknown:

$$\nu(\chi_3) = 1076 \text{ (cm}^{-1}\text{)} + 29 \text{ (cm}^{-1}\text{)} \cos(2\chi_3) + 9 \text{ (cm}^{-1}\text{)} \cos(\chi_3 + 99^\circ)$$

It can be applied, for example, to determine the side chain $\chi_3$ and $\chi_2$ dihedral angles of Gln and Asn residues located in turn structures of proteins. For these residues, it may not be clear if the side chains are hydrogen bonded to water, to other side chains, or the peptide backbone. Eq. 4.5 is shown by the yellow curve in Figure 4.6.

### 4.4.5 Predicting Side Chain $\chi_3$ and $\chi_2$ Dihedral Angles in Gln and Asn as a Function of Ramachandran ($\Phi$, $\Psi$) Angles

Shapovalov and Dunbrack [179] recently developed a new peptide backbone dependent rotamer library, which includes the non-rotameric Gln and Asn side chain $\chi_3$ and $\chi_2$ dihedral angles. Their database was compiled by analyzing high resolution crystal structures from
Figure 4.7: Gln and Asn side chain $\chi_3$ and $\chi_2$ dihedral angle dependence on secondary structure. Plots showing Ramachandran angles for PDB entries from the Shapovalov and Dunbrack database of (a) Gln and (b) Asn. The colored boxes correspond to canonical PPII ($\Phi = -65^\circ$, $\Psi = 145^\circ$) angles (red), $\beta$-sheet ($\Phi = -115^\circ$, $\Psi = 130^\circ$) angles (blue), and $\alpha$-helix ($\Phi = -63^\circ$, $\Psi = -43^\circ$) angles (cyan). Distributions of $\chi_3$ and $\chi_2$ dihedral angles for Gln and Asn residues that have ($\Phi$, $\Psi$) angles close to canonical: (c, d) PPII-like, (e, f) $\beta$-sheet, and (g, h) $\alpha$-helical structures. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

the protein data bank (PDB), and consists of $\sim$30,000 entries for Asn and $\sim$20,000 entries for Gln. Figure 4.7a and Figure 4.7b show Ramachandran plots of all of the Gln and Asn entries in the Shapovalov and Dunbrack database. The Gln and Asn side chains populate similar regions of the Ramachandran plot, and both show a preference for $\alpha$-helical region ($\Phi$, $\Psi$) angles. Asn populates a much broader range of ($\Phi$, $\Psi$) angles, especially in the nearly forbidden “bridge” region between $\beta$-sheet and $\alpha$-helical regions of the Ramachandran plot.
We used the Shapovalov and Dunbrack database to examine the side chain \( \chi_3 \) and \( \chi_2 \) dihedral angle preferences of Gln and Asn residues that possess canonical PPII, \( \beta \)-sheet, or \( \alpha \)-helix Ramachandran angle values. Based on work by Richardson [190] and Karplus [191], we assume \((\Phi, \Psi)\) angles centered around \((-65^\circ, 145^\circ)\) for canonical PPII structures, \((-115^\circ, 130^\circ)\) for canonical \( \beta \)-sheets, and \((-63^\circ, -43^\circ)\) for canonical \( \alpha \)-helices. Figure 4.7c-h depict histograms of the \( \chi_3 \) and \( \chi_2 \) dihedral angles observed for the population of Gln and Asn residues with canonical PPII, \( \beta \)-sheet, or \( \alpha \)-helical Ramachandran angles.

The Gln and Asn side chain \( \chi_3 \) and \( \chi_2 \) dihedral angles clearly depend upon the peptide bond \( \Phi \) and \( \Psi \) angles. This correlation could result from a preference for particular \( \chi_3 \) or \( \chi_2 \) dihedral angles for stretches of consecutive peptide bonds with \((\Phi, \Psi)\) angles that result in PPII, \( \beta \)-sheet, or \( \alpha \)-helical secondary structures. Alternatively, it could result from a preference for \( \chi_3 \) or \( \chi_2 \) dihedral angles for the \((\Phi, \Psi)\) angle values of their individual peptide bonds.

The \( \chi_3 \) and \( \chi_2 \) dihedral angle histograms of Gln and Asn residues that populate the canonical PPII region of the Ramachandran plot are shown in Figure 4.7c and Figure 4.7d. The distribution of \( \chi_3 \) angles adopted by Gln is broader than that of the \( \chi_2 \) angles of Asn. Both histograms are centered about negative dihedral angles, with Gln showing a peak at around \( \chi_3 \sim -8^\circ \) and Asn showing a peak near \( \chi_2 \sim -36^\circ \). It should be noted that the bias due to the L- amino acid chirality gives rise to a clear preference for negative \( \chi_2 \) dihedral angles for the shorter side chain Asn residues.

The \( \chi_3 \) and \( \chi_2 \) dihedral angle histograms of Gln and Asn with \( \beta \)-sheet \((\Phi, \Psi)\) angles in Figure 4.7e and Figure 4.7f differ dramatically from one another. The population of Gln \( \chi_3 \) dihedral angles (Figure 4.7e) is nearly symmetric about \( \chi_3 \sim 0^\circ \). The histogram is bimodal, with two peaks located near \( \chi_3 \) angles of \( \sim -44^\circ \) and \( \sim 41^\circ \). In contrast, the population of Asn residues (Figure 4.7f) predominately adopts negative dihedral angles and is peaked around \( \chi_2 \sim -61^\circ \). A minor peak also occurs around \( \chi_2 \sim 56^\circ \).

Figure 4.7g and Figure 4.7h show histograms of the \( \chi_3 \) and \( \chi_2 \) dihedral angles of Gln and Asn residues that adopt canonical \( \alpha \)-helical Ramachandran angles. As in Figure 4.7e, the Figure 4.7g Gln \( \chi_3 \) dihedral angle population is roughly bimodal and nearly symmetric about \( \chi_3 \sim 0^\circ \). It is peaked at \( \chi_3 \) angles of \( \sim -34^\circ \) and \( \sim 45^\circ \). In contrast, in Figure 4.7h the
population of Asn $\chi_2$ dihedral angles is narrow and sharply peaked at $\chi_2 \sim -19^\circ$ with two minor peaks at $\chi_2 \sim -49^\circ$ and $\sim 62^\circ$.

The $\chi_3$ and $\chi_2$ dihedral angle dependencies on the peptide bond Ramachandran angles, shown by the Shapovalov and Dunbrack database, enable us to predict the most probable Am III$^P$ frequencies of Gln and Asn residues that adopt canonical PPII, $\beta$-sheet, and $\alpha$-helix ($\Phi$, $\Psi$) angles (shown in Table 4.3). For example, using eq. 4.3, we calculate that Gln and Asn side chains with PPII ($\Phi$, $\Psi$) angles will have a maximum probability of showing Am III$^P$ bands centered at $\sim 1111$ cm$^{-1}$ and $\sim 1096$ cm$^{-1}$, respectively. Similarly, we calculate that the Am III$^P$ bands of Gln residues with $\beta$-sheet Ramachandran angles will have the greatest probability of being located at $\sim 1080$ cm$^{-1}$ and/or $\sim 1089$ cm$^{-1}$. In contrast, the Am III$^P$ bands for Asn residues with $\beta$-sheet ($\Phi$, $\Psi$) angles will have the largest probability of being located at $\sim 1064$ cm$^{-1}$ and/or $\sim 1075$ cm$^{-1}$. For $\alpha$-helical Ramachandran angles, we calculate that the probability maxima for Am III$^P$ bands will be at $\sim 1076$ cm$^{-1}$ and/or $\sim 1098$ cm$^{-1}$ for Gln and $\sim 1058$ cm$^{-1}$, $\sim 1085$ cm$^{-1}$, and/or $\sim 1107$ cm$^{-1}$ for Asn residues.

<table>
<thead>
<tr>
<th></th>
<th>Gln</th>
<th>Asn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Phi$ (°) $\Psi$ (°)</td>
<td>$\chi_3$ (°) $\chi_2$ (°)</td>
</tr>
<tr>
<td>PPII</td>
<td>-65 145</td>
<td>-8 (-22, -32)$^a$</td>
</tr>
<tr>
<td>$\beta$-sheet</td>
<td>115 130</td>
<td>-44, 41</td>
</tr>
<tr>
<td>$\alpha$-helix</td>
<td>-63 130</td>
<td>-34, 45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Values in parentheses were measured experimentally for $Q_3$ and $D_2Q_{10}K_2$. We can calculate the expected Raman spectral Am III$^P$ band shapes from the Gln $\chi_3$ and Asn $\chi_2$ dihedral angle histograms in Figure 4.7 using the Am III$^P$ Raman band frequency dependencies of eqs. 4.2–4.5. These calculated band shapes (not shown) are unphysically broad ($> 100$ cm$^{-1}$). This clearly indicates that these histograms derive from the inhomogeneous distribution of $\chi_3$ and $\chi_2$ angles of individual Gln and Asn residues within the proteins found in the Shapovalov and Dunbrack database. This distribution of Raman frequencies from the calculated Am III$^P$ band is much broader than the homogeneous linewidth of an Am III$^P$ band expected for a single Gln and Asn residue in a typical PPII, $\beta$-sheet, or $\alpha$-helix.
helix conformation in proteins. The large widths of the Gln $\chi_3$ and Asn $\chi_2$ dihedral angle histograms result because the residues in the Shapovalov and Dunbrack database exist in a larger distribution of conformations, hydrogen bonding states, and chemical environments than we have so far encountered in our UVRR investigations.

4.4.6 Experimentally Determined Gln PPII-like Structure Peptide $\chi_3$ Dihedral Angles

4.4.6.1 UVRR Spectra of Gln Peptides in PPII-like Structures  We examined the UVRR spectra of two peptides, $Q_3$ and $D_2Q_{10}K_2$, in order to determine their solution-state $\chi_3$ angles. Xiong et al. [125] previously showed that $D_2Q_{10}K_2$ exists in predominately PPII-like and 2.5$_1$-helix-like conformations when prepared using a “disaggregation” protocol developed by Wetzel and coworkers [77]. In this protocol, the $D_2Q_{10}K_2$ peptide is initially dissolved in a mixture of trifluoroacetic acid and hexafluoroisopropanol. These solvents are subsequently evaporated under dry N$_2$ gas, and the peptide is redissolved in pure water.

The UVRR spectra indicate that $Q_3$ has predominately PPII-like peptide bond conformations. Figure 4.8a shows the peak fitted $\sim$204 nm excitation UVRR spectrum of $Q_3$ in the region between 1050 cm$^{-1}$ to 1500 cm$^{-1}$. The AmIII$_3$ region, between $\sim$1200 cm$^{-1}$ to 1280 cm$^{-1}$, is most sensitive to the secondary structure of the peptide since its frequency depends on the Ramachandran $\Psi$ angle [96, 97]. This region is well fit by two Gaussian bands located at $\sim$1210 cm$^{-1}$ and $\sim$1260 cm$^{-1}$. Using the methodology of Mikhonin et al. [97], we correlated the band peak positions to their $\Psi$ angles. We used their eq. 6A to correlate the 1210 cm$^{-1}$ frequency of the AmIII$_3$ band to a $\Psi$ angle of 103°±3° and the 1260 cm$^{-1}$ frequency to a $\Psi$ angle of 157°±2°. The $\Psi$ angle of $\sim$157° derives from peptide bonds situated in PPII-like conformations, while the $\Psi$ angle of $\sim$103° derives from peptide bond situated in $\beta$-strand-like conformations. Assuming identical Raman cross sections for these two different species, we find that the peptide bonds are dominated by PPII-like $\Psi$ angles ($\sim$87±2%), while a small fraction adopt $\beta$-strand-like $\Psi$ angles ($\sim$13±2%). This is supported by the circular dichroism spectra of $Q_3$ shown in Figure C7, which show a predominantly PPII spectral signature.
Figure 4.8: Deconvolution of the UVRR spectra of Q$_3$ and D$_2$Q$_{10}$K$_2$. (a) Fitting the 204 nm excitation UVRR spectrum of Q$_3$. (b) 198-204 nm difference spectrum of D$_2$Q$_{10}$K$_2$ taken from Xiong et al. [125] The inset shows the Am III$^P$ region of D$_2$Q$_{10}$K$_2$. The $\chi^2_{red}$ statistics for the spectral fits shown in (a) and (b) are 1.1 and 0.74, respectively. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
4.4.6.2 \( \chi_3 \) Dihedral Angle Determination of Side Chains in Gln Peptides  

The Am III\(^P\) bands of Q\(_3\) and D\(_2\)Q\(_{10}\)K\(_2\) are found in the region between \( \sim 1050 \text{ cm}^{-1} \) and \( 1150 \text{ cm}^{-1} \). Based on our normal mode analysis of Gln, we fit this region with four bands that derive from C\(_\alpha\)H rocking/C\(_\gamma\)H\(_2\) twisting, C\(_{\beta}\)–C\(_\gamma\) stretching/N\(_3\)H\(_2\) rocking, NH\(_3\) rocking/C\(_\alpha\)–N stretching, and the Am III\(^P\) vibrations. For Q\(_3\), these bands are located within a broad asymmetric spectral feature at \( \sim 1050 \text{ cm}^{-1} \), \( \sim 1106 \text{ cm}^{-1} \), \( \sim 1130 \text{ cm}^{-1} \), and \( \sim 1160 \text{ cm}^{-1} \).

We assigned these bands based on our analysis of Gln. We assign the \( \sim 1160 \text{ cm}^{-1} \) band to a C\(_\alpha\)H rocking/C\(_\gamma\)H\(_2\) twisting mode, the \( \sim 1130 \text{ cm}^{-1} \) band to a C\(_{\beta}\)–C\(_\gamma\) stretching/N\(_3\)H\(_2\) rocking vibration, and the \( \sim 1080 \text{ cm}^{-1} \) band to a NH\(_3\) rocking/C\(_\alpha\)–N stretching mode. The \( \sim 1106 \text{ cm}^{-1} \) band appears as a low-frequency shoulder feature and is assigned to the Am III\(^P\) vibration. This is very close to the predicted Am III\(^P\) vibrational frequency band center from the Gaussian fit of PPII-like structures in Figure 4.7c, as listed in Table 4.3. In fact, the Am III\(^P\) frequency band center of Q\(_3\) differs by only \( \sim 5 \text{ cm}^{-1} \) from the predicted frequency band center (\( \sim 1111 \text{ cm}^{-1} \)) for PPII Ramachandran angles.

Figure 4.8b shows the 198 nm – 204 nm difference spectrum of disaggregated D\(_2\)Q\(_{10}\)K\(_2\) published by Xiong et al. [125] Xiong et al. showed that excitation at 198 nm enhances the primary amide UVRR bands more than does excitation at 204 nm. Thus, the Figure 4.8b D\(_2\)Q\(_{10}\)K\(_2\) difference spectrum is dominated by the primary amide Gln side chain bands with little interference from the secondary amide peptide bond UVRR bands.

The inset in Figure 4.8b shows the region where the Am III\(^P\) band of D\(_2\)Q\(_{10}\)K\(_2\) is located. We parsimoniously peak fit this region to three Gaussian bands located at \( \sim 1099 \text{ cm}^{-1} \), \( \sim 1118 \text{ cm}^{-1} \), and \( \sim 1140 \text{ cm}^{-1} \). Using prior knowledge from our analysis of Gln, we assign the bands at \( \sim 1118 \text{ cm}^{-1} \) and \( \sim 1140 \text{ cm}^{-1} \) to the the C\(_{\beta}\)–C\(_\gamma\) stretching/N\(_3\)H\(_2\) rocking and C\(_\alpha\)H rocking/C\(_\gamma\)H\(_2\) twisting vibrations, respectively. The \( \sim 1099 \text{ cm}^{-1} \) band is assigned to the Am III\(^P\) band.

The Am III\(^P\) bandwidths of Q\(_3\) and D\(_2\)Q\(_{10}\)K\(_2\) are \( \sim 30 \text{ cm}^{-1} \), which is similar to that of Gln in H\(_2\)O (Figure 4.2a). These bandwidths are roughly twice as large as those found in the Raman spectra of the different Gln and Asn derivative crystals, which we measure to be on average \( \sim 13.3(50) \text{ cm}^{-1} \). This bandwidth is significantly larger than our spectrometer resolution of \( \sim 4.5 \text{ cm}^{-1} \). Thus, if we assume a Lorentzian band shape, we estimate that the
Am III$^P$ band homogeneous linewidth for a Gln compound with a well-defined $\chi_3$ angle is $\sim 6.6 \text{ cm}^{-1}$. The fact that the Am III$^P$ bandwidths of solution-state Gln, Q$_3$, and D$_2$Q$_{10}$K$_2$ are much broader than those measured in our crystals suggests that there is a distribution of hydrogen bonding states and $\chi_3$ angles in these compounds.

Given the estimated homogeneous linewidth, we can roughly calculate the distribution of $\chi_3$ angles of Gln, Q$_3$, and D$_2$Q$_{10}$K$_2$ by using a methodology that is similar to that of Asher et al. [105] To do this, we assume that the inhomogeneously broadened Am III$^P$ bands derive from a distribution of different $\chi_3$ dihedral angles, which can be represented as the sum of $M$ Lorentzian bands:

$$A(\nu) = \frac{1}{\pi} \sum_{i=1}^{M} I_i \frac{\Gamma^2}{\Gamma^2 + (\nu - \nu_i)^2}$$

where $I_i$ is the intensity of a Lorentzian band that occurs at a given center frequency, $\nu_i$, and $\Gamma$ is the homogeneous linewidth.

We can apply eq. 4.3 to correlate the $\nu_i$ Am III$^P$ frequencies of the $M$ Lorentzian bands to their corresponding $\chi_3$ dihedral angles. As shown in Figure 4.6, a single Am III$^P$ frequency can correspond to as many as four possible $\chi_3$ dihedral angles. However, the Shapovalov and Dunbrack database show that $\chi_3$ dihedral angles that are greater than $+90^\circ$ and less than $-90^\circ$ are nearly forbidden (Figure 4.7). Thus, we consider only the two $\chi_3$ dihedral angle solutions that are found in the region between $-90^\circ$ and $+90^\circ$, as shown in Figure 4.9.

To determine which of the two remaining $\chi_3$ dihedral angle solutions is occurring in our peptides, we first fit the histograms to the sum of two Gaussians with identical amplitudes, $A$, and widths, $w$, but different center $\chi_3$ angles, $\bar{\chi}_{3,1}$ and $\bar{\chi}_{3,2}$:

$$I(\chi_3) = Ae^{-(\chi_3 - \bar{\chi}_{3,1})^2} + Ae^{-(\chi_3 - \bar{\chi}_{3,2})^2}$$

The Gln, Q$_3$, and D$_2$Q$_{10}$K$_2$ results all show one Gaussian centered at negative $\chi_3$ angles and another Gaussian centered at positive $\chi_3$ angles (Figure 4.9). For Gln, we assume that the Gaussian centered at $\chi_3 \sim -13^\circ$ is the physically relevant solution based on the neutron diffraction study of Rhys et al. [187]. For Q$_3$ and D$_2$Q$_{10}$K$_2$, we conclude that the Gaussians centered at negative $\chi_3$ angles correspond to the physically relevant solutions to eq. 4.3.
Figure 4.9: $\chi_3$ dihedral angle histograms calculated by decomposing Am III$^P$ bands into a sum of Lorentzians for (a) Gln, (b) Q_3, and (c) D_2Q_{10}K_2 in water. Because the solution to eq. 4.3 is double valued between $\pm 90^\circ$, the histograms show two peaks. The histograms were fit to two identical Gaussians that differed only in center $\chi_3$ dihedral angles (shown in dashed lines). The sum of the Gaussians is shown in the solid red lines. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
because they fall within the range of $\chi_3$ dihedral angles most commonly adopted by Gln residues that populate PPII ($\Phi$, $\Psi$) angles (Figure 4.7c).

Figure 4.10 shows the resulting $\chi_3$ dihedral angle distributions for Gln, $Q_3$, and $D_2Q_{10}K_2$ by assuming the physically relevant solutions to eq. 4.3. The distributions of $Q_3$ and $D_2Q_{10}K_2$ populate $\chi_3$ angles similar to that of Gln. This suggests that primary amides of $Q_3$ and $D_2Q_{10}K_2$ are fully solvated like that of monomeric Gln in water. Thus, the Gln side chains are not engaged in side chain-backbone peptide bond hydrogen bonding as previously hypothesized [33].

4.4.6.3 Determination of the Gibbs Free Energy Landscape for Gln and Gln peptides Along the $\chi_3$ Dihedral Angle Reaction Coordinate The structure sensitivity of the Am III$^P$ band enables us to determine the Gibbs free energy landscape of the Gln side chains along the $\chi_3$ dihedral angle structure coordinate. To do this, we assume that the probability of each $\chi_3,i$ angle in the $\chi_3$ dihedral angle distributions of Gln, $Q_3$, and $D_2Q_{10}K_2$ shown in Figure 4.10a-c is given by a Boltzmann distribution:

$$\frac{p(\chi_3,i)}{p(\chi_3,0)} = e^{-\frac{\Delta G(\chi_3,i)}{RT}}$$ (4.8)

where $p(\chi_3,i)/p(\chi_3,0)$ is the ratio of populations with $\chi_3$ angles $\chi_3,i$ and $\chi_3,0$. The angle, $\chi_3,0$, is the minimum energy $\chi_3$ angle, $R$ is the molar gas constant, $T$ is the experimental temperature (293 K), and $\Delta G(\chi_3,i) = G(\chi_3,i) - G(\chi_3,0)$. We assume in eq. 4.8 that each $\chi_3,i$ dihedral angle state has a degeneracy of one.

To calculate the free energy difference, $\Delta G(\chi_3,i)$, between a particular $\chi_3,i$ angle and the equilibrium $\chi_3,0$ angle, we rearrange eq. 4.8:

$$\Delta G(\chi_3,i) = -RT \ln \left( \frac{p(\chi_3,i)}{p(\chi_3,0)} \right)$$ (4.9)

Figure 4.11 shows the calculated Gibbs free energy landscapes of Gln, $Q_3$, and $D_2Q_{10}K_2$ along the $\chi_3$ dihedral angle structure coordinate. We model the side chain free energies
Figure 4.10: Comparison of $\chi_3$ dihedral angle distributions between (a) Gln, (b) Q$_3$ in a predominately PPII-like conformation, and (c) D$_2$Q$_{10}$K$_2$ in a PPII/2.5$_1$-helix equilibrium. (d) $\chi_3$ angle distribution of Gln residues with PPII-like Ramachandran angles from the Shapovalov and Dunbrack database. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure 4.11: Gibbs free energy landscapes of (a) Gln, (b) Q₃, and (c) D₂Q₁₀K₂. The energy wells can be modeled by assuming a harmonic oscillator model torsional spring:

\[ \Delta G(\chi_3,i) = \frac{\tau}{2} (\chi_{3,i} - \chi_{3,0})^2 \]

where \( \tau \) is the torsional spring force constant and \( (\chi_{3,i} - \chi_{3,0}) \) is the displacement from the equilibrium position. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

about the equilibrium \( \chi_{3,0} \) angles in terms of a simple Hooke’s Law torsional model:

\[ \Delta G(\chi_3,i) = \frac{\tau}{2} (\chi_{3,i} - \chi_{3,0})^2 \]  

(4.10)

where \( \tau \) is the torsional force constant. We can fit the free energy landscapes in Figure 4.11 to eq. 4.10 to determine the torsional force constants along the \( \chi_3 \) dihedral angle coordinate of Gln, Q₃, and D₂Q₁₀K₂. We find that \( \tau \sim 12 \text{ J} \cdot \text{mol}^{-1} \cdot \text{deg}^{-2} \) for Gln, \( \sim 16 \text{ J} \cdot \text{mol}^{-1} \cdot \text{deg}^{-2} \) for Q₃, and \( \sim 13 \text{ J} \cdot \text{mol}^{-1} \cdot \text{deg}^{-2} \) for D₂Q₁₀K₂. The similarity of the Q₃ and D₂Q₁₀K₂ \( \chi_3 \) angle torsional force constants of Gln most likely results from the similar side chain constraints and solvation states of these compounds.
4.5 CONCLUSIONS

We determined the dependence of the Am III\(^{P}\) band frequency on the \(\chi_3\) and \(\chi_2\) dihedral angles of Gln and Asn side chains. The Am III\(^{P}\) vibration is complex and consists of \(C_\delta-N_\epsilon\) (\(C_\gamma-N_\delta\)) stretching and \(N_\epsilon H_2\) (\(N_\delta H_2\)) rocking motions that are out-of-phase with \(C_\beta-C_\gamma\) (\(C_\alpha-C_\beta\)) stretching in Gln (Asn). The frequency of the Am III\(^{P}\) vibration shows a cosinusoidal dependence on the \(\chi_3\) and \(\chi_2\) dihedral angles of the Gln and Asn side chains. The structural sensitivity of the Am III\(^{P}\) vibration derives from hyperconjugation between the \(C_\beta-C_\gamma\) (\(C_\alpha-C_\beta\)) \(\sigma\) and the \(C_\delta=O_\epsilon\) (\(C_\gamma=O_\delta\)) \(\pi^*\) orbitals. Hyperconjugation between these two orbitals increases the \(C_\beta-C_\gamma\) (\(C_\alpha-C_\beta\)) bond length, which decreases the \(C_\beta-C_\gamma\) (\(C_\alpha-C_\beta\)) stretching force constant and causes a downshift in the Am III\(^{P}\) frequency. In this case, hyperconjugation gives rise to spectroscopic markers diagnostic of local dihedral angles. This suggests that future studies of conformationally dependent hyperconjugation interactions will enable the discovery of new, structurally sensitive spectroscopic makers.

The correlations between the Am III\(^{P}\) frequency and the \(\chi_3\) and \(\chi_2\) dihedral angles of Gln and Asn side chains will be useful for protein conformational investigations, particularly for amyloid-like fibril and prion aggregates. In general, fibril and prion aggregates are insoluble and cannot be crystallized. Therefore, there are few approaches to obtain molecular-level structural information. As a result, little is known about the structure of Gln and Asn side chains in fibrils. The Am III\(^{P}\) spectroscopic marker band enables us to experimentally probe conformations of the Gln side chains of polyQ fibrils in order to obtain new, molecular-level insights into fibril structures.

4.6 ACKNOWLEDGMENT

Funding for this work was provided by the University of Pittsburgh. E.M.D. gratefully acknowledges support through the NIH Molecular Biophysics and Structural Biology Training Grant (T32 GM88119). The computational work was supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources pro-
vided. We thank Dr. Sergei V. Bykov for useful discussions, and Prof. Roland Dunbrack, Jr. for providing his rotamer database.
Understanding the structure of polyQ amyloid-like fibril aggregates is crucial to gaining insights into the etiology of at least ten neurodegenerative disorders, including Huntington’s disease. Here, we determine the structure of $D_2Q_{10}K_2$ (Q10) fibrils using UVRR spectroscopy and Molecular Dynamics (MD). Using UVRR, we determine the fibril peptide backbone $\Psi$ and Gln side chain $\chi_3$ dihedral angles. We find that most of the fibril peptide bonds adopt antiparallel $\beta$-sheet conformations; however, a small population of peptide bonds exist in parallel $\beta$-sheet structures. Using MD, we simulate three different potential fibril structural models that consist of either $\beta$-strands or $\beta$-hairpins. Comparing the experimentally measured $\Psi$ and $\chi_3$ angle distributions to those obtained from the MD simulated models, we conclude that the basic structural motif of Q10 fibrils is an extended $\beta$-strand structure.
Importantly, we determine from our MD simulations that Q10 fibril antiparallel $\beta$-sheets are thermodynamically more stable than parallel $\beta$-sheets. This accounts for why polyQ fibrils preferentially adopt antiparallel $\beta$-sheet conformations, instead of in-register parallel $\beta$-sheets like most amyloidogenic peptides. In addition, we directly determine, for the first time, the structures of Gln side chains. Our structural data give new insights into the role that the Gln side chains play in the stabilization of polyQ fibrils. Finally, our work demonstrates the synergistic power and utility of combining UVRR measurements and MD modeling in order to determine the structure of amyloid-like fibrils.

5.1 INTRODUCTION

There are at least ten neurodegenerative disorders, including Huntington’s disease, that are associated with mutational expansions in genomic CAG codon repeats [1]. These expansions increase the length of polyQ repeats in proteins. The increase in the repeat length of polyQ segments greatly enhances protein misfolding and aggregation. Although the exact mechanism of neurotoxicity is still heavily debated, the pathological hallmark of all CAG repeat diseases is the formation of large neuronal inclusions composed of polyQ-rich aggregates [67, 192, 193]. Given their potential role in neurotoxicity, it is therefore crucial to understand the structure of polyQ-rich aggregates.

Numerous structures have been proposed for polyQ fibrils based on the results of many different biophysical methods. For example, X-ray diffraction studies indicate that polyQ fibrils of various Gln repeat lengths all show similar X-ray diffraction patterns. Despite this, these studies assign very different structures from surprisingly similar X-ray data, including “polar-zippers” [52], $\beta$-helices [53], and canonical $\beta$-sheet structures [54, 55]. More recent structural studies of polyQ fibril aggregates use solid state NMR. One of these studies [56], concludes that the basic structural motif of polyQ fibrils prepared from pathologically relevant peptides is a “$\beta$-arc,” similar to that of A$\beta$ [57, 194]. The $\beta$-arc model has been challenged by other solid state NMR and biochemical studies [58, 59], which alternatively suggest that polyQ fibrils are composed of extended $\beta$-strands that contain reverse hairpin
In recent years, several Molecular Dynamics (MD) approaches have also been used to investigate the structural properties of polyQ-rich fibrils. The bulk of these computational studies focus mainly on the kinetic or thermodynamic stability of different fibril structures [51, 63, 66, 195]. However, these MD studies are conducted independently of experimental studies. Thus, there is little direct validation of the computational results against experimental data.

The lack of consensus regarding the structure of polyQ fibrils underscores the need for new and incisive biophysical methods that can quantitatively discriminate between the many proposed models. A fundamental factor in understanding polyQ fibrils is determining the structures and hydrogen bonding environments of the Gln side chains, which are thought to play an important role in stabilizing the aggregates. Another important structural property to understand is the propensity of polyQ fibrils to adopt antiparallel $\beta$-sheets instead of in-register parallel $\beta$-sheets, like most amyloidogenic peptides and prions.

UVRR spectroscopy is a powerful biophysical tool for studying the conformational ensembles and aggregation dynamics of amyloidogenic peptides [90–95]. An advantage of UVRR is that quantitative information can be obtained quickly, under dilute conditions, and without the need for extensive or complex isotopic labeling of peptides and proteins [89]. In recent years, numerous UVRR spectroscopic markers have been discovered. Some of these markers are sensitive to the Ramachandran $\Psi$ angles of the peptide bonds [96, 97], while others are sensitive to the dihedral angles of amino acid side chains [169, 171, 173], including Asn and Gln [180]. Other marker bands are sensitive to the hydrogen bonding and the dielectric environments of peptide bonds and side chains [101, 103, 123, 124, 146, 171].

We can combine structural information obtained from interpreting these spectral markers with results from MD simulations to determine the structure of polyQ and other amyloid-like fibrils. An elegant example of this approach was recently published by Buchanan et al. [197], who combined 2D IR spectroscopy with MD simulations to determine the structure of $K_2Q_{24}K_2W$ fibrils. They concluded that $K_2Q_{24}K_2W$ fibrils adopt an antiparallel $\beta$-sheet structure that contains $\beta$-turns, but not $\beta$-arc structures, after comparing their experimental Amide I spectra with those calculated from simulated models.
In this work, we synergistically couple UVRR and MD in order to determine the structures of polyQ amyloid-like fibrils prepared from the model peptide D$_2$Q$_{10}$K$_2$ (Q10). Xiong et al. [125] previously showed that this peptide can exist in two distinct solution-state conformations, a putative $\beta$-hairpin-like structure (called NDQ10) and a PPII-like structure (called DQ10). We show that both NDQ10 and DQ10 peptide solutions can aggregate into amyloid-like fibrils. We use UVRR to measure the Ramachandran $\Psi$ angle distributions of the NDQ10 and DQ10 fibril peptide bonds, as well as their Gln $\chi_3$ ($O_\epsilon$–C$_\delta$–C$_\gamma$–C$_\beta$) side chain dihedral angles.
To determine the structure of NDQ10 and DQ10 fibrils, we compare our experimentally determined $\Psi$ and $\chi_3$ angle distributions to those obtained from three MD simulated fibril models (Figure 5.1). On the basis of these comparisons, we find that Q10 fibrils consist of extended $\beta$-strands that predominately assemble into antiparallel $\beta$-sheets; however, small populations of Q10 peptide bonds are in parallel $\beta$-sheet structures. From our MD simulations, we determine that polyQ fibrils in antiparallel $\beta$-sheets are lower in free energy compared to parallel $\beta$-sheets. This energetic preference appears unique for polyQ fibrils compared to typical amyloid-like fibrils, where in-register parallel $\beta$-sheets are thought to be at lower energy [198]. Finally, our $\chi_3$ dihedral angle measurements and MD simulations of Q10 fibrils leads us to propose a new model for the structure of Gln side chains in polyQ fibrils.

5.2 EXPERIMENTAL SECTION

5.2.1 Materials

The 14 residue peptide, $D_2Q_{10}K_2$ (Q10), was purchased from AnaSpec Inc. at $\geq 95\%$ purity. Trifluoroacetic acid (TFA) was purchased from Acros at 99.5\% purity, and 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was obtained from Fluka at $\geq 99\%$ purity. HPLC-grade $H_2O$ was purchased from Fisher Scientific, and $D_2O$ (99.9 atom \% D) was purchased from Cambridge Isotope Laboratories, Inc. NaOD (40 wt \% solution in $D_2O$, 99+ atom \% D) and L-glutamine (99\% purity) was purchased from Sigma Aldrich.

5.2.2 Sample Preparation

NDQ10 peptide solutions were prepared by dissolving Q10 directly in $H_2O$ or $D_2O$. DQ10 solutions were prepared using a standard protocol based on a method developed by Wetzel and coworkers [77]. Briefly, DQ10 samples were prepared by suspending the lyophilized Q10 peptide powder received from AnaSpec Inc. in a 1:1 (v/v) mixture of TFA and HFIP. The samples were sonicated for 20 min and incubated at room temperature for 2 h. The solvents
were evaporated under a gentle stream of dry N\textsubscript{2} gas for 1 h. The peptide film was dissolved in H\textsubscript{2}O or D\textsubscript{2}O and ultracentrifuged at 627,000 \times g for 30 min. at 4 °C. The top 2/3 of the solution was decanted and used for the aggregation reaction.

Fibril aggregates were prepared by initially dissolving NDQ10 and DQ10 prepared peptides in H\textsubscript{2}O or D\textsubscript{2}O at 5 mg mL\textsuperscript{-1} concentrations and incubating the samples at 60 °C in vials sealed with Teflon tape. Solutions were titrated to pH 7 (pD 8) using NaOH (NaOD) solutions. After incubation for 6 days, aggregates were harvested via centrifugation, and the pellets resuspended in 120 μL of H\textsubscript{2}O or D\textsubscript{2}O. For the hydrogen-deuterium exchange (HX) experiments, harvested aggregates prepared in H\textsubscript{2}O (D\textsubscript{2}O) were washed in 500 μL D\textsubscript{2}O (H\textsubscript{2}O), centrifuged, and the supernatant removed. The pellets were resuspended in 120 μL of D\textsubscript{2}O (H\textsubscript{2}O) and incubated for 3 h at room temperature in sealed vials.

5.2.3 Transmission Electron Microscopy (TEM)

A 10 μL aliquot of aggregate solution was placed onto a freshly glow discharged carbon-coated grid for 2 min before blotting dry with filter paper. Samples were stained with 10 μL of 1% (w/v) uranyl acetate for 2 min, and the excess stain removed by blotting the grid. Grids were imaged using a Tenai F20 electron microscope (FEI Co.) operating at 200 kV and equipped with a 4k × 4k CCD camera (Gatan).

5.2.4 X-ray Diffraction of Fibril Films

Aggregates prepared from the NDQ10 peptide solutions were placed into the wide end of a 0.7 mm diameter quartz X-ray capillary tube (Charles Supper Company). The wide end of the tube was sealed with melted beeswax and the thin end of the tube left open for drying. DQ10 aggregates were mounted on the broken end of a quartz capillary tube and dried. Aggregate samples were placed on the end of a goniometer head, and centered in the X-ray beam path. X-ray diffraction patterns were collected using a Bruker X8 Prospector Ultra with a Copper micro-focus tube (λ = 1.54178 Å) and equipped with an Apex II CCD detector. All data were collected at room temperature with exposure times of 60 min.
5.2.5 UVRR spectroscopy

The UVRR instrumentation used is described in detail by Bykov et al. [131] Briefly, ∼204 nm light was obtained by mixing the third harmonic with the 816 nm fundamental generated by a tunable Ti:sapphire system (Photonics Industries) operating at 1 kHz. An Indigo S tunable Ti:sapphire system (Positive Light), operating at 5 kHz, generated ∼197 nm light by mixing the third harmonic with the ∼788 nm fundamental. For fibrils measured in solution, the laser light was focused onto a spinning Suprasil quartz NMR tube containing the sample. The average laser power at the sample ranged from ∼0.4 - 0.5 mW. A ∼165° backscattering geometry was used. The total acquisition time to collect spectra was only ∼10 min. The scattered light was imaged into a home-built subtractive double monochromator and detected with a liquid N₂ cooled, back-thinned Spec-10:400B CCD camera (Princeton Instruments) with a Lumogen E coating. The spectrometer resolution was ∼5 cm⁻¹ at the excitation wavelengths used. A description of how the spectra were processed is in APPENDIX D.

5.3 COMPUTATIONAL SECTION

We considered three model fibril systems, as shown in Figure 5.1. These systems are composed of eight Q10 peptides that are assembled into parallel and antiparallel β-sheets. The two β-sheets were oriented parallel to each other, but rotated by 180°, in order to maximize attractive electrostatic interactions between terminal Asp and Lys residues. Models a and b were constructed using canonical β-sheet Ramachandran dihedral angles [199]. In the case of model a, we used canonical antiparallel β-sheet (Φ, Ψ) angles of (−140°, 135°). For model b, we used canonical parallel β-sheet (Φ, Ψ) angles of (−120°, 113°). Model c was constructed using β-hairpin geometries observed in metadynamics simulations (data not shown). These model fibrils were constructed using the Molecular Operating Environment (MOE 2013.10) software suite [199] and were solvated in a water box. Files containing the initial structure coordinates used in the MD simulations, as well as the NAMD configuration
templates, are available for download.

After constructing the models, the solvated fibril systems were energy minimized for 10,000 steps using the conjugate gradient method and then equilibrated for 50 ps. During equilibration, the fibril atoms were initially restrained by harmonic potentials. After relaxing the water around the restrained fibrils, the water molecules and peptide side chains were then energy-minimized for 10,000 steps. This was followed by 100 ps of equilibration, in which the peptide backbone atoms were restrained. The model fibril systems were then equilibrated for an additional 50 ns without restraints. All fibril models retained structural integrity throughout the energy minimization and equilibration. The fibril models were simulated using classical MD for 200 ns.

MD simulations were performed with the NAMD software package (version 2.10) [200]. The potential energies were calculated with the CHARMM22/CMAP force field [201]. This force field was chosen for its torsional energy corrections intended to decrease α-helix bias and stabilize β-strand secondary structures. Other force fields, such as Amber99fs [202], also implement these corrections; however, CHARMM22/CMAP [201] was selected due to its reported accuracy in describing α-helix, β-sheet, and disordered structures [203]. The fibril models were solvated using the solvate module of VMD 1.9.1 [204], resulting in a periodic box of 5,000 water molecules with dimensions of 70 × 50 × 50 Å³ for a total system size of 17,189 atoms. The TIP3P water model [205] was employed in all simulations, and the particle mesh Ewald algorithm [206] was used with a grid spacing of 1.0 Å to calculate full system electrostatics. An integration time step of 2 fs was employed. Simulations were performed under the NPT ensemble, with a Langevin thermostat and piston utilized to regulate the temperature of 300 K and pressure of 1.01325 bar, respectively. The pair interaction cutoff was 12.0 Å, and the switch distance was 10.0 Å.

Simulation analysis was performed using VMD 1.9.1 and Tcl scripting. The Ψ and χ₃ dihedral angles were obtained for all Gln residues with a Tcl script in VMD. The χ₃ dihedral angles were obtained from Gln side chains that were not significantly solvent exposed, so as to best simulate the interior of a fibril environment. The extent of β-sheet dissociation was determined qualitatively, by viewing the trajectories and monitoring the dissociation of peptides. This qualitative analysis was also paired with a quantitative root mean square
deviation (RMSD) metric (see APPENDIX D for details).

The Gibbs free energy difference between the Figure 5.1 model a antiparallel and model b parallel β-strand fibril structures was calculated using a Python implementation of the Bennett acceptance ratio method, called Pymbar [207]. Potential energies were obtained from the NAMD log output files and used as inputs for Pymbar via a Python script that can be found in APPENDIX D.

Hydrogen bonding analysis on models a and b (from Figure 5.1) was also done by using VMD 1.9.1. Hydrogen bond contacts were defined by a heavy atom (N - - - O) distance of <3.0 Å and a N–H O angle between −30° and 30°. The number of peptide backbone-backbone, backbone-side chain, side chain-side chain, and peptide-water hydrogen bonds were calculated for models a and b over the course of the entire trajectory. The data for each hydrogen bond category was binned for each model and plotted with the statistical computing package R (see APPENDIX D).

5.4 RESULTS AND DISCUSSION

5.4.1 Q10 Forms Amyloid-like Fibril Aggregates

Figure 5.2a,b show the TEM images of NDQ10 and DQ10 aggregates. The TEM images of both NDQ10 and DQ10 aggregates resemble those of amyloid-like fibrils and exhibit morphologies similar to those of polyQ peptides with larger, more pathologically relevant Gln repeat lengths [32]. NDQ10 aggregates (Figure 5.2a) cluster into dense meshworks composed of small, thin fibrils, while DQ10 (Figure 5.2b) forms long, ribbon-like fibrils.

We used X-ray diffraction to further characterize the fibril-like nature of NDQ10 and DQ10 aggregates. Oriented amyloid fibril films exhibit a characteristic diffraction pattern with a “meridional reflection” at ∼4.8 Å and an “equatorial reflection” at ∼10 Å to 12 Å. This diffraction pattern is the hallmark of cross-β structures, wherein constituent β-strands orient perpendicular to the long axis of the fibril. The meridional reflection is indicative of the spacing between β-strands, and the equatorial reflection is indicative of the spacing
Figure 5.2: Transmission electron micrographs (TEM) and X-ray diffraction patterns of NDQ10 and DQ10 fibril aggregates. TEM of (a) NDQ10 and (b) DQ10 fibrils. The scale bars represent 200 nm. X-ray diffraction pattern of (c,d) NDQ10 fibril films, (e) DQ10 fibril films, and (f) the quartz capillary tube. The patterns shown in (c) and (d) are the same except for the contrast settings, which have been set in (d) to highlight the weak reflection at \( \sim 8.2 \) Å. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

between \( \beta \)-sheets in the fibrils.

Figure 5.2c-e show the “powder-like” diffraction patterns of unoriented NDQ10 and DQ10 aggregate films. There are reflections at \( \sim 4.1 \) Å, \( \sim 4.8 \) Å, and \( \sim 8.2 \) Å observed in the diffraction patterns of NDQ10 and DQ10. Similar reflections have been observed in larger sequences of polyQ fibrils in previous studies [54, 55]. Based on these previous studies, we assign the prominent \( \sim 4.8 \) Å reflection to the repeat distance between neighboring \( \beta \)-strands within the fibril \( \beta \)-sheets. The weaker \( \sim 8.2 \) Å reflection is assigned to the inter-\( \beta \)-sheet stacking repeat distance.

The equatorial reflections of most amyloid-like fibrils are very diffuse, which indicates
limited ordering and disordered growth in the inter-β-sheet dimension [208]. In contrast, polyQ fibrils usually show very sharp reflections between ~8 Å to 9 Å. We observe reflections at ~8.15 Å and ~8.23 Å in NDQ10 and DQ10, respectively. A higher order reflection occurs at ~4.1 Å for both NDQ10 and DQ10. The presence of these higher orders and the narrowness of the ~8.2 Å reflections suggests that there is long range ordering in the inter-β-sheet dimension of polyQ fibrils [54, 55]. As noted by Atkins and Sikorski [54], this long range ordering presumably arises from the tight interdigitation of the Gln side chains from neighboring sheets. In addition, there is strong inter-amide hydrogen bonding interactions between neighboring side chains of the same β-sheet.

5.4.2 UVRR of Polyglutamine Fibrils in H₂O

We utilized UVRR to investigate the molecular structure of NDQ10 and DQ10 fibrils in solution. The ~197 nm and ~204 nm excited UVRR spectra of NDQ10 and DQ10 fibrils are shown in Figure 5.3. Raman excitation at ~200 nm occurs within the NV₁ electronic transitions of secondary amide peptide bonds and to the long wavelength side on the NV₁ transitions of Gln side chain primary amide groups [125, 134]. Thus, the ~200 nm excitation UVRR spectra of polyQ peptides are dominated mainly by resonance enhanced bands that derive from primary and secondary amide (Am) vibrations, which we label with the superscripts P and S, respectively.

5.4.2.1 Assignment of Gln Side Chain Bands The main difference between the ~197 nm and ~204 nm excited UVRR spectra are the relative intensities of the primary (side chain) and secondary (peptide bond) amide bands. We previously showed that excitation at ~197 nm enhances primary amide UVRR bands significantly more than ~204 nm excitation [125]. As a result, primary amide vibrations can be highlighted by calculating the difference spectrum between the ~197 nm and ~204 nm UVRR spectra.

Figure 5.3 shows that the 197 nm – 204 nm difference spectra highlights the primary amide vibrations of the Gln side chains in NDQ10 and DQ10 fibrils. The AmI P (predominately a Cδ=Oε₁ stretching vibration) and the AmII P (mainly Nε2H₂ scissoring) bands are
Figure 5.3: UVRR spectra (197 nm and 204 nm excitation) of (a) NDQ10 and (b) DQ10 fibrils prepared in H$_2$O. The spectra were measured on precipitates that were resuspended in H$_2$O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

located at $\sim$1660 cm$^{-1}$ and $\sim$1615 cm$^{-1}$, respectively, for both NDQ10 and DQ10 fibrils. In addition, both NDQ10 and DQ10 show a band at $\sim$1100 cm$^{-1}$ that derives from an in-phase combination of C$_\delta$–N$_{\epsilon2}$ stretching and N$_{\epsilon2}$H$_2$ rocking motions. As discussed in detail below, this vibration, which we call the AmIII$^P$, is sensitive to the structure of the Gln side chains.

The primary amide band frequencies and Raman cross sections are very sensitive to
the hydrogen bonding and the local dielectric environment of Gln side chains [146]. For example, the AmIP band frequency is diagnostic of $C_\delta$=O$_{\epsilon 1}$ hydrogen bonding. The AmIP$^P$ band frequency reports on hydrogen bonding of the N$_{\epsilon 2}$H$_2$ group. Compared to monomeric Gln in water [180], the NDQ10 and DQ10 fibril AmIP$^P$ bands are downshifted by $\sim$20 cm$^{-1}$, while the AmIP$^P$ bands are downshifted by $\sim$7 cm$^{-1}$. These frequency downshifts indicate that the inter-amide hydrogen bonding within the fibrils is much stronger than the amide-H$_2$O hydrogen bonding that occurs for monomeric Gln. In addition, the NDQ10 and DQ10 fibril AmIP$^P$ and AmII$^P$ bands show very narrow linewidths, similar to those seen in UVRR spectra of Gln crystals [180], which indicates that the primary amide groups are in very well-defined hydrogen bonding states.

The most noticeable difference between the NDQ10 and DQ10 fibril 197 nm - 204 nm difference spectra occur for strong bands located between $\sim$1400 cm$^{-1}$ to 1500 cm$^{-1}$. A band located at $\sim$1415 cm$^{-1}$ in NDQ10 fibrils upshifts $\sim$15 cm$^{-1}$ in DQ10 fibrils. This band derives from a complex vibration that contains CH$_2$ wagging, C–C stretching, CH$_2$ scissoring, and C$_\delta$–N$_{\epsilon 2}$ stretching motions. Based on our previous work [146], an upshifted CH$_2$ wagging band signals that the Gln side chains are in an environment of higher dielectric constant. Thus, we conclude that the DQ10 Gln side chain methylene groups are in a higher dielectric constant environment than are those of NDQ10. This presumably correlates with the $\sim$0.8 Å larger inter-β-sheet spacing for DQ10 fibrils compared to NDQ10 fibrils (Figure 5.2c-e). The larger inter-sheet spacing allows solvating water molecules to penetrate deeper into DQ10 fibrils, increasing the local dielectric constant. This hypothesis is supported by the results shown in Figure D3, wherein the CH$_2$ wagging band downshifts to $\sim$1415 cm$^{-1}$ and $\sim$1407 cm$^{-1}$ in dried DQ10 and NDQ10 fibril films, respectively.

5.4.2.2 Assignment of Peptide Backbone Bands We subtracted the 197 nm - 204 nm difference spectra from the 204 nm excited UVRR spectra in order to reveal the peptide bond secondary amide bands. The AmIS (mainly peptide backbone C=O stretching) appears as two bands (labelled as AmIS$_A$ and AmIS$_B^2$) in the spectra of both NDQ10 and DQ10 fibrils. This “excitonic splitting” is diagnostic of β-sheet structures and derives from through-space transition dipole coupling between the AmIS oscillators. The Raman spectral AmIS splitting
patterns of antiparallel and parallel $\beta$-sheets are similar. Thus, it is usually difficult to discriminate between these two structures using this band alone [209, 210].

The intense, high frequency AmI$^S$ band (labelled as the AmI$^S_A$) appears at $\sim 1665$ cm$^{-1}$ and $\sim 1660$ cm$^{-1}$ in NDQ10 and DQ10 fibrils, respectively. The less intense, low frequency AmI$^S$ band (AmI$^S_B2$) appears at $\sim 1620$ cm$^{-1}$ in NDQ10 and at $\sim 1615$ cm$^{-1}$ in DQ10. These $\sim 5$ cm$^{-1}$ decreases in the AmI$^S$ mode frequencies suggest slightly stronger peptide backbone C=O hydrogen bonding between $\beta$-strands in DQ10 fibrils than in NDQ10 fibrils [98]. We are, however, aware that the Raman and IR AmI$^S$ bands can also be impacted by $\beta$-sheet twisting and stacking, as well as the registry of the $\beta$-strands, as described in detail by Keiderling and coworkers [209]. Thus, these AmI$^S$ frequency differences between NDQ10 and DQ10 could also signal subtle differences in the twisting and stacking of the fibril $\beta$-sheets. We are continuing to examine these issues.

The extended AmIII$^S$ UVRR spectral region between 1200 cm$^{-1}$ to 1350 cm$^{-1}$ is generally considered to be the most structurally informative [89]. This region in polyQ peptides is complicated due to the overlap of bands from Gln side chain CH$_2$ twisting and wagging modes that occur between $\sim 1280$ cm$^{-1}$ to 1350 cm$^{-1}$. In addition, the AmIII$^S$ region consists of three sub-bands (called the AmIII$^S_1$, AmIII$^S_2$, and AmIII$^S_3$), which derive from vibrations that are composed of in-phase combinations of peptide bond N–H in-plane bending and C–N stretching motions.

Mikhonin et al. [127] previously assigned the AmIII$^S_1$, AmIII$^S_2$, and AmIII$^S_3$ bands in detail. In NDQ10 and DQ10 fibrils, the AmIII$^S_1$ occurs at $\sim 1315$ cm$^{-1}$, while the AmIII$^S_2$ occurs at $\sim 1290$ cm$^{-1}$, and the AmIII$^S_3$ occurs between $\sim 1200$ cm$^{-1}$ to 1280 cm$^{-1}$. Both NDQ10 and DQ10 show peaks at $\sim 1230$ cm$^{-1}$ and low-frequency shoulders at $\sim 1210$ cm$^{-1}$. As discussed in detail below, the AmIII$^S_3$ band frequency is sensitive to the Ramachandran $\Psi$ dihedral angle, and thus can be used to obtain quantitative information on the secondary structure of polyQ fibrils.
Figure 5.4: UVRR spectra (204 nm excitation) of (a) NDQ10 and (b) DQ10 fibrils prepared in D$_2$O. The spectra were measured on precipitates that were resuspended in D$_2$O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

5.4.3 UVRR of Polyglutamine Fibrils in D$_2$O

Figure 5.4 shows the UVRR spectra of NDQ10 and DQ10 fibrils prepared and measured in D$_2$O. Deuteration of the polyQ peptide backbone N–H and primary amide side chain N$_\epsilon$H$_2$ groups leads to significant spectral changes. Upon N-deuteration, the AmI$^S$ downshifts to $\sim$1640 cm$^{-1}$ (AmI$'^S$) and overlaps the AmI$^P$. In the case of the AmIII$^S$ mode, deuteration decouples N–H in-plane bending from C–N stretching [127]. This leaves a weak AmIII$'^S$ band (mainly N–D in-plane bending) in the $\sim$950 cm$^{-1}$ to 1050 cm$^{-1}$ region. The AmIII$^S$ region also contains bands that derive from side chain N$_\epsilon$D$_2$ rocking modes [127]. The loss of the AmIII$^S$ band reveals the presence of several weak bands between $\sim$1300 cm$^{-1}$ to 1400 cm$^{-1}$, which derive mainly from side chain CH$_2$ and peptide backbone C–H deformation modes.
5.4.4 Hydrogen-Deuterium (HX) Exchange of Polyglutamine Fibrils

A comparison of Figure 5.3 and Figure 5.4 show that UVRR can be employed to differentiate between N–H and N–D peptide bonds. Hydrogen-deuterium exchange (HX) can selectively probe solvent exposed versus solvent shielded peptide bonds in fibrils since the AmIII\textsuperscript{S} completely disappears upon N-deuteration [91]. For example, fibrils prepared in H\textsubscript{2}O and subjected to HX in D\textsubscript{2}O will show AmIII\textsubscript{3} bands that derive mainly from peptide bonds shielded from solvent, such as those that are buried within the cross-β core. In contrast, fibrils prepared in D\textsubscript{2}O that are subjected to HX in H\textsubscript{2}O will show AmIII\textsubscript{3} bands that derive mainly from peptide bonds that are solvent accessible, generally because they are located on the aggregate surface or because they exist in exchangeable conformations such as “disordered” regions, turns, or loops.

Figure 5.5: UVRR spectrum (204 nm excitation) of Gln measured in a 50%H\textsubscript{2}O/50%D\textsubscript{2}O mixture. The spectral contributions of solvent, as well as fully protonated and fully deuterated Gln have been subtracted. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

5.4.4.1 UVRR Bands of Partially Deuterated Primary Amides Extensive HX of polyQ fibrils may not completely deuterate the primary amide N\textsubscript{ε}H\textsubscript{2} groups since the Gln side chains may be involved in extensive hydrogen bonding interactions. Based on Saito and coworkers’ normal mode analyses of acetamide [211, 212], the partial deuteration of primary amides results in decoupling of N–H and N–D motions. As shown below, mono-deuteration
of primary amides can give rise to secondary amide-like vibrational modes! Thus, it is conceivable that mono-deuterated Gln side chains can result in AmIII$^S$-like vibrations.

To investigate the potential presence of AmIII$^S$-like vibrations in mono-deuterated primary amides, we measured the UVRR spectrum of Gln in a 50%/50% mixture of H$_2$O and D$_2$O. The spectrum, shown in Figure 5.5, was assigned with the aid of DFT calculations (see APPENDIX D for details). We considered two geometrical isomers in calculating the vibrational normal modes of partially deuterated Gln, as shown in Scheme 5.1. The “trans-NHD” Gln species resembles the trans-isomer configuration of the peptide bond, and thus, is expected to give rise to vibrations that resemble the canonical AmI$^S$, AmII$^S$, and AmIII$^S$ vibrations of secondary amides.

Scheme 5.1: Geometric isomers of the mono-deuterated primary amide group of Gln side chains. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

We present a detailed assignment of the Figure 5.5 spectrum in APPENDIX D, as shown in Tables S1 - S3. According to our normal mode analysis, we assign an AmI$^S$-like vibration to an $\sim$1660 cm$^{-1}$ band, and two AmII$^S$-like vibrations to bands located at $\sim$1550 cm$^{-1}$ and $\sim$1475 cm$^{-1}$. The DFT calculations also indicate that the $\sim$1250 cm$^{-1}$ to 1400 cm$^{-1}$ region contains four vibrations with significant C$_\delta$–N$_\epsilon$ stretchings character. Two of these vibrations derive from the cis-NHD species of Gln and are assigned to the $\sim$1335 cm$^{-1}$ and $\sim$1280 cm$^{-1}$ bands. The other two vibrations at $\sim$1310 cm$^{-1}$ and $\sim$1250 cm$^{-1}$ derive from the trans-NHD species of Gln. These modes resemble AmIII$^S$-like vibrations since they contain a significant combination of NHD scissoring and C$_\delta$–N$_\epsilon$ stretching motions.

5.4.4.2 HX of NDQ10 and DQ10 fibrils Our normal mode analysis and band assignments of the Figure 5.5 Gln spectrum indicate that HX of polyQ fibrils can result in partially deuterated Gln side chains that give rise to AmIII$^S$-like vibrations. These AmIII$^S$-like bands
Figure 5.6: HX-UVRR spectra (204 nm excitation) of (a, b) NDQ10 and (c,d) DQ10 fibrils. For (a, c), fibrils were prepared in H$_2$O and exchanged in D$_2$O. For (b, d), fibrils were prepared in D$_2$O and exchanged in H$_2$O. The spectra were measured on precipitates that were resuspended in either H$_2$O or D$_2$O. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

appear in the high frequency side of the AmIII$_3^S$ region; however, they do not significantly overlap with that of the AmIII$_3^S$ region of β-sheets, which occurs between ∼1200 cm$^{-1}$ to 1240 cm$^{-1}$. Thus, we can straightforwardly assign the β-sheet AmIII$_3^S$ bands of NDQ10 and DQ10 fibrils in the HX-UVRR spectra.

Figure 5.6 shows the curve-resolved AmIII$_3^S$ region of the UVRR spectra of NDQ10 and DQ10 fibrils following HX. The spectra labeled as H→D (D→H) were measured from fibrils prepared in H$_2$O (D$_2$O) and subjected to HX in D$_2$O (H$_2$O). The bands shown in blue are
assigned to true AmIII\textsubscript{3}S vibrations, while those shown in green are assigned to the AmIII\textsubscript{3}-like vibrations that derive from partially deuterated primary amides. The H→D spectra of NDQ10 and DQ10 fibrils show AmIII\textsubscript{3}S bands at ∼1210 cm\textsuperscript{−1} and ∼1230 cm\textsuperscript{−1}. These bands upshift ∼10 cm\textsuperscript{−1} in the D→H spectra. As discussed in detail by Mikhonin et al. [97], this peptide bond frequency upshift can result from an increased hydrogen bonding to water which does not involve any structural changes. Thus, we conclude that the Q10 fibril exterior peptide bonds are more extensively hydrogen bonded to water than those in the fibril interior.

5.4.5 Ramachandran Ψ Angle Distributions

The AmIII\textsubscript{3} band is the most conformationally sensitive secondary amide band because its frequency depends sinusoidally on the Ramachandran Ψ dihedral angle (see Scheme 5.2) of the peptide backbone [96, 97]. This sinusoidal dependence derives from the coupling of the Cα–H bending vibration with the N–H bending component of the AmIII\textsubscript{3} vibration. For example, this coupling is strong for β-strand and PPII-like peptide bond Ψ angles, when the Cα–H and N–H groups are in an approximately cis-configuration. The strong coupling between N–H and Cα–H bending motions downshifts the AmIII\textsubscript{3} frequency. In contrast, for α-helical-like Ψ angles, the Cα–H and N–H groups are in a trans-configuration to each other, which results in the Cα–H and N–H bending motions decoupling. This decoupling results in the AmIII\textsubscript{3} band upshifting.

We utilized the structural sensitivity of the AmIII\textsubscript{3} band to determine Ramachandran Ψ dihedral angle distributions for the NDQ10 and DQ10 fibril peptide bonds (Figure 5.7). To do this, we employed the methodology of Asher and coworkers [97, 105] (see APPENDIX D for details), which correlates the different frequencies of the AmIII\textsubscript{3} band envelope to different peptide bond Ψ angles. This enables us to determine a probability distribution of peptide bond Ψ angles from the inhomogeneously broadened AmIII\textsubscript{3} bandshapes shown in Figure 5.6. The Ψ distributions shown in black derive from the H→D HX-UVRR (Figure 5.6a,c) AmIII\textsubscript{3} band profiles, while those shown in blue are from the D→H HX-UVRR (Figure 5.6b, d) AmIII\textsubscript{3} band profiles.
As shown in Figure 5.7, the $\Psi$ angle distributions of NDQ10 and DQ10 are similar, which indicates that the fibril secondary structures are essentially the same. All the distributions are bimodal, showing peaks near $\sim 145^\circ$ and $\sim 125^\circ$. According to Hovmöller et al.’s analysis [213] of protein data bank crystal structures, Gln residues in antiparallel $\beta$-sheet conformations show an average $\Psi$ angle of $\sim 137(15)^\circ$, while those in parallel $\beta$-sheet structures show an average $\Psi$ angle of $\sim 129(15)^\circ$. Given these values, we attribute the NDQ10 and DQ10 peaks centered at $\sim 145^\circ$ to fibril peptide bonds that are in antiparallel $\beta$-sheet conformations, and the peaks at $\sim 125^\circ$ to peptide bonds in parallel $\beta$-sheet conformations.

A comparison of the black and blue $\Psi$ angle distributions indicates that there are structural differences between solvent accessible and inaccessible peptide bonds in both NDQ10 and DQ10 fibrils. The blue distributions are much broader than the black distributions, which indicates that the solvent accessible peptide bonds exhibit greater conformational heterogeneity than the solvent inaccessible peptide bonds. Most of this increased structural heterogeneity stems from the peptide bonds in parallel $\beta$-sheet conformations. Indeed, the standard deviations ($\sigma$) of the antiparallel $\beta$-sheet distributions corresponding to solvent accessible and inaccessible peptide bonds are not significantly different. However, the standard deviations of the parallel $\beta$-sheet distributions are $\sim 12^\circ$-$13^\circ$ for solvent accessible peptide bonds, but collapse to a narrower range of $\Psi$ angles for solvent inaccessible peptide bonds.
Figure 5.7: UVRR determined $\Psi$ angle distributions for NDQ10 and DQ10 fibrils. (a) Distribution corresponding to the NDQ10 fibril Figure 5.6a HX spectrum. (b) Distribution corresponding to NDQ10 fibril Figure 5.6b HX spectrum. (c) Distribution corresponding to DQ10 fibril Figure 5.6c HX spectrum. (d) Distribution corresponding to DQ10 fibril Figure 5.6d HX spectrum. The distributions were least-squares fit to one or two Gaussians. The peak $\Psi$ angles and distribution standard deviations ($\sigma$) are reported, along with their standard errors from the fits (in parentheses). Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

The $\Psi$ angle distributions shown in Figure 5.7 indicate that the solvent inaccessible peptide bonds in NDQ10 and DQ10 fibrils preferentially adopt antiparallel over parallel
β-sheet conformations. The solvent inaccessible peptide bonds derive primarily from the fibril interiors, where primary fibril nucleation occurs. This suggests that nascent polyQ (proto)fibrils form around antiparallel β-sheet nuclei. In contrast, peptide bonds that are solvent accessible are located predominately on the surface of polyQ aggregates, which is more disordered since peptides can aggregate onto the fibril by forming parallel β-sheet structures, in addition to adding as antiparallel β-sheets.

5.4.6 Developing a Molecular-level Structural Model

![Figure 5.8](image-url)

Figure 5.8: Time evolution of (a) antiparallel β-sheet, (b) parallel β-sheet, and (c) β-hairpin fibril models in MD simulations. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

The Figure 5.7 distributions show that no peptide bonds populate β-turn Ψ angles. Thus, we conclude that the basic structural motifs of NDQ10 and DQ10 fibrils are extended β-strands. To investigate this hypothesis, we utilized atomistic MD to examine Q10 peptides arranged in three different fibril architectures. As shown in Figure 5.1, these models are composed of eight Q10 peptides arranged into two β-sheet layers. Models a and b consist of extended β-strands that are arranged into antiparallel and parallel β-sheet architectures,
respectively. Model c consists of β-hairpins with Type I turn structures that are arranged into an antiparallel β-sheet configuration. Further details regarding the construction of the fibril models for the MD simulations can be found in APPENDIX D.

The Figure 5.1 simulated fibril models retained structural integrity throughout the energy minimization and equilibration processes of the computations. We utilized an RMSD metric (Figure D2) to monitor the extent of dissociation of the fibril models during the simulation production runs. The RMSD of peptide backbone atoms relative to that of the respective initial, energy-minimized structure was used. An RMSD of 3 Å was used as a dissociation threshold. Based on this criterion, the Figure 5.1c β-hairpin model dissociates at ~60 ns into the production simulation, while models a and b remain intact throughout the full 200 ns simulation (Figure 5.8).

As shown in Figure 5.9, we compared the Ramachandran Ψ angle distributions obtained from the production runs of the simulated fibril models to those measured experimentally. The distributions corresponding to models a and b show large peaks at ~141° and ~127°, respectively, which are very close to the experimentally observed Ψ angle peaks for the antiparallel and parallel β-sheets. In contrast, the model c distribution shows very poor agreement with the experimentally determined distributions since the calculated peak Ψ angle distribution is downshifted ~12° from the experimentally observed antiparallel β-sheet peak distribution. In addition, the model c Ψ angle distribution shows a doublet located at ~−19° and ~−43° that is not experimentally observed. This doublet corresponds to Ψ angles that derive from the i+1 and i+2 Type I β-turn residues. The strong agreement between the model a and b Ψ angle distributions with those measured experimentally supports our conclusion that the basic structural motif of NDQ10 and DQ10 fibrils are stacked β-strands organized into β-sheets.

Our β-strand models for NDQ10 and DQ10 fibrils are consistent with other studies. For example, Schneider et al. [56] suggest, on the basis of EM and solid-state NMR, that D2Q15K2 peptides adopt extended β-strands in fibrils. In another study, Thakur and Wetzel [69] probed polyQ fibril structure by replacing Gln-Gln residue pairs with Pro-Gly pairs to increase the formation of β-turn structures. They found that peptides, which had continuous stretches of ~9 to 10 Gln residues, mimicked the aggregation behavior and morphologies of...
unmutated fibrils. They interpret these results to mean that simple polyQ peptides form β-strands that are optimally ~9 to 10 Gln residues in length.
5.4.7 Antiparallel $\beta$-sheets are more Favorable than Parallel $\beta$-sheets in polyQ fibrils

Numerous studies have shown that the most common structural motif of typical prions and amyloid fibrils is the in-register parallel cross-$\beta$ structure \[81, 82, 167, 168, 214\]. For these fibrils, in-register parallel $\beta$-sheets appear to maximize hydrophobic and steric zipper interactions \[198\]. Reports of antiparallel $\beta$-sheets structures are far less common, although they have been reported for some amyloidogenic peptide microcrystals \[215\] and fibrils \[83\]. PolyQ fibrils are unique in that most studies \[55, 56, 58, 69\] suggest that they predominately form antiparallel $\beta$-sheets.

This preference indicates that antiparallel $\beta$-sheet structures is be more energetically favorable than parallel $\beta$-sheets in polyQ fibrils. To examine this possibility, we used the Bennett acceptance ratio method (see APPENDIX D, Figure D4) to calculate the free energy difference between our simulated parallel and antiparallel $\beta$-sheet fibrils. The free energy of our simulated antiparallel $\beta$-sheet model system (model $a$) was found to be $160.5(20) \text{kJ mol}^{-1}$ (per fibril system) lower than the parallel $\beta$-sheet model system (model $b$). Since the simulated fibril systems were composed of eight Q10 peptides, this means that the free energy is $\sim 1.5 \text{kJ mol}^{-1}$ per peptide bond (there are 104 total peptide bonds in each simulated fibril system) lower for the antiparallel $\beta$-sheet structure than the parallel $\beta$-sheet structure.

The MD simulation results suggest that antiparallel $\beta$-sheet formation is favored in polyQ fibrils. This accounts for the greater fraction of antiparallel $\beta$-sheets over parallel $\beta$-sheets experimentally observed in NDQ10 and DQ10 fibrils. Using the integrated areas of the AmIII$_3^S$ bands in the Figure 5.6 spectra, we calculate the apparent Gibbs free energy difference between parallel and antiparallel $\beta$-sheets to be $\sim 6-7 \text{kJ mol}^{-1}$ per solvent inaccessible peptide bond and $\sim 1 \text{kJ mol}^{-1}$ per solvent accessible peptide bond. We note that the experimentally measured free energy difference of $1 \text{kJ mol}^{-1}$ per solvent accessible peptide bond is very close to the value of $\sim 1.5 \text{kJ mol}^{-1}$ per peptide bond calculated from the simulated, well-hydrated fibril models.

We examined electrostatic and hydrogen bonding interactions of models $a$ and $b$ in order to understand the origins of the energetic favorability of antiparallel $\beta$-sheets over parallel
\(\beta\)-sheets. In the case of electrostatics, we find that models \(a\) and \(b\) are both stabilized by favorable electrostatic interactions between N-terminal Asp residues and C-terminal Lys residues from opposing \(\beta\)-sheets. However, within a \(\beta\)-sheet, the antiparallel \(\beta\)-sheet model \(a\) is stabilized by favorable inter-strand electrostatic interactions between oppositely charged terminal residues. In contrast, the parallel \(\beta\)-sheet model \(b\) is slightly destabilized due to inter-strand electrostatic repulsions of similarly charged terminal residues. The inter-strand repulsion between like-charged terminal residues in model \(b\) may disrupt nearby peptide backbone hydrogen bonding interactions.

Most of the energetic favorability of the antiparallel \(\beta\)-sheet derives from hydrogen bonding interactions. In our analysis, we considered three different types of peptide-peptide hydrogen bonding: peptide backbone-backbone, side chain-backbone, or side chain-side chain (see Figure D5 and Table D4 in APPENDIX D). Our analysis indicates that, on average, model \(a\) forms more peptide-peptide hydrogen bonds than does model \(b\). Specifically, model \(a\) forms significantly more hydrogen bonds between peptide backbone amides than does model \(b\) (Figure D5d). In contrast, model \(b\) forms more side chain-backbone and peptide-water hydrogen bonds than does model \(a\), as shown in Figure D5b,c. Thus, it appears that model \(a\) is stabilized by more peptide-peptide hydrogen bonds, and less destabilized due to fewer peptide-water hydrogen bonds. In contrast, model \(b\) is less energetically favorable due to fewer stabilizing peptide-peptide hydrogen bonds and more destabilizing peptide-water hydrogen bonds. We also note that it is well-known that antiparallel \(\beta\)-sheets are enthalpically more favorable than are parallel \(\beta\)-sheets structures \[216\] due to their optimal hydrogen bonding geometries. Our MD results support the hypothesis that thermodynamics, not kinetics, drive polyQ aggregation into antiparallel \(\beta\)-sheet architectures.

5.4.8 Structure of NDQ10 and DQ10 Gln side chains

5.4.8.1 Determination of Gln Side Chain $\chi_3$ Angle Distributions We recently discovered \[180\] a vibrational spectral marker band that we call the AmIII$^P$ band, which shows a cosinusoidal frequency dependence on the O–C–C–C dihedral angles of Gln and Asn side chains (the $\chi_3$ and $\chi_2$ angles, respectively). The AmIII$^P$ vibration is somewhat
Figure 5.10: Peak fitting of the 197 nm – 204 nm UVRR difference spectra of (a) NDQ10 and (b) DQ10 fibrils. Also shown are the residuals between the fitted and measured spectra. The bands shown in blue are assigned to the AmIII\(^P\) of the Gln side chains. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

reminiscent of the AmIII\(^S\)_3 vibration since it derives from an in-phase combination of C\(_\delta\)–N\(_e\)2 stretching and N\(_e\)2H\(_2\) rocking motions (replacing N–H bending in the AmIII\(^S\)_3\(^\delta\)). However, C\(_\beta\)–C\(_\gamma\) stretching also contributes significantly to the AmIII\(^P\) vibrational potential energy distribution.
The structural sensitivity of the AmIII$^P$ band derives mainly from the hyperconjugation of the $C_\delta=O_{\epsilon 1}$ $\pi^*$ and $C_\beta-C_\gamma$ $\sigma$ orbitals [180]. When hyperconjugation is strong (e.g. at $\chi_3 \sim \pm 90^\circ$) electron density is transferred from the $C_\beta-C_\gamma$ to the $C_\delta=O_{\epsilon 1}$ bond. This elongates the $C_\beta-C_\gamma$ bond and reduces the $C_\beta-C_\gamma$ stretching force constant. As a result, the AmIII$^P$ frequency downshifts. In contrast, in the absence of hyperconjugation (e.g. at $\chi_3 \sim 0^\circ$), the $C_\beta-C_\gamma$ bond length shortens and the AmIII$^P$ frequency upshifts.

To locate the AmIII$^P$ bands of NDQ10 and DQ10 fibrils, we curve-resolved the 197 nm - 204 nm difference spectra. As shown in Figure 5.10, the AmIII$^P$ band is located between $\sim 1050$ cm$^{-1}$ and 1150 cm$^{-1}$. Curve-resolving this spectral region for both NDQ10 and DQ10 reveals four underlying bands located at $\sim 1060$ cm$^{-1}$, $\sim 1100$ cm$^{-1}$, $\sim 1120$ cm$^{-1}$, and $\sim 1140$ cm$^{-1}$. Based on our previous work [180], we assign the 1060 cm$^{-1}$, 1120 cm$^{-1}$, and 1140 cm$^{-1}$ bands to C–N stretching, $C_\beta-C_\gamma$ stretching/$N_{\epsilon 2}H_2$ rocking, and CH$_2$ twisting vibrations, respectively. The AmIII$^P$ band is assigned to the $\sim 1100$ cm$^{-1}$ shoulder feature [180].

We utilized the structural sensitivity of the AmIII$^P$ vibration to determine the distributions of $\chi_3$ dihedral angles for the NDQ10 and DQ10 fibrils. The methodology employed to calculate the $\chi_3$ angle distributions is similar to that used to determine the peptide bond Ramachandran $\Psi$ angle distributions (see APPENDIX D on for details). As shown in Figure 5.11a and b, the $\chi_3$ distributions are doubly peaked since the AmIII$^P$ band frequencies gives rise to two physically possible $\chi_3$ angle solutions (see the discussion of eq. D.8 in APPENDIX D). The AmIII$^P$ band of NDQ10 gives rise to distributions of $\chi_3$ angles centered at $\chi_3 \sim -14^\circ$ or $5^\circ$. The AmIII$^P$ band of DQ10 gives rise to similar $\chi_3$ angle distributions that are centered at $\chi_3 \sim 12^\circ$ or $3^\circ$. The Gaussian-like distributions of both NDQ10 and DQ10 show standard deviations of $\sim 15^\circ$.

We compared the Figure 5.11a,b distributions to those calculated from the MD simulated fibril models. The $\chi_3$ dihedral angle distributions corresponding to the simulated fibril models are shown in Figure 5.11c-e. The antiparallel $\beta$-sheet model $a$ shows a dominating peak at $\sim 4^\circ$, whereas the parallel $\beta$-sheet model $b$ shows its largest peak centered at $\sim 10^\circ$. These dihedral angle maxima of models $a$ and $b$ are very close to the experimentally measured for NDQ10 and DQ10.
Our combined UVRR and validated MD results provide new detailed insights into the structure of the Gln side chains in polyQ fibrils. The experimentally measured $\chi_3$ dihedral angles for both NDQ10 and DQ10 are distributed around $\sim 0^\circ$. As shown by the structures in Figure 5.12a,b, the model $a$ and $b$ Gln side chains are approximately planar. This allows Gln side chains from opposing $\beta$-sheets to form tightly interdigitated steric zippers, as well as to enable the primary amide groups to both accept and donate hydrogen bonds between neighboring $\beta$-strands.

5.4.8.2 Comparisons of Side Chain Structures with Other Models To our knowledge, the structure of Gln side chains in polyQ fibrils have only been investigated previously by the Sikorski [54] and Kirschner [55] groups. Both of these studies examined polyQ fibril structure with X-ray diffraction. However, despite obtaining similar diffraction patterns, the Sikorski [54] and Kirschner [55] groups proposed different fibril structures based upon modeling the data. For example, the Sikorski [54] group proposes that Gln side chains show alternating $\chi_1$ (N–C$_\alpha$–C$_\beta$–C$_\gamma$) torsion angles along each $\beta$-strand of $\sim 69^\circ$ and $\sim -113^\circ$. The Sikorski group’s model [54] is similar to our structure since the Gln side chains are approximately planar; however, their final model indicates that the $\chi_3$ angle is $\sim 180^\circ$, which is energetically unfavorable and sterically nearly forbidden [180].

The Kirschner [55] group proposed a different structure for Gln side chains in polyQ fibrils. In their study, an electron density maps were calculated from their low-resolution powder-like X-ray diffraction patterns by combining the structure factors determined from the experimental reflection intensities, and the phase angles from a model that satisfactorily accounted for the observed d-spacings. They then modeled the electron densities to generate atomic models for Ac$–$Q$_8$–NH$_2$ and D$_2$Q$_{45}$K$_2$ fibrils. Their modeling of the electron density maps suggests that polyQ fibril Gln side chains adopt unusual bent conformations, as shown in Figure 5.12c,d. It should be noted that the fibril models reported by the Kirschner [55] group result in very high R-factors (24% for Ac$–$Q$_8$–NH$_2$ and 35% for D$_2$Q$_{45}$K$_2$).

The Kirschner [55] group’s side chain structures are inconsistent with our model. The Kirschner [55] group structure shows side chain $\chi_3$ angles that approach values of $\sim \pm 90^\circ$, which differs significantly from our experimentally determined values that are close to $\sim 0^\circ$. 
The fact that we measure $\chi_3$ angles near $\sim 0^\circ$ is important because it means that the Gln side chains are roughly planar, which allows the steric zipper interactions that are believed to stabilize amyloid-like fibrils. It is difficult to envision Gln primary amide stabilizing steric zipper interactions in the Kirschner [55] group model, since their Gln side chains conformations are bent, and should not enable tight interdigitation of neighboring $\beta$-sheets.

5.4.9 Dependence of Fibril Structure on Deposition of Different NDQ10 and DQ10 Solution Conformations

The NDQ10 and DQ10 peptides have significantly different solution conformations. The NDQ10 and DQ10 peptide solutions are composed of putative $\beta$-hairpin-like and PPII-like conformations, respectively [125]. These two solution conformations aggregate into fibrils that are composed of similar, but not identical, $\beta$-sheet structures, which have similar planar Gln side chain conformations. The similarity between these fibril structures probably signals that the extended $\beta$-strand structure is the most energetically favorable Q10 fibril structure.

However, the subtle differences observed between the NDQ10 and DQ10 fibril structures must result from the variation in the growth mechanisms due to different Q10 solution conformations. A speculative hypothesis is that the increased water content and spacing of the DQ10 fibrils results from the preferential addition of the well hydrated PPII-like DQ10 peptides that lead to incorporation of water into the hydrophobic interdigitating side chain domains. This gives rise to the longer DQ10 fibril inter-$\beta$-sheet spacing compared to NDQ10, as observed in Figure 5.2. We are presently investigating this possibility.

5.5 CONCLUSION

We performed a detailed structural analysis of NDQ10 and DQ10 fibrils using UVRR and MD simulations. On the basis of comparing our UVRR and MD simulation results, we determine that the basic structural element of Q10 fibrils is an extended $\beta$-strand. The solvent inaccessible interiors of NDQ10 and DQ10 fibrils are a predominately antiparallel $\beta$-sheet
structures that are highly ordered and composed of these extended $\beta$-strands. However, the water accessible peptide bonds, which are located predominately on the fibril surfaces, show greater conformational heterogeneity and contain significant subpopulations of $\beta$-strands that adopt parallel $\beta$-sheet architectures.

Our MD simulation results indicate that Q10 antiparallel $\beta$-sheets are energetically more favorable than parallel $\beta$-sheets. This is an important insight because it may explain why polyQ fibrils, in contrast to most amyloid-like aggregates, preferentially adopt antiparallel $\beta$-sheets instead of in-register parallel $\beta$-sheets. Our results indicate that the origin of the energetic favorability of Q10 fibril antiparallel $\beta$-sheets is mainly due to hydrogen bonding. Antiparallel $\beta$-sheets form, on average, more hydrogen bonds between peptide backbone amides than do parallel $\beta$-sheets.

This study also provides important new insights into the structure and chemical environment of Gln side chains in polyQ fibrils. In contrast to previous, low-resolution X-ray studies, we show that that the Gln side chains in polyQ fibrils adopt conformations that are roughly planar, where the $\chi_3$ dihedral angles are narrowly distributed around 0°. This enables the formation of steric zippers, wherein the side chains of neighboring $\beta$-sheets tightly interdigitate.

The UVRR spectra are also highly sensitive to the local hydrogen bonding and dielectric environments of the Gln side chains. For example, a major difference observed between NDQ10 and DQ10 fibrils is a different local dielectric environment of the Gln side chains. The primary amides of both NDQ10 and DQ10 fibrils are strongly hydrogen bonded; however, in DQ10, the side chain methylene groups experience a higher dielectric constant environment. This is likely correlated with the larger DQ10 fibril inter-sheet spacing compared to NDQ10. The larger inter-sheet spacing of DQ10 fibrils presumably results from an increased content of water between $\beta$-sheets. The NDQ10 and DQ10 fibrils experience different growth processes due to their different Q10 solution conformations. Deposition of the well-hydrated PPII-like DQ10 peptides results in fibrils with higher water content, with a $\beta$-sheet structure showing larger inter-sheet spacings than occurs for growth with $\beta$-hairpin NDQ10 solution peptides.

Finally, our study demonstrates the utility of synergistically coupling UVRR with MD simulations. Understanding the structure of polyQ and other amyloid-like fibrils remains
of great importance because these aggregates are implicated in numerous neurodegenerative
diseases. Knowing the fibril structures will provide important insights into the aggrega-
tion mechanism(s) of polyQ peptide sequences. However, determining the molecular-level
structure of fibrils is challenging because of the insoluble and non-crystalline nature of fibril
aggregates. A key advantage of our approach is that UVRR can be utilized to measure the
peptide backbone $\Psi$ and the side chain $\chi_3$ dihedral angle distributions. This information
can be used to generate hypotheses on the structure of fibrils, which can then be tested in silico
with MD simulations. Our work combines UVRR and MD into a novel approach for
investigating fibrils. Our approach complements existing methods such as solid-state NMR.
However, the short experimental UVRR acquisition time frames of our measurements enable
the examination of peptide structural changes during fibril aggregation on timescales that
are inaccessible to conventional biophysical methods.

5.6 NOTE ADDED IN REVISION

While this paper was in review, Hoop et al. [61] published a study that examined
fibrils prepared from the huntingtin exon 1 domain (htt exon 1). Overall, their solid-state
NMR (ssNMR) measurements on the polyQ fibril core of htt exon 1 are in remarkable
agreement with our UVRR-based measurements on Q10. Their findings suggest that the
polyQ fibril core in htt exon 1 is arranged in $\beta$-hairpins that form antiparallel $\beta$-sheets. Using
sophisticated magic angle spinning ssNMR techniques, Hoop et al. measure Ramachandran
$\Psi$ angles that are very close to our values for antiparallel $\beta$-sheets. This is particularly true
for their observed “b-type” conformer, where they measure $\Psi$ angles of $\sim 150^\circ$.

Hoop et al. also measured the Gln side chain $\chi_2$ dihedral angles ($C_\alpha - C_\beta - C_\gamma - C_\delta$) to
be $\sim 180^\circ$ in htt exon1 fibrils. Their results lead them to also conclude that the Gln side
chains in polyQ fibrils are extended in structure. Although they did not directly measure
Gln side chain $\chi_3$ angles, Hoop et al. suggest values of $\pm 150^\circ$, which differ significantly
from our experimentally determined values reported here. It is interesting to note that our
experimentally validated MD-simulated $\beta$-sheet fibril structure (model a) shows a mean $\chi_2$
value of $\sim 180^\circ$, which is exactly the angle that Hoop et al. measure. Combining these ssNMR results with our UVRR and MD data, leads us to propose that the most likely fibril structures of the Gln side chains will have $\chi_2$ and $\chi_3$ angles of $\sim 180^\circ$ and $\sim 0^\circ$, respectively. We are examining this issue in greater detail.

5.7 ACKNOWLEDGMENTS

Funding for this work was provided by the University of Pittsburgh (D.P., Z.H., S.A.A.) and partially supported by NIH R01 DA027806 (J.D.M. and R.J.W.). The MD simulation computer time was supported by XSEDE MCB060069, and computer equipment was purchased from NSF funds (CHE-1126465 and P116Z080180). The DFT calculations were supported by the University of Pittsburgh Center for Simulation and Modeling through the super-computing resources provided. We thank Dr. Sergei V. Bykov, Liqi Feng, Jonathan Weisberg, and Jonathan Wert for useful discussions. We are also grateful to Dr. Steven Geib and Dr. Alexander Makhov for technical assistance with the X-ray diffraction measurements and electron microscopy, respectively.
Figure 5.11: $\chi_3$ dihedral angle distributions of Gln side chains. Distributions obtained experimentally are shown for (a) NDQ10 fibrils and (b) DQ10 fibrils. Distributions obtained from MD simulated structures correspond to (c) model $a$ in Figure 5.1; (d) model $b$ in Figure 5.1; and (e) model $c$ in Figure 5.1. The missing $\chi_3$ angles in (a,b) around the region of $\sim 0^\circ$ is due to the fact that those corresponding AmIII$^P$ frequencies are outside the domain of the semi-empirically derived equation [180] used to correlate the frequencies of the AmIII$^P$ band envelope to their respective $\chi_3$ angles (see eq. D.8 in APPENDIX D). Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
Figure 5.12: Comparison of side chain geometries from (a,b) our MD simulations and (c,d) Sharma et al. [55] The (c,d) figure panels were adapted from reference [55]. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
APPENDIX A

DERIVATIONS

A.1 DERIVATION OF THE TRANSITION SCATTERING RATE

A.1.1 Second Order Perturbation Theory

In this section, the transition scattering rate for Rayleigh and Raman (eq. 2.1) is derived using second order perturbation theory in the interaction picture of quantum mechanics. We begin by expanding the Dyson series to its second order term [113]:

\[
\hat{U}_I(t,0) = 1 + \left(\frac{1}{i\hbar}\right) \int_0^t dt_1 \hat{H}_I(t_1) + \left(\frac{1}{i\hbar}\right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1) \hat{H}_I(t_2)
\]

\[
= \hat{U}_I^{(0)}(t,0) + \hat{U}_I^{(1)}(t,0) + \hat{U}_I^{(2)}(t,0) \tag{A.1}
\]

where \(\hat{U}_I(t,0)\) is the time evolution operator. The operator, \(\hat{H}_I(t)\), is the interaction picture Hamiltonian [113], which is defined as:

\[
\hat{H}_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{H}'(t) e^{-i\hat{H}_0 t/\hbar} \tag{A.2}
\]

\(\hat{H}_0\) is the time-independent Hamiltonian of the system. \(\hat{H}'(t)\) is the time-dependent Hamiltonian due to the interaction between the molecule being excited and the perturbing radiation field. The perturbation is small and acts over the time interval from \(t = 0\) to \(t\). For simplicity, this perturbation is treated classically. We assume that the perturbation is a linearly
polarized, continuous light wave oscillating at a frequency, \( \omega \), that is given by the following equation:

\[
\vec{E} = E \hat{\epsilon} \cos(\hat{k} \cdot \hat{r} - \omega t)
\]

\[
= \frac{E \hat{\epsilon}}{2} \left[ e^{i(\hat{k} \cdot \hat{r} - \omega t)} + e^{-i(\hat{k} \cdot \hat{r} - \omega t)} \right]
\]

(A.3)

where \( E \) is the amplitude of the electric field, \( \hat{\epsilon} \) is a unit vector specifying the direction of the electric field polarization, \( \hat{k} \) is the propagation vector, and \( \hat{r} \) is the position vector.

Eq. A.3 is further simplified by invoking the electric dipole approximation. This approximation assumes that the wavelength of the perturbing electromagnetic field is much larger than the dimensions of the molecules being excited. This means that the molecules being excited experience an electric field that is spatially constant, so that:

\[
e^{\pm i(\hat{k} \cdot \hat{r})} = 1 \pm i \hat{k} \cdot \hat{r} \pm \ldots \approx 1
\]

(A.4)

\[
\vec{E} \approx \frac{E \hat{\epsilon}}{2} \left[ e^{-i\omega t} + e^{i\omega t} \right]
\]

(A.5)

Given eq. A.5, \( \hat{H}'(t) \) can be be written as the following:

\[
\hat{H}'(t) = -\frac{E_I}{2} \hat{\epsilon}_I \cdot \hat{\mu} e^{-i\omega t} - \frac{E_S}{2} \hat{\epsilon}_S \cdot \hat{\mu} e^{i\omega t}
\]

\[
= V e^{-i\omega t} + V^\dagger e^{i\omega t}
\]

(A.6)

where the subscripts \( I \) and \( S \) denote the incident and scattered photons. The operator, \( \hat{\mu} \), is the transition dipole moment that couples the molecule to the perturbing radiation field.

We assume that the molecule starts in an initial state \( |i\rangle \) and, following the perturbation, ends in state \( |f\rangle \). The probability amplitude associated with this transition is:
\[ \langle f|\hat{U}_I(t, 0)|i\rangle = \langle f|\hat{U}_I^{(0)}(t, 0)|i\rangle + \langle f|\hat{U}_I^{(1)}(t, 0)|i\rangle + \langle f|\hat{U}_I^{(2)}(t, 0)|i\rangle \]

\[ = \delta_{fi} + \left( \frac{1}{i\hbar} \right) \langle f| \int_0^t dt_1 \hat{H}_I(t_1) |i\rangle + \left( \frac{1}{i\hbar} \right)^2 \langle f| \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1) \hat{H}_I(t_2) |i\rangle \]

(A.7)

The Kronecker delta, \(\delta_{fi}\), indicates that the \(\langle f|\hat{U}_I^{(0)}(t, 0)|i\rangle\) term is non-zero only if no transition occurs, \((i.e., i = f)\). In contrast, the \(\langle f|\hat{U}_I^{(1)}(t, 0)|i\rangle\) term represents one-photon absorption and emission processes. Thus, only the \(\langle f|\hat{U}_I^{(2)}(t, 0)|i\rangle\) term of eq. A.7 contributes to the probability amplitude associated with two photon scattering processes such as Rayleigh and Raman:

\[ \langle f|\hat{U}_I^{(2)}(t, 0)|i\rangle = \left( \frac{1}{i\hbar} \right)^2 \langle f| \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{H}_I(t_1) \hat{H}_I(t_2) |i\rangle \]

(A.8)

To evaluate eq. A.8, we use the following identity:

\[ \sum_r |r\rangle \langle r| = 1 \]

(A.9)

Substituting eqs. A.6 and A.9 into A.8, gives the following equation:

\[ \langle f|\hat{U}_I^{(2)}(t, 0)|i\rangle = \left( \frac{1}{i\hbar} \right)^2 \sum_r \int_0^t dt_1 \langle f| \hat{H}_I(t_1) |r\rangle \int_0^{t_1} dt_2 \langle r| \hat{H}_I(t_2) |i\rangle \]

\[ = \left( \frac{1}{i\hbar} \right)^2 \sum_r \int_0^t dt_1 \langle f| e^{i\hat{H}_0 t_1/\hbar} [V e^{-i\omega_1 t_1} + V^\dagger e^{i\omega_s t_1}] e^{-i\hat{H}_0 t_1/\hbar} |r\rangle \]

\[ \times \int_0^{t_1} dt_2 \langle r| e^{i\hat{H}_0 t_2/\hbar} [V e^{-i\omega_1 t_2} + V^\dagger e^{i\omega_s t_2}] e^{-i\hat{H}_0 t_2/\hbar} |i\rangle \]

\[ = \left( \frac{1}{i\hbar} \right)^2 \sum_r \int_0^t dt_1 \left[ \langle f|V|r\rangle e^{i(\omega_{fr} - \omega_r)t_1} + \langle f|V^\dagger|r\rangle e^{i(\omega_r + \omega_s)t_1} \right] \]

\[ \times \int_0^{t_1} dt_2 \left[ \langle r|V|i\rangle e^{i(\omega_r - \omega_l)t_2} + \langle r|V^\dagger|i\rangle e^{i(\omega_r + \omega_s)t_2} \right] \]

(A.10)

where \(\omega_{fr} = \omega_f - \omega_r\) and \(\omega_{fr} = \omega_f - \omega_r\). Integrating eq. A.10 with respect to \(t_2\) gives:
gives the following probability amplitude:

\[
\left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \left[ \frac{\langle f|V|r\rangle \langle r|V|i\rangle e^{i(\omega_f t - 2\omega_I)t_1} - \langle f|V|r\rangle \langle r|V|i\rangle e^{i(\omega_f - \omega_I)t_1}}{i(\omega_r - \omega_I)} \right. \\
+ \frac{\langle f|V|r\rangle \langle r|V|i\rangle e^{i(\omega_f - \omega_I + \omega_S)t_1} - \langle f|V|r\rangle \langle r|V|i\rangle e^{i(\omega_f + \omega_I)t_1}}{i(\omega_r + \omega_S)} \\
+ \frac{\langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f - \omega_I + \omega_S)t_1} - \langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f + \omega_I)t_1}}{i(\omega_r - \omega_I)} \\
+ \frac{\langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f + \omega_S)t_1} - \langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f + \omega_I + \omega_S)t_1}}{i(\omega_r + \omega_S)} \right]
\] (A.11)

where we have utilized the fact that \(\omega_f = \omega_f - \omega_I + \omega_r\).

Expression A.11 is modulated in five frequencies: \(\omega_f - \omega_I + \omega_S\), \(\omega_f - 2\omega_I\), \(\omega_f + 2\omega_S\), \(\omega_f - \omega_I\), and \(\omega_f + \omega_S\). The terms containing \(\omega_f - \omega_I\) and \(\omega_f + \omega_S\) represent transitions that are non-resonant with the initial or final states. Integrating these terms with respect to \(t_1\) averages to zero over short timescales. Thus, the contributions of these terms in A.11 can be neglected.

The term containing \(\omega_f - 2\omega_I\) corresponds to two-photon absorption (Figure 2.3d) since the frequency of the incident photon is half the frequency associated with the transition from the initial to the final state. Similarly, the term containing \(\omega_f + 2\omega_S\) corresponds to a two-photon emission (Figure 2.3e). The only relevant terms for Rayleigh and Raman scattering contain \(\omega_f - \omega_I + \omega_S\). Integrating the terms that contain \(\omega_f - \omega_I + \omega_S\) with respect to \(t_1\) gives the following probability amplitude:

\[
\langle f|\hat{U}^{(2)}_I(t,0)|i\rangle = \left(\frac{1}{i\hbar}\right)^2 \sum_r \int_0^t dt_1 \left[ \frac{\langle f|V|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f - \omega_I + \omega_S)t_1}}{i(\omega_r + \omega_S)} \right. \\
+ \frac{\langle f|V^\dagger|r\rangle \langle r|V^\dagger|i\rangle e^{i(\omega_f - \omega_I + \omega_S)t_1}}{i(\omega_r - \omega_I)} \\
= \frac{1}{\hbar^2} \sum_r \left[ \frac{\langle f|V|r\rangle \langle r|V|i\rangle}{\omega_r - \omega_I} + \frac{\langle f|V|r\rangle \langle r|V^\dagger|i\rangle}{\omega_r + \omega_S} \right] \left[ \frac{e^{i(\omega_f - \omega_I + \omega_S)t} - 1}{\omega_f - \omega_I + \omega_S} \right] \\
= \frac{E_I E_S}{4\hbar^2} \sum_r \left[ \frac{\langle f|\hat{e}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\mu}_\rho \cdot \hat{e}_S|i\rangle}{\omega_r - \omega_I} + \frac{\langle f|\hat{e}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\mu}_\sigma \cdot \hat{e}_I|i\rangle}{\omega_r + \omega_S} \right] \\
\times \frac{e^{i(\omega_f - \omega_I + \omega_S)t} - 1}{\omega_f - \omega_I + \omega_S}
\] (A.12)

where the subscripts \(\rho\) and \(\sigma\) represent the directions of the transition moments.
A.1.2 Fermi’s Golden Rule and the Transition Scattering Rate

The transition probability, $P_{fi}(t)$, of finding the system in state $|f\rangle$ at time $t$ is the square of eq. A.12:

$$P_{i\rightarrow f}(t) = |\langle f | U^{(2)}_I (t, 0) | i \rangle|^2$$

$$= \frac{E_i^2 E_f^2}{16\hbar^4} \sum_r \left| \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | r \rangle \langle r | \hat{I}_I \cdot \hat{\mu}_\sigma | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | i \rangle}{\omega_{ri} + \omega_S} \right|^2$$

$$\times \left| \frac{e^{i(\omega_{fi} - \omega_I + \omega_S)t} - 1}{\omega_{fi} - \omega_I + \omega_S} \right|^2$$

$$= \frac{E_i^2 E_f^2}{16\hbar^4} \sum_r \left| \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | r \rangle \langle r | \hat{I}_I \cdot \hat{\mu}_\sigma | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | i \rangle}{\omega_{ri} + \omega_S} \right|^2$$

$$\times \left[ \frac{4 \sin^2[(\omega_{fi} - \omega_I + \omega_S)t/2]}{2} \right]$$

(A.13)

Eq. A.13 is the probability of a transition to occur between two discrete states, $|i\rangle$ and $|f\rangle$. However, we assume a central tenant of Fermi’s Golden Rule, viz. that the molecule is being excited into a continuum of $|f\rangle$ states that span a range of frequencies. If we assume that, in this continuum, each $|f\rangle$ state is independent of the others, then the total transition probability is simply the sum of the individual transition probabilities:

$$P(t) = \sum_f P_{i\rightarrow f}(t)$$

$$= \sum_f \frac{E_i^2 E_f^2}{16\hbar^4} \sum_r \left| \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | r \rangle \langle r | \hat{I}_I \cdot \hat{\mu}_\sigma | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_\sigma | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_\rho | i \rangle}{\omega_{ri} + \omega_S} \right|^2$$

$$\times \left[ \frac{4 \sin^2[(\omega_{fi} - \omega_I + \omega_S)t/2]}{2} \right]$$

(A.14)

The index, $f$, is continuous. To convert the summation to an integral, we introduce $\rho_f$, the density of states, which is defined as the number of levels per unit energy:

$$\rho_f = \frac{dn_f}{dE_f} = \frac{1}{\hbar} \frac{dn_f}{d\omega_f}$$

(A.15)
where $\rho_f$ is assumed to be a continuous function over $\hbar\omega_f$. Substituting eq. A.15 into A.14 gives the following:

$$
P(t) = \frac{E_i^2 E_S^2}{16\hbar^3} \int_{\omega_i - \Delta \omega/2}^{\omega_i + \Delta \omega/2} d\omega_f \rho_f \sum_r \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} + \omega_S} \bigg|^2 \quad (A.16)
$$

Further simplifying assumptions are necessary in order to integrate eq. A.16. Although $\rho_f$ and the transition moment integrals are functions of $\omega_f$, we assume that they vary slowly, and thus are approximately constant over the range of frequencies being considered. This means that eq. A.16 can be written as:

$$
P(t) = \frac{E_i^2 E_S^2}{8\hbar^3} \rho_f \sum_r \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} + \omega_S} \bigg|^2 \quad (A.17)
$$

To integrate eq. A.17, the following substitution is made:

$$
u = (\omega_{fi} - \omega_I + \omega_S)t/2 \quad (A.18)
$$

$$
P(t) = \frac{E_i^2 E_S^2}{8\hbar^3} \rho_f \sum_r \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} + \omega_S} \bigg|^2 \quad (A.19)
$$

The limits of integration in eq. A.19 can be extended to $\pm \infty$ without serious error so that:

$$
P(t) = \frac{E_i^2 E_S^2}{8\hbar^3} \rho_f \sum_r \frac{\langle f | \hat{\epsilon}_S \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_I \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} - \omega_I} + \frac{\langle f | \hat{\epsilon}_I \cdot \hat{\mu}_r | r \rangle \langle r | \hat{\epsilon}_S \cdot \hat{\mu}_r | i \rangle}{\omega_{ri} + \omega_S} \bigg|^2 \rho_f \quad (A.20)
$$
Eq. A.20 indicates that the total transition probability is linearly proportional to time, so that the scattering rate, $W_{fi}$, can be defined as follows:

$$W_{fi} = \frac{dP(t)}{dt} = \frac{\pi E_I^2 E_S^2}{8\hbar^3} \sum_r \left| \frac{\langle f|\hat{\varepsilon}_S \cdot \hat{\mu}_\rho|r\rangle \langle r|\hat{\varepsilon}_I \cdot \hat{\mu}_\sigma|i\rangle}{\omega_{ri} - \omega_I} + \frac{\langle f|\hat{\varepsilon}_I \cdot \hat{\mu}_\sigma|r\rangle \langle r|\hat{\varepsilon}_S \cdot \hat{\mu}_\rho|i\rangle}{\omega_{ri} + \omega_S} \right|^2 \rho_f$$ (A.21)

### A.2 DERIVATION OF THE ALBRECHT A, B, AND C TERMS

In this section, the Albrecht $A$, $B$, and $C$ terms (eqs. 2.16–2.18) are derived. We begin by using eqs. 2.11 and 2.12 to write the relevant electronic transition moment integrals of eq. 2.10 as:

$$\langle \phi_g | \hat{r}_\rho | \phi_e \rangle \approx \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle + \frac{1}{\hbar} \sum_{a,s \neq e} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{es}} \right) Q_a \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle$$

$$+ \frac{1}{\hbar} \sum_{a,t \neq g} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{gt}} \right) Q_a \langle \phi_t^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle$$ (A.22)

$$\langle \phi_e | \hat{r}_\sigma | \phi_g \rangle \approx \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle + \frac{1}{\hbar} \sum_{a,t \neq g} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{gt}} \right) Q_a \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_t^{(0)} \rangle$$

$$+ \frac{1}{\hbar} \sum_{a,s \neq e} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{es}} \right) Q_a \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle$$ (A.23)

$$\langle \phi_g | \hat{r}_\sigma | \phi_e \rangle \approx \langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_e^{(0)} \rangle + \frac{1}{\hbar} \sum_{a,s \neq e} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{es}} \right) Q_a \langle \phi_g^{(0)} | \hat{r}_\sigma | \phi_s^{(0)} \rangle$$

$$+ \frac{1}{\hbar} \sum_{a,t \neq g} \left( \frac{h_a^{(0)}}{\omega^{(0)}_{gt}} \right) Q_a \langle \phi_t^{(0)} | \hat{r}_\sigma | \phi_e^{(0)} \rangle$$ (A.24)
\[
\langle \phi_e | \hat{r}_\rho | \phi_g \rangle \approx \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle + \frac{1}{\hbar} \sum_{a,t \neq g} \frac{(h_a^{(0)})_{gt}}{\omega_{gt}^0} Q_a \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \\
+ \frac{1}{\hbar} \sum_{a,s \neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} Q_a \langle \phi_s^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle
\]

These expressions can be utilized to write the products of the vibronic transition moment integrals in eq. 2.10 as:

\[
\langle \chi_{gf'} | \langle \phi_g | \hat{r}_\rho | \phi_e \rangle | \chi_{er'} \rangle \langle \chi_{er'} | \langle \phi_e | \hat{r}_\sigma | \phi_g \rangle | \chi_{g'f'} \rangle = \\
\left[ \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \chi_{gf'} | \chi_{er'} \rangle + \frac{1}{\hbar} \sum_{a,t \neq g} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \chi_{gf'} | Q_a | \chi_{er'} \rangle \right] \left[ \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \langle \chi_{er'} | \chi_{g'f'} \rangle + \frac{1}{\hbar} \sum_{a,s \neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \langle \chi_{er'} | Q_a | \chi_{g'f'} \rangle \right] \\
\approx \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | \chi_{g'f'} \rangle + \\
+ \frac{1}{\hbar} \sum_{a,t \neq g} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_g^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_t^{(0)} \rangle \langle \partial H / \partial Q_a | \phi_t^{(0)} \rangle \langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | Q_a | \chi_{g'f'} \rangle \\
+ \frac{1}{\hbar} \sum_{a,s \neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \partial H / \partial Q_a | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \langle \chi_{gf'} | \chi_{er'} \rangle \langle \chi_{er'} | Q_a | \chi_{g'f'} \rangle \\
+ \frac{1}{\hbar} \sum_{a,s \neq e} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_g^{(0)} | \hat{r}_\rho | \phi_s^{(0)} \rangle \langle \phi_s^{(0)} | \hat{r}_\sigma | \phi_g^{(0)} \rangle \langle \chi_{gf'} | Q_a | \chi_{er'} \rangle \langle \chi_{er'} | Q_a | \chi_{g'f'} \rangle \\
+ \frac{1}{\hbar} \sum_{a,t \neq g} \frac{(h_a^{(0)})_{es}}{\omega_{es}^0} \langle \phi_e^{(0)} | \hat{r}_\rho | \phi_e^{(0)} \rangle \langle \phi_e^{(0)} | \hat{r}_\sigma | \phi_t^{(0)} \rangle \langle \partial H / \partial Q_a | \phi_t^{(0)} \rangle \langle \chi_{gf'} | Q_a | \chi_{er'} \rangle \langle \chi_{er'} | Q_a | \chi_{g'f'} \rangle \quad (A.26)
\]
\[ \langle \chi g' | \langle \phi_g | \hat{r}_a | \phi_e \rangle | \chi_{er'} \rangle | \langle \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{g'v'} \rangle = \cdots \approx \]

\[ \langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{er'} \rangle | \langle \phi_g | \hat{r}_p | \phi_e \rangle | \chi_{g'v'} \rangle + \]

\[ + \frac{1}{\hbar} \sum_{a, t \neq g} \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{er'} \rangle}{\omega_g^0 - \omega_t^0} \langle \chi_{g'} | \chi_{er'} \rangle | \langle \chi_{g'} | Q_a | \chi_{g'v'} \rangle \]

\[ + \frac{1}{\hbar} \sum_{a, s \neq e} \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{er'} \rangle}{\omega_e^0 - \omega_s^0} \langle \chi_{g'} | Q_a | \chi_{g'v'} \rangle | \langle \chi_{g'} | \chi_{er'} \rangle | \langle \chi_{g'} | Q_a | \chi_{g'v'} \rangle \]

\[ + \frac{1}{\hbar} \sum_{a, t \neq g} \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{er'} \rangle}{\omega_g^0 - \omega_t^0} \langle \chi_{g'} | \chi_{er'} \rangle | \langle \chi_{g'} | Q_a | \chi_{g'v'} \rangle \]

\[ + \frac{1}{\hbar} \sum_{a, t \neq g} \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle | \chi_{er'} \rangle}{\omega_g^0 - \omega_t^0} \langle \chi_{g'} | \chi_{er'} \rangle | \langle \chi_{g'} | Q_a | \chi_{g'v'} \rangle \]

\[ \text{(A.27)} \]

Eqs. A.26 and A.27 can be substituted into eq. 2.10 to derive eq. 2.15:

\[ \langle \alpha_{pa} g_f, g_i \rangle = \frac{1}{\hbar} \sum_{c \neq g} \sum_{v'} \left[ \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 - \omega_t} + \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 + \omega_S} \right] \langle \chi_{g'} | \chi_{er'} \rangle | \chi_{g'v'} \rangle \]

\[ + \frac{1}{\hbar} \sum_{c \neq g} \sum_{v'} \sum_{a} \sum_{s \neq e} \left\{ \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 - \omega_t} \right\} \]

\[ + \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 + \omega_S} \right\} \]

\[ + \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 - \omega_t} \right\} \]

\[ + \frac{\langle \phi_g | \hat{r}_p | \phi_e \rangle \langle \phi_e | \phi_e | \hat{r}_a | \phi_g \rangle}{\omega_e^0 - \omega_g^0 + \omega_S} \right\} \]

\[ = A + B + C \quad \text{(A.28)} \]

where the A, B, and C terms are defined by eqs. 2.16–2.18.
APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER 3.0

B.1 X-RAY CRYSTAL STRUCTURE OF PROPANAMIDE

X-ray crystal diffraction was performed using a Bruker X8 Prospector Ultra equipped with a copper micro-focus tube (λ = 1.54178 Å). A propanamide crystal specimen with approximate dimensions of 0.020 mm × 0.090 mm × 0.110 mm was used for structure determination. The crystal was mounted and placed in a cold N₂ stream (~230 K) for data collection. Crystals were not well-formed, showing diffuse diffraction spots; however, we were able to determine a unit cell and solve a crystal structure.

The frames were integrated with the Bruker SAINT software package. The integration of the data using a monoclinic unit cell yielded a total of 2212 reflections to a maximum θ angle of 68.09° (0.83 Å resolution), of which 689 were independent (average redundancy 3.210, completeness = 83.5%, R_{int} = 8.52%, R_{sig} = 7.03%) and 364 (52.83%) were greater than 2σ(F²).

The final cell constants (Table B1) of a = 8.851(4) Å, b = 5.750(2) Å, c = 9.766(3) Å, β = 114.780(15)°, volume = 451.3(3) Å³, are based upon the refinement of the XYZ-centroids of reflections above 20 σ(I). Unit-cell parameters and analysis of systematic absences indicated propanamide crystallized in a P2₁/c space group.

The structure (Figure B1) was solved via direct methods, which located all of the non-hydrogen atoms. Idealized atom positions were calculated for all hydrogen atoms, except for NH₂ hydrogen atoms (see Table B2–Table B6). The final anisotropic least-squares refinement
on $F^2$ converged at $R1 = 9.53\%$, for the observed data and $wR2 = 24.72\%$ for all data, and the goodness-of-fit was 1.304. The final Fourier map contained no significant residual electron density.

## B.2 DECONVOLUTION OF UVRR SPECTRA

UVRR spectra of propanamide in different acetonitrile and water mixtures were deconvoluted into a sum of a minimum number of Gaussian or Lorentzian bands on a linear background using the GRAMS/AI 8.0 software suite (Thermo Fisher). Figure B2 and Figure B3 show the deconvolution of the 204 nm excitation UVRR spectra of propanamide in pure acetonitrile and pure water.

Table B1: Summary of Crystallographic Data for CH$_3$CH$_2$CONH$_2$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Molecular formula</th>
<th>C$_3$H$_7$NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (K)</td>
<td>230(2)</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.54178</td>
</tr>
<tr>
<td>Theta range (°)</td>
<td>9.18-68.09</td>
</tr>
<tr>
<td>Cell setting</td>
<td>monoclinic</td>
</tr>
<tr>
<td>Space group</td>
<td>P2$_1$/c</td>
</tr>
<tr>
<td>a (Å)</td>
<td>8.851(4)</td>
</tr>
<tr>
<td>b (Å)</td>
<td>5.750(2)</td>
</tr>
<tr>
<td>c (Å)</td>
<td>9.766(3)</td>
</tr>
<tr>
<td>α (°)</td>
<td>90</td>
</tr>
<tr>
<td>β (°)</td>
<td>114.780(15)</td>
</tr>
<tr>
<td>γ (°)</td>
<td>90</td>
</tr>
<tr>
<td>V (Å$^3$)</td>
<td>451.259</td>
</tr>
<tr>
<td>Z</td>
<td>4</td>
</tr>
<tr>
<td>Calc. density (g·cm$^{-1}$)</td>
<td>1.076</td>
</tr>
<tr>
<td>R1</td>
<td>0.0953</td>
</tr>
<tr>
<td>wR2</td>
<td>0.2472</td>
</tr>
</tbody>
</table>
Table B2: Atomic Coordinates and Equivalent Isotropic Displacement Parameters (Å).
Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>x/a</th>
<th>y/b</th>
<th>z/c</th>
<th>U(eq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O9</td>
<td>0.6355(5)</td>
<td>0.7687(5)</td>
<td>0.6186(3)</td>
<td>0.1029(15)</td>
</tr>
<tr>
<td>N10</td>
<td>0.6057(7)</td>
<td>0.8705(7)</td>
<td>0.3890(4)</td>
<td>0.0958(16)</td>
</tr>
<tr>
<td>C1</td>
<td>0.8330(8)</td>
<td>0.3790(10)</td>
<td>0.6341(7)</td>
<td>0.120(2)</td>
</tr>
<tr>
<td>C2</td>
<td>0.7834(9)</td>
<td>0.5471(11)</td>
<td>0.5093(6)</td>
<td>0.121(2)</td>
</tr>
<tr>
<td>C6</td>
<td>0.6678(7)</td>
<td>0.7373(7)</td>
<td>0.5094(5)</td>
<td>0.0867(16)</td>
</tr>
</tbody>
</table>

Table B3: Anisotropic Atomic Displacement Parameters (Å²). Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>U_{11}</th>
<th>U_{22}</th>
<th>U_{33}</th>
<th>U_{13}</th>
<th>U_{12}</th>
</tr>
</thead>
<tbody>
<tr>
<td>O9</td>
<td>0.166(3)</td>
<td>0.095(2)</td>
<td>0.0723(19)</td>
<td>0.0141(14)</td>
<td>0.075(2)</td>
</tr>
<tr>
<td>N10</td>
<td>0.159(4)</td>
<td>0.084(3)</td>
<td>0.067(2)</td>
<td>0.004(2)</td>
<td>0.069(3)</td>
</tr>
<tr>
<td>C1</td>
<td>0.152(6)</td>
<td>0.092(3)</td>
<td>0.116(4)</td>
<td>0.007(3)</td>
<td>0.055(4)</td>
</tr>
<tr>
<td>C2</td>
<td>0.169(6)</td>
<td>0.121(4)</td>
<td>0.098(4)</td>
<td>0.022(3)</td>
<td>0.080(4)</td>
</tr>
<tr>
<td>C6</td>
<td>0.135(4)</td>
<td>0.072(3)</td>
<td>0.070(2)</td>
<td>0.002(2)</td>
<td>0.060(3)</td>
</tr>
</tbody>
</table>

Table B4: Comparison of Bond Lengths (Å) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Bond</th>
<th>Crys. Struc.</th>
<th>DFT Calc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(C6=O9)</td>
<td>1.228(6)</td>
<td>1.223</td>
</tr>
<tr>
<td>r(C6-N10)</td>
<td>1.315(5)</td>
<td>1.353</td>
</tr>
<tr>
<td>r(C6-C2)</td>
<td>1.498(8)</td>
<td>1.519</td>
</tr>
<tr>
<td>r(C2-C1)</td>
<td>1.470(8)</td>
<td>1.521</td>
</tr>
<tr>
<td>r(C1-H3)</td>
<td>0.97</td>
<td>1.090</td>
</tr>
<tr>
<td>r(C1-H4)</td>
<td>0.97</td>
<td>1.091</td>
</tr>
<tr>
<td>r(C1-H5)</td>
<td>0.97</td>
<td>1.091</td>
</tr>
<tr>
<td>r(C2-H7)</td>
<td>0.98</td>
<td>1.095</td>
</tr>
<tr>
<td>r(C2-H8)</td>
<td>0.98</td>
<td>1.093</td>
</tr>
<tr>
<td>r(N10-H11)</td>
<td>0.83(5)</td>
<td>1.007</td>
</tr>
<tr>
<td>r(N10-H12)</td>
<td>0.97(6)</td>
<td>1.008</td>
</tr>
</tbody>
</table>
Table B5: Comparison of Torsion Angles (°) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>Crys. Struc.</th>
<th>DFT Calc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$(N10C6C2C1)</td>
<td>-171.5(5)</td>
<td>169.60</td>
</tr>
<tr>
<td>$\tau$(O9C6C2C1)</td>
<td>9.8(7)</td>
<td>-10.89</td>
</tr>
<tr>
<td>$\tau$(H12N10C6O9)</td>
<td>-16(4)</td>
<td>-2.69</td>
</tr>
<tr>
<td>$\tau$(H12N10C6C2)</td>
<td>166(4)</td>
<td>176.83</td>
</tr>
<tr>
<td>$\tau$(H11N10C6O9)</td>
<td>-176(3)</td>
<td>-178.79</td>
</tr>
<tr>
<td>$\tau$(H11N10C6C2)</td>
<td>5(3)</td>
<td>0.72</td>
</tr>
<tr>
<td>$\tau$(H5C1C2H8)</td>
<td>-56.4</td>
<td>-58.32</td>
</tr>
<tr>
<td>$\tau$(H5C1C2H7)</td>
<td>59.5</td>
<td>59.54</td>
</tr>
<tr>
<td>$\tau$(H5C1C2C6)</td>
<td>-178.4</td>
<td>179.58</td>
</tr>
<tr>
<td>$\tau$(H3C1C2H8)</td>
<td>-176.4</td>
<td>-178.37</td>
</tr>
<tr>
<td>$\tau$(H3C1C2H7)</td>
<td>-60.4</td>
<td>-60.51</td>
</tr>
<tr>
<td>$\tau$(H3C1C2C6)</td>
<td>61.6</td>
<td>59.54</td>
</tr>
<tr>
<td>$\tau$(H4C1C2H8)</td>
<td>63.6</td>
<td>61.64</td>
</tr>
<tr>
<td>$\tau$(H4C1C2H7)</td>
<td>179.6</td>
<td>179.50</td>
</tr>
<tr>
<td>$\tau$(H4C1C2C6)</td>
<td>-58.4</td>
<td>-60.46</td>
</tr>
<tr>
<td>$\tau$(H8C2C6O9)</td>
<td>-112.2</td>
<td>-134.50</td>
</tr>
<tr>
<td>$\tau$(H8C2C6N10)</td>
<td>66.5</td>
<td>-45.99</td>
</tr>
</tbody>
</table>

Table B6: Comparison of Bond Angles (°) Between the DFT-optimized and X-ray Crystal Structure. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>Crys. Struc.</th>
<th>DFT Calc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$(H12N10H11)</td>
<td>121(4)</td>
<td>118.56</td>
</tr>
<tr>
<td>$\theta$(H12N10C6)</td>
<td>118(3)</td>
<td>119.60</td>
</tr>
<tr>
<td>$\theta$(H11N10C6)</td>
<td>118(3)</td>
<td>121.73</td>
</tr>
<tr>
<td>$\theta$(H5C1H3)</td>
<td>109.4</td>
<td>108.43</td>
</tr>
<tr>
<td>$\theta$(H5C1H4)</td>
<td>109.5</td>
<td>108.32</td>
</tr>
<tr>
<td>$\theta$(H5C1C2)</td>
<td>109.5</td>
<td>110.17</td>
</tr>
<tr>
<td>$\theta$(H3C1H4)</td>
<td>109.5</td>
<td>107.97</td>
</tr>
<tr>
<td>$\theta$(H3C1C2)</td>
<td>109.5</td>
<td>110.90</td>
</tr>
<tr>
<td>$\theta$(H4C1C2)</td>
<td>109.5</td>
<td>110.17</td>
</tr>
<tr>
<td>$\theta$(C1C2H8)</td>
<td>108.2</td>
<td>111.15</td>
</tr>
<tr>
<td>$\theta$(C1C2H7)</td>
<td>108.2</td>
<td>110.72</td>
</tr>
<tr>
<td>$\theta$(C1C2C6)</td>
<td>116.5(5)</td>
<td>112.88</td>
</tr>
<tr>
<td>$\theta$(H8C2H7)</td>
<td>107.3</td>
<td>106.26</td>
</tr>
<tr>
<td>$\theta$(H8C2C6)</td>
<td>108.2</td>
<td>108.44</td>
</tr>
<tr>
<td>$\theta$(H7C2C6)</td>
<td>108.2</td>
<td>107.09</td>
</tr>
<tr>
<td>$\theta$(O9C6N10)</td>
<td>121.9(4)</td>
<td>122.07</td>
</tr>
<tr>
<td>$\theta$(O9C6C2)</td>
<td>121.1(4)</td>
<td>122.91</td>
</tr>
<tr>
<td>$\theta$(N10C6C2)</td>
<td>117.0(4)</td>
<td>115.02</td>
</tr>
</tbody>
</table>
Figure B1: Crystal structure of propanamide. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.
Figure B2: Spectral deconvolution of the 204 nm UVRR spectrum of propanamide in (a) water and (b) acetonitrile in the region from 1200–1800 cm$^{-1}$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.
Figure B3: Spectral deconvolution of the 204 nm UVRR spectrum of propanamide in (a) water and (b) acetonitrile in the region from 800–1200 cm$^{-1}$. Adapted with permission from [146]. Copyright © (2015), American Chemical Society.
C.1 DEPENDENCE OF THE AMIII\(^P\) BAND FREQUENCY ON THE OCCC DIHEDRAL ANGLE OF BUTYRAMIDE

In order to understand the asymmetry of the AmIII\(^P\) frequency dependence on the \(\chi_3\) dihedral angle of L-Gln, we investigated the achiral molecule butyramide, which is a model of the L-Gln side chain. DFT calculations were performed on butyramide using the same methods as described in the Computational Details section for L-Gln. Figure C1 shows the cosinuisoidal dependence of the AmIII\(^P\) band frequency on the OCCC dihedral angle of butyramide. Unlike L-Gln, there is no asymmetry in the points about 0\(^\circ\), and the data can be satisfactorily fit to the following equation:

\[
\nu(\chi_3) = 1081 \text{ cm}^{-1} + (16 \text{ cm}^{-1}) \cos(2\chi_3)
\]  

(C.1)

C.2 X-RAY DIFFRACTION OF GLUTAMINE AND DERIVATIVES

We determined the unit cells and lattice constants of the following molecules: L-Gln, D-Gln, NAcGln, Gly-Gln, and Ser-Asn. We solved for the unit cells of each specimen by refining the XYZ-centroids of the reflections above 20\(\sigma(I)\). Information on the unit cells of each molecule are found in Table C1. All of these compounds show crystal lattice constants
that were essentially the same as those found in the Cambridge Crystallographic database [188, 217–219].

Also, we determined the crystal structures of all specimens examined, except for NAcGln. The crystal structures and atomic labeling schemes used in Tables S8-S13 are shown in Figure C2. All structures were solved and refined using the Bruker SHELXTL software package [220]. With the exception of GlnTBE, all the crystal structures that we determined match the Cambridge Crystallographic database entries.

The crystal structure of GlnTBE does not exist in the Cambridge Crystallographic database. The details regarding the data collection and structure refinement for GlnTBE are listed in Table C2 and Table C3. The integration of the data using a monoclinic unit cell yielded a total of 12444 reflections, of which 2246 were independent (average redundancy 5.541, completeness = 99.8%, R_{int} = 5.32%, R_{sig} = 5.00%), to a maximum θ angle of 68.25° (0.83 Å). There were 2181 independent reflections (~97%) that were greater than 2σ(F^2). The intensities were corrected for absorption effects using the multi-scan method, SADABS. The calculated minimum and maximum transmission coefficients (based on the crystal size) for GlnTBE were 0.7800 and 0.8800, respectively. The structure of GlnTBE was solved using the P 1 2_1 1 space group with Z = 2 for the formula unit C_9H_{18}N_2O_3 HCl. The final anisotropic full-matrix least-squares refinement on F^2 with 161 variables converged at R1 = 6.09% for the observed data and wR2 = 15.75% for all data. The goodness-of-fit was 1.812. The largest peak in the final difference electron density synthesis was 1.207 e·Å^{-3} and the largest hole was -0.585 e·Å^{-3}. The calculated density, on the basis of the final model, was 1.239 g·cm^{-3} and F(000) was 256 e\text{\textasciicircum}. Table C2 to Table C7 list the atomic coordinates, equivalent isotropic displacement parameters, anisotropic displacement parameters, bond lengths, and bond angles for the atoms in GlnTBE. The atomic labeling scheme of GlnTBE used for Table C2 to Table C7 is shown in Figure C3.
C.3 RAMAN BAND ASSIGNMENTS OF CRYSTALLINE GLUTAMINE AND DERIVATIVES

C.3.1 Spectral Deconvolution

Figure C4 and Figure C5 show the 633 nm excitation Raman and the 229 nm UVRR spectra, respectively, for each of the crystals examined. We used the GRAMS software suite (Version 8.0, Thermo Fisher Scientific, Inc.) to peak fit the spectra in order to locate bands in the region from $\sim 950 \text{ cm}^{-1}$ to $1200 \text{ cm}^{-1}$. We modeled the visible Raman spectra shown in Figure C4 as a minimum sum of Lorentzian bands.

We modeled the UVRR spectra, shown in Figure C5, as a minimum sum of Voigt bands in order to account for the spectrometer transfer function. We determined the transfer function by illuminating the CCD camera with the 229 nm line from our frequency doubled Ar$^+$ laser. For the slit widths used in our measurements ($\sim 50 \mu\text{m}$), the spectrometer transfer function was well modeled by a Gaussian with a full-width-half-height (FWHH) of $\sim 2.3 \text{ cm}^{-1}$. We fixed the FWHH of the Gaussian component of the Voigt function to $\sim 2.3 \text{ cm}^{-1}$ and allowed the Lorenztian width to vary. We found the average half-width-half height of the AmIII$^P$ vibration to be $\sim 6.6 \pm 2.4 \text{ cm}^{-1}$. Thus, we assume that this value is the homogeneous linewidth, $\Gamma$, of the AmIII$^P$ vibration.

C.3.2 Band Assignments

Table C8–Table C13 show our band assignments of the crystals examined. We used DFT calculations to aid in our band assignments. We also measured the spectra of N-deuterated crystals (data not shown) to verify the band assignments of any vibrations that contain NH$_3$ or NH$_2$ rocking components. The AmIII$^P$ band shifts upon N-deuteration since the vibration contains significant NH$_2$ rocking.

C.3.2.1 L-glutamine and D-glutamine  L-Gln and D-Gln have essentially identical Raman spectra. The experimentally observed and calculated Raman frequencies are shown in Table C8 and Table C9 for L-Gln and D-Gln, respectively. We assign the $\sim 1205 \text{ cm}^{-1}$ band to
the $\text{C}_\beta\text{H}_2$ twisting/$\text{C}_\alpha$–$\text{C}_\beta$ stretching vibration, and the $\sim1166$ cm$^{-1}$ band is assigned to a $\text{C}_\alpha$–H rocking/$\text{C}_\gamma\text{H}_2$ twisting vibration. The $\sim1135$ cm$^{-1}$ and $\sim1105$ cm$^{-1}$ bands exhibit large shifts in their N-deuterated spectra and are therefore assigned to $\text{C}_\beta$–$\text{C}_\gamma$ stretching/$\text{N}_\epsilon\text{H}_2$ rocking and $\text{NH}_3$ rocking vibrations, respectively. The AmIII$^P$ vibration is assigned to the $\sim1097$ cm$^{-1}$ band since this band shows a large shift upon N-deuteration. The remaining peaks are assigned to a $\text{C}_\alpha$–N stretching mode ($\sim1086$ cm$^{-1}$), $\text{NH}_3$ rocking/$\text{C}_\alpha$–N stretching mode ($\sim1052$ cm$^{-1}$), and a $\text{NH}_3$ rocking/$\text{C}_\alpha$–$\text{C}_\beta$ stretching mode ($\sim1000$ cm$^{-1}$).

C.3.2.2 N-Acetyl-L-glutamine The band frequencies and assignments for NAcGln are found in Table C10. The $\sim1180$ cm$^{-1}$ Raman band is assigned to the $\text{C}_\alpha$–N stretching mode, which agrees with that of Kausar et al.’s assignment for a similar band observed in N-Acetyl-L-glutamic acid [221]. The $\sim1138$ cm$^{-1}$ band is assigned to $\text{C}_\alpha$–$\text{C}_\beta$ stretching/$\text{C}_\alpha$–N stretching mode, while the $\sim1111$ cm$^{-1}$ band is assigned to a $\text{N}_\epsilon\text{H}_2$ rocking vibration. The assignment of the 1111 cm$^{-1}$ band to a primary amide $\text{N}_\epsilon\text{H}_2$ rocking mode is consistent with the absence of a peak in this region in N-Acetyl-L-glutamic acid [221]. The $\sim1071$ cm$^{-1}$ band, which appears as a shoulder feature, is assigned to the AmIII$^P$ vibration. The bands at $\sim1061$ cm$^{-1}$ and $\sim1022$ cm$^{-1}$ are assigned to $\text{CH}_3$ and $\text{C}_\beta\text{H}_2$ rocking modes, respectively. The $\sim997$ cm$^{-1}$ band is assigned to a C–C stretching vibration of the acetyl group.

C.3.2.3 L-glutamine t-butyl ester HCl Table C11 displays the frequencies and assignments of crystalline GlnTBE. The $\sim1195$ cm$^{-1}$ band is assigned to a C–O stretching/$\text{CH}_3$ rocking mode of the butyl ester group. Most of the remaining bands in the spectra contain significant $\text{N}_\epsilon\text{H}_2$ or $\text{NH}_3$ character since they shift upon N-deuteration. However, the $\sim1043$ cm$^{-1}$ and $\sim1030$ cm$^{-1}$ bands do not shift appreciably upon N-deuteration and are therefore assigned to $\text{C}_\alpha$–N stretching, and $\text{CH}_3$ rocking vibrations, respectively. The bands located at $\sim1117$ cm$^{-1}$ and $\sim1105$ cm$^{-1}$ are assigned to the $\text{NH}_3$ rocking/$\text{C}_\beta$–$\text{C}_\gamma$ stretching and $\text{NH}_3$ rocking/$\text{N}_\epsilon\text{H}_2$ rocking vibrations, respectively. We assign the $\sim1151$ cm$^{-1}$ band to a $\text{C}_\alpha$–H rocking/$\text{NH}_3$ rocking mode. The band located at $\sim1082$ cm$^{-1}$ is assigned to the AmIII$^P$ vibration due to its sensitivity to N-deuteration. The $\sim998$ cm$^{-1}$ band is assigned to a $\text{NH}_3$ rocking vibration.
C.3.2.4 Glycyl-L-glutamine  Band assignments of crystalline Gly-Gln are shown in Table C12. The band observed at \( \sim 1124 \) cm\(^{-1} \) is assigned to NH\(_3\) rocking mode, while the \( \sim 1093 \) cm\(^{-1} \) band is assigned to the AmIII\(^{P}\) vibration. We do not see the NH\(_3\) rocking mode calculated to be at \( \sim 1105 \) cm\(^{-1} \). The remaining band assignments are shown in Table C12.

C.3.2.5 L-seryl-L-asparagine  Table C13 shows the band assignments of crystalline Ser-Asn. The \( \sim 1188 \) cm\(^{-1} \) band is assigned to a C\(_{\beta}\)H\(_2\) rocking/NH\(_3\) rocking vibration due to its shift upon N-deuteration. The \( \sim 1159 \) cm\(^{-1} \) band is assigned to a C\(_{\alpha}\)–N stretching vibration, while the bands at \( \sim 1121 \) cm\(^{-1} \) and \( \sim 1018 \) cm\(^{-1} \) are both assigned to C\(_{\beta}\)–O\(_{\gamma}\)H stretching modes of the serine side chain. We assign the AmIII\(^{P}\) vibration to the \( \sim 1051 \) cm\(^{-1} \) band. The remaining bands located at \( \sim 1083 \) cm\(^{-1} \), \( \sim 1065 \) cm\(^{-1} \), \( \sim 1019 \) cm\(^{-1} \), and \( \sim 1004 \) cm\(^{-1} \) are assigned to N\(_{\delta}\)H\(_2\) rocking, NH\(_3\) rocking, C\(_{\beta}\)H\(_2\) rocking, and C\(_{\beta}\)H\(_2\) rocking/NH\(_3\) rocking modes, respectively.

C.4 DEPENDENCE OF THE AMIII\(^{P}\) FREQUENCY ON HYDROGEN BONDING AND DIELECTRIC ENVIRONMENT

Figure C6 shows the \( \sim 204 \) nm excitation UVRR spectra of butyramide (Sigma Aldrich, \( \geq 98\% \) purity) in water and acetonitrile (Acros Organics, HPLC, far-UV grade) in the region of 950 cm\(^{-1} \) to 1250 cm\(^{-1} \). In the water spectrum (Figure C6a), there are prominent bands at \( \sim 1132 \) cm\(^{-1} \) and \( \sim 1076 \) cm\(^{-1} \), as well as a shoulder feature at \( \sim 1050 \) cm\(^{-1} \). Based on our DFT calculations (see Figure C6), we assign the band at 1132 cm\(^{-1} \) to a C–C stretching mode that contains significant CH\(_3\) and NH\(_2\) rocking character. The \( \sim 1076 \) cm\(^{-1} \) band is assigned to the AmIII\(^{P}\) vibration. The shoulder feature at \( \sim 1050 \) cm\(^{-1} \) is assigned to a C–CH\(_3\) stretching vibration.

As shown in Figure C6b, all the bands in the 950 cm\(^{-1} \) to 1250 cm\(^{-1} \) region downshift in acetonitrile. The AmIII\(^{P}\) band downshifts 12 cm\(^{-1} \) to \( \sim 1064 \) cm\(^{-1} \). This is similar to that in propanamide [146], where the AmIII\(^{P}\) band downshifts \( \sim 18 \) cm\(^{-1} \) in acetonitrile compared to water. Therefore we conclude that on average the AmIII\(^{P}\) band is downshifted \( \sim 15 \) cm\(^{-1} \).
in acetonitrile compared to water.

C.5 CIRCULAR DICHROISM (CD) OF THE GLN$_3$ PEPTIDE

The temperature dependent CD spectra of Glu$_3$ are shown in Figure C7. The spectra were measured at a concentration of 0.5 mg·mL$^{-1}$ using a Jasco-715 spectropolarimeter with a 0.1 cm pathlength cuvette. We averaged six individual CD spectra for each temperature.
Table C1: Summary of Crystallographic Data of Gln and Derivatives. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>L-glutamine</th>
<th>D-glutamine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Formula</td>
<td>C$<em>5$H$</em>{10}$N$_2$O$_3$</td>
<td>C$<em>5$H$</em>{10}$N$_2$O$_3$</td>
</tr>
<tr>
<td>Temperature</td>
<td>230(2) K</td>
<td>230(2) K</td>
</tr>
<tr>
<td>Crystal System</td>
<td>orthorhombic</td>
<td>orthorhombic</td>
</tr>
<tr>
<td>Space Group</td>
<td>P 2$_1$ 2$_1$ 2$_1$</td>
<td>P 2$_1$ 2$_1$ 2$_1$</td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>a = 5.1054(4) Å, α = 90°</td>
<td>a = 5.1030(4) Å, α = 90°</td>
</tr>
<tr>
<td></td>
<td>b = 7.7641(5) Å, β = 90°</td>
<td>b = 7.7634(7) Å, β = 90°</td>
</tr>
<tr>
<td></td>
<td>c = 15.9993(11) Å, γ = 90°</td>
<td>c = 16.0056(13) Å, γ = 90°</td>
</tr>
<tr>
<td>Volume</td>
<td>634.19(8) Å$^3$</td>
<td>634.09(9) Å$^3$</td>
</tr>
<tr>
<td>Z</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N-Acetyl-L-glutamine</th>
<th>L-glycyl-L-glutamine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Formula</td>
<td>C$<em>7$H$</em>{13}$N$_2$O$_4$</td>
<td>C$<em>7$H$</em>{15}$N$_3$O$_5$</td>
</tr>
<tr>
<td>Temperature</td>
<td>230(2) K</td>
<td>230(2) K</td>
</tr>
<tr>
<td>Crystal System</td>
<td>orthorhombic</td>
<td>orthorhombic</td>
</tr>
<tr>
<td>Space Group</td>
<td>P 2$_1$ 2$_1$ 2$_1$</td>
<td>P 2$_1$ 2$_1$ 2$_1$</td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>a = 5.0732(47) Å, α = 90°</td>
<td>a = 5.4025(2) Å, α = 90°</td>
</tr>
<tr>
<td></td>
<td>b = 12.9033(95) Å, β = 90°</td>
<td>b = 11.5771(4) Å, β = 90°</td>
</tr>
<tr>
<td></td>
<td>c = 13.7830(104) Å, γ = 90°</td>
<td>c = 15.4651(5) Å, γ = 90°</td>
</tr>
<tr>
<td>Volume</td>
<td>902.274(2) Å$^3$</td>
<td>967.27(6) Å$^3$</td>
</tr>
<tr>
<td>Z</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>L-seryl-L-asparagine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Formula</td>
<td>C$<em>7$H$</em>{15}$N$_3$O$_6$</td>
</tr>
<tr>
<td>Temperature</td>
<td>230(2) K</td>
</tr>
<tr>
<td>Crystal System</td>
<td>triclinic</td>
</tr>
<tr>
<td>Space Group</td>
<td>P 1</td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>a = 4.7643(8) Å, α = 63.988(5)°</td>
</tr>
<tr>
<td></td>
<td>b = 7.5452(12) Å, β = 76.646(5)°</td>
</tr>
<tr>
<td></td>
<td>c = 8.5942(14) Å, γ = 75.190(7)°</td>
</tr>
<tr>
<td>Volume</td>
<td>265.90(8) Å$^3$</td>
</tr>
<tr>
<td>Z</td>
<td>1</td>
</tr>
</tbody>
</table>
Table C2: Summary of Crystallographic Data for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value/Detailed Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Formula</td>
<td>$C_9H_{18}N_2O_3\cdot HCl$</td>
</tr>
<tr>
<td>Temperature</td>
<td>230(2) K</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.54178 Å</td>
</tr>
<tr>
<td>Crystal Size</td>
<td>0.04 mm $\times$ 0.140 mm $\times$ 0.160 mm</td>
</tr>
<tr>
<td>Crystal Habit</td>
<td>clear, colorless rectangular prism</td>
</tr>
<tr>
<td>Crystal System</td>
<td>monoclinic</td>
</tr>
<tr>
<td>Space Group</td>
<td>$P 1 2_1 1$</td>
</tr>
<tr>
<td>Unit Cell Dimensions</td>
<td>$a = 10.4579(4)$ Å, $\alpha = 90'$</td>
</tr>
<tr>
<td></td>
<td>$b = 5.2517(2)$ Å, $\beta = 90.1490(10)'$</td>
</tr>
<tr>
<td></td>
<td>$c = 11.6493(4)$ Å, $\gamma = 90'$</td>
</tr>
<tr>
<td>Volume</td>
<td>639.80(4) Å³</td>
</tr>
<tr>
<td>$Z$</td>
<td>2</td>
</tr>
<tr>
<td>Density (calc.)</td>
<td>1.239 g$\cdot$cm$^{-3}$</td>
</tr>
<tr>
<td>Absorption Coefficient</td>
<td>2.601 mm$^{-1}$</td>
</tr>
<tr>
<td>$F(000)$</td>
<td>256</td>
</tr>
</tbody>
</table>
Table C3: Data Collection and Structure Refinement for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theta range (°)</td>
<td>3.79'-68.25'</td>
</tr>
<tr>
<td>Index ranges</td>
<td>-12≤h≤12, -6≤k≤6, -14≤l≤14</td>
</tr>
<tr>
<td>Reflections collected</td>
<td>12444</td>
</tr>
<tr>
<td>Independent reflections</td>
<td>2246 [R_{int} = 0.0532]</td>
</tr>
<tr>
<td>Refinement method</td>
<td>Full-matrix least-squares on F²</td>
</tr>
<tr>
<td>Refinement program</td>
<td>SHELXL-2014/7 [220]</td>
</tr>
<tr>
<td>Function minimized</td>
<td>( \sum w(F_o^2 - F_c^2)^2 )</td>
</tr>
<tr>
<td>Data / restraints / parameters</td>
<td>2246 / 1 / 161</td>
</tr>
<tr>
<td>Goodness-of-fit on F²</td>
<td>1.812</td>
</tr>
<tr>
<td>( \Delta/\sigma_{max} )</td>
<td>0.001</td>
</tr>
<tr>
<td>Final R indices</td>
<td>2181 data; I&gt;2σ(I), R1 = 0.0609, wR2 = 0.1566</td>
</tr>
<tr>
<td></td>
<td>all data, R1 = 0.0618, wR2 = 0.1575</td>
</tr>
<tr>
<td>Weighting scheme</td>
<td>( w = 1/[(\sigma^2(F_o^2) + (0.0680P)^2)] )</td>
</tr>
<tr>
<td></td>
<td>where, ( P = (F_o^2 + 2F_c^2)/3 )</td>
</tr>
<tr>
<td>Absolute structure parameter</td>
<td>0.2(0)</td>
</tr>
<tr>
<td>Extinction coefficient</td>
<td>0.0310(50)</td>
</tr>
<tr>
<td>Largest diff. peak and hole</td>
<td>1.207 and -0.585 e·Å⁻³</td>
</tr>
<tr>
<td>R.M.S. deviation from mean</td>
<td>0.127 e·Å⁻³</td>
</tr>
</tbody>
</table>
Table C4: Atomic Coordinates and Equivalent Isotropic Atomic Displacement Parameters (Å²) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>x/a</th>
<th>y/b</th>
<th>z/c</th>
<th>U(eq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl1</td>
<td>0.92521(7)</td>
<td>0.46852(17)</td>
<td>0.35473(6)</td>
<td>0.0280(3)</td>
</tr>
<tr>
<td>O1</td>
<td>0.8535(4)</td>
<td>0.1912(7)</td>
<td>0.0330(3)</td>
<td>0.0450(10)</td>
</tr>
<tr>
<td>N1</td>
<td>0.8493(4)</td>
<td>0.6054(10)</td>
<td>0.0824(3)</td>
<td>0.0373(9)</td>
</tr>
<tr>
<td>C1</td>
<td>0.8498(4)</td>
<td>0.4164(10)</td>
<td>0.0059(3)</td>
<td>0.0299(10)</td>
</tr>
<tr>
<td>O2</td>
<td>0.7073(3)</td>
<td>0.7877(7)</td>
<td>0.6433(3)</td>
<td>0.0400(8)</td>
</tr>
<tr>
<td>N2</td>
<td>0.9045(3)</td>
<td>0.4597(9)</td>
<td>0.6235(2)</td>
<td>0.0241(7)</td>
</tr>
<tr>
<td>C2</td>
<td>0.8456(4)</td>
<td>0.4992(10)</td>
<td>0.8807(3)</td>
<td>0.0311(9)</td>
</tr>
<tr>
<td>O3</td>
<td>0.5703(3)</td>
<td>0.4777(7)</td>
<td>0.6963(3)</td>
<td>0.0343(7)</td>
</tr>
<tr>
<td>C3</td>
<td>0.8066(4)</td>
<td>0.2834(8)</td>
<td>0.8016(3)</td>
<td>0.0239(8)</td>
</tr>
<tr>
<td>C4</td>
<td>0.7839(4)</td>
<td>0.3639(8)</td>
<td>0.6766(3)</td>
<td>0.0215(8)</td>
</tr>
<tr>
<td>C5</td>
<td>0.6835(4)</td>
<td>0.5721(8)</td>
<td>0.6684(3)</td>
<td>0.0227(8)</td>
</tr>
<tr>
<td>C6</td>
<td>0.4556(4)</td>
<td>0.6448(10)</td>
<td>0.7142(5)</td>
<td>0.0379(10)</td>
</tr>
<tr>
<td>C7</td>
<td>0.4806(6)</td>
<td>0.8073(17)</td>
<td>0.8189(6)</td>
<td>0.0632(18)</td>
</tr>
<tr>
<td>C8</td>
<td>0.3495(6)</td>
<td>0.4559(16)</td>
<td>0.7351(12)</td>
<td>0.107(4)</td>
</tr>
<tr>
<td>C9</td>
<td>0.4293(7)</td>
<td>0.8023(19)</td>
<td>0.6099(6)</td>
<td>0.067(2)</td>
</tr>
</tbody>
</table>

*a*U(eq) is defined as one third of the trace of the orthogonalized U_{ij} tensor.
Table C5: Anisotropic Atomic Displacement Parameters\(^a\) (\(\text{Å}^2\)) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>(U_{11})</th>
<th>(U_{22})</th>
<th>(U_{33})</th>
<th>(U_{23})</th>
<th>(U_{13})</th>
<th>(U_{12})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl1</td>
<td>0.0312(5)</td>
<td>0.0323(5)</td>
<td>0.0204(5)</td>
<td>0.0026(4)</td>
<td>-0.0058(3)</td>
<td>-0.0052(4)</td>
</tr>
<tr>
<td>O1</td>
<td>0.070(3)</td>
<td>0.038(2)</td>
<td>0.0274(15)</td>
<td>0.0104(15)</td>
<td>-0.0012(15)</td>
<td>0.0098(17)</td>
</tr>
<tr>
<td>N1</td>
<td>0.051(2)</td>
<td>0.041(3)</td>
<td>0.0205(18)</td>
<td>-0.0004(18)</td>
<td>-0.0050(15)</td>
<td>-0.001(2)</td>
</tr>
<tr>
<td>C1</td>
<td>0.0295(18)</td>
<td>0.041(3)</td>
<td>0.0196(18)</td>
<td>0.0031(16)</td>
<td>-0.0031(14)</td>
<td>0.0021(17)</td>
</tr>
<tr>
<td>O2</td>
<td>0.0344(16)</td>
<td>0.0229(15)</td>
<td>0.063(2)</td>
<td>0.0082(15)</td>
<td>0.0047(15)</td>
<td>0.0011(14)</td>
</tr>
<tr>
<td>N2</td>
<td>0.0280(15)</td>
<td>0.0262(16)</td>
<td>0.0181(14)</td>
<td>0.0022(15)</td>
<td>-0.0003(11)</td>
<td>0.0010(16)</td>
</tr>
<tr>
<td>C2</td>
<td>0.044(2)</td>
<td>0.031(2)</td>
<td>0.0182(16)</td>
<td>0.0027(17)</td>
<td>-0.0071(15)</td>
<td>-0.003(2)</td>
</tr>
<tr>
<td>O3</td>
<td>0.0249(13)</td>
<td>0.0205(12)</td>
<td>0.0576(18)</td>
<td>0.0017(17)</td>
<td>0.0039(11)</td>
<td>0.0021(14)</td>
</tr>
<tr>
<td>C3</td>
<td>0.0297(18)</td>
<td>0.0231(17)</td>
<td>0.0189(16)</td>
<td>0.0056(15)</td>
<td>-0.0012(14)</td>
<td>0.0006(15)</td>
</tr>
<tr>
<td>C4</td>
<td>0.0239(19)</td>
<td>0.0230(18)</td>
<td>0.0175(16)</td>
<td>-0.0002(14)</td>
<td>-0.0045(14)</td>
<td>-0.0017(14)</td>
</tr>
<tr>
<td>C5</td>
<td>0.0264(18)</td>
<td>0.0189(17)</td>
<td>0.0227(16)</td>
<td>-0.0006(14)</td>
<td>-0.0051(14)</td>
<td>-0.0013(14)</td>
</tr>
<tr>
<td>C6</td>
<td>0.0252(19)</td>
<td>0.027(2)</td>
<td>0.062(3)</td>
<td>0.002(2)</td>
<td>0.0069(18)</td>
<td>0.0020(18)</td>
</tr>
<tr>
<td>C7</td>
<td>0.059(3)</td>
<td>0.072(4)</td>
<td>0.059(3)</td>
<td>-0.015(3)</td>
<td>0.007(3)</td>
<td>0.022(3)</td>
</tr>
<tr>
<td>C8</td>
<td>0.036(3)</td>
<td>0.035(3)</td>
<td>0.251(13)</td>
<td>0.010(6)</td>
<td>0.048(5)</td>
<td>0.000(3)</td>
</tr>
<tr>
<td>C9</td>
<td>0.050(3)</td>
<td>0.090(5)</td>
<td>0.060(4)</td>
<td>0.004(4)</td>
<td>-0.009(3)</td>
<td>0.039(4)</td>
</tr>
</tbody>
</table>

\(^a\)The anisotropic atomic displacement factor exponent takes the form:

\[-2\pi^2[h^2 \cdot a^2 \cdot U_{11} + \cdots + 2h \cdot k \cdot a \cdot b \cdot U_{12}]\]
Table C6: Hydrogen Atomic Coordinates and Isotropic Atomic Displacement Parameters (Å²) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th>x/a</th>
<th>y/b</th>
<th>z/c</th>
<th>U(eq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1NB</td>
<td>0.858(5)</td>
<td>0.744(15)</td>
<td>1.062(5)</td>
<td>0.029(16)</td>
</tr>
<tr>
<td>H1NA</td>
<td>0.870(7)</td>
<td>0.582(15)</td>
<td>1.156(7)</td>
<td>0.053(18)</td>
</tr>
<tr>
<td>H2NC</td>
<td>0.889(4)</td>
<td>0.477(13)</td>
<td>0.546(4)</td>
<td>0.024(10)</td>
</tr>
<tr>
<td>H2NB</td>
<td>0.926(6)</td>
<td>0.597(14)</td>
<td>0.656(5)</td>
<td>0.032(14)</td>
</tr>
<tr>
<td>H2NA</td>
<td>0.980(6)</td>
<td>0.355(12)</td>
<td>0.645(5)</td>
<td>0.027(13)</td>
</tr>
<tr>
<td>H2A</td>
<td>0.7846</td>
<td>0.6398</td>
<td>0.8723</td>
<td>0.037</td>
</tr>
<tr>
<td>H2B</td>
<td>0.9301</td>
<td>0.5618</td>
<td>0.8581</td>
<td>0.037</td>
</tr>
<tr>
<td>H3A</td>
<td>0.7281</td>
<td>0.2061</td>
<td>0.8312</td>
<td>0.029</td>
</tr>
<tr>
<td>H3B</td>
<td>0.8736</td>
<td>0.1529</td>
<td>0.8033</td>
<td>0.029</td>
</tr>
<tr>
<td>H7A</td>
<td>0.4972</td>
<td>0.6984</td>
<td>0.8845</td>
<td>0.095</td>
</tr>
<tr>
<td>H7B</td>
<td>0.5542</td>
<td>0.9153</td>
<td>0.8053</td>
<td>0.095</td>
</tr>
<tr>
<td>H7C</td>
<td>0.4063</td>
<td>0.9126</td>
<td>0.8342</td>
<td>0.095</td>
</tr>
<tr>
<td>H8A</td>
<td>0.369</td>
<td>0.3564</td>
<td>0.8031</td>
<td>0.161</td>
</tr>
<tr>
<td>H8B</td>
<td>0.2696</td>
<td>0.5463</td>
<td>0.7463</td>
<td>0.161</td>
</tr>
<tr>
<td>H8C</td>
<td>0.3418</td>
<td>0.3434</td>
<td>0.6694</td>
<td>0.161</td>
</tr>
<tr>
<td>H9A</td>
<td>0.355</td>
<td>0.9086</td>
<td>0.6234</td>
<td>0.1</td>
</tr>
<tr>
<td>H9B</td>
<td>0.5027</td>
<td>0.9092</td>
<td>0.5938</td>
<td>0.1</td>
</tr>
<tr>
<td>H9C</td>
<td>0.4131</td>
<td>0.6913</td>
<td>0.545</td>
<td>0.1</td>
</tr>
<tr>
<td>H4</td>
<td>0.749(5)</td>
<td>0.212(10)</td>
<td>0.627(4)</td>
<td>0.020(12)</td>
</tr>
</tbody>
</table>
Table C7: Bond Lengths (Å) and Bond Angles (°) for L-glutamine t-butyl ester HCl. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Bond Lengths</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O1-C1</td>
<td>1.225(6)</td>
</tr>
<tr>
<td>N1-H1NB</td>
<td>0.77(8)</td>
</tr>
<tr>
<td>C1-C2</td>
<td>1.522(5)</td>
</tr>
<tr>
<td>N2-C4</td>
<td>1.494(5)</td>
</tr>
<tr>
<td>N2-H2NB</td>
<td>0.85(7)</td>
</tr>
<tr>
<td>C2-C3</td>
<td>1.515(6)</td>
</tr>
<tr>
<td>C2-H2B</td>
<td>0.98</td>
</tr>
<tr>
<td>O3-C6</td>
<td>1.501(5)</td>
</tr>
<tr>
<td>C3-H3A</td>
<td>0.98</td>
</tr>
<tr>
<td>C4-C5</td>
<td>1.519(5)</td>
</tr>
<tr>
<td>C6-C9</td>
<td>1.494(9)</td>
</tr>
<tr>
<td>C6-C7</td>
<td>1.511(9)</td>
</tr>
<tr>
<td>C7-H7B</td>
<td>0.97</td>
</tr>
<tr>
<td>C8-H8A</td>
<td>0.97</td>
</tr>
<tr>
<td>C8-H8C</td>
<td>0.97</td>
</tr>
<tr>
<td>C9-H9B</td>
<td>0.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond Angles</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1-N1-H1NB</td>
<td>119.(5)</td>
</tr>
<tr>
<td>H1NB-N1-H1NA</td>
<td>114.(7)</td>
</tr>
<tr>
<td>O1-C1-C2</td>
<td>121.6(4)</td>
</tr>
<tr>
<td>C4-N2-H2NC</td>
<td>107.3(3)</td>
</tr>
<tr>
<td>H2NC-N2-H2NB</td>
<td>114.(6)</td>
</tr>
<tr>
<td>H2NC-N2-H2NA</td>
<td>115.(5)</td>
</tr>
<tr>
<td>C3-C2-C1</td>
<td>112.1(4)</td>
</tr>
<tr>
<td>C1-C2-H2A</td>
<td>109.2</td>
</tr>
<tr>
<td>C5-C3-C6</td>
<td>122.0(4)</td>
</tr>
<tr>
<td>C2-C3-H3A</td>
<td>108.7</td>
</tr>
<tr>
<td>C2-C3-H3B</td>
<td>108.7</td>
</tr>
<tr>
<td>H3A-C3-H3B</td>
<td>107.6</td>
</tr>
<tr>
<td>N2-C4-C3</td>
<td>110.9(3)</td>
</tr>
<tr>
<td>N2-C4-H4</td>
<td>109.3</td>
</tr>
<tr>
<td>C3-C4-H4</td>
<td>111.(3)</td>
</tr>
<tr>
<td>O2-C5-C4</td>
<td>123.6(4)</td>
</tr>
<tr>
<td>C9-C6-O3</td>
<td>110.8(4)</td>
</tr>
<tr>
<td>O3-C6-C8</td>
<td>103.1(4)</td>
</tr>
<tr>
<td>O3-C6-C7</td>
<td>107.8(4)</td>
</tr>
<tr>
<td>C6-C7-H7A</td>
<td>109.5</td>
</tr>
<tr>
<td>H7A-C7-H7B</td>
<td>109.5</td>
</tr>
<tr>
<td>H7A-C7-H7C</td>
<td>109.5</td>
</tr>
<tr>
<td>C6-C8-H8A</td>
<td>109.5</td>
</tr>
<tr>
<td>H8A-C8-H8B</td>
<td>109.5</td>
</tr>
<tr>
<td>H8A-C8-H8C</td>
<td>109.5</td>
</tr>
<tr>
<td>C6-C9-H9A</td>
<td>109.5</td>
</tr>
</tbody>
</table>
Table C8: Raman Frequencies (cm⁻¹) and Assignments of Crystalline L-glutaminea. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PEDb (≥5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1205</td>
<td>1215</td>
<td>-τCβH₂ (21), -νCαCβ (18), -τC₂H₂ (16), -ρ'NH₃ (13), δNCαC(OO) (5)</td>
<td>τCH₂/νC–C</td>
</tr>
<tr>
<td>1166</td>
<td>1153</td>
<td>-ρC₂H₂ (20), τC₂H₂ (17), -ρ'NH₃ (13), τCβH₂ (12), -νCαCβ (8), τC₂H₂ (6)</td>
<td>ρCH/τCH₂</td>
</tr>
<tr>
<td>1135</td>
<td>1122</td>
<td>νCβCγ (34), ρN₂H₂ (17), -νCαCβ (7), νCαN (6), -βN₂C₂Cγ (5)</td>
<td>νC–C/ρNH₂</td>
</tr>
<tr>
<td>1105</td>
<td>1109</td>
<td>ρNH₂ (27), -τC₂H₂ (10), -ρN₂H₂ (10), -δNCαC(OO) (9), -ρC₂H₂ (7), νCαN (7)</td>
<td>ρNH₂</td>
</tr>
<tr>
<td>1097</td>
<td>1097</td>
<td>νCβCγ (26), -ρN₂H₂ (26), -τC₂N (13), -ρNH₃ (8)</td>
<td>AmIIP</td>
</tr>
<tr>
<td>1086</td>
<td>1038</td>
<td>νCαN (36), -νCαCβ (9), ρC₂H₂ (8), ρ'NH₃ (6), ρC₂H₂ (5), ρ'NH₂ (5)</td>
<td>νC–N</td>
</tr>
<tr>
<td>1052</td>
<td>1003</td>
<td>ρNH₂ (25), -τC₂N (19), ρC₂H₂ (14), ρC₂H₂ (14), νCαCβ (7), -τC₂H₂ (5)</td>
<td>ρNH₂/νC–N</td>
</tr>
<tr>
<td>1000</td>
<td>974</td>
<td>-ρ'NH₃ (38), νCαCβ (25), -τC₂C (8), -σC₂Cβγ (7), νCαN (6)</td>
<td>ρNH₂/νC–C</td>
</tr>
</tbody>
</table>

aFrequencies correspond to visible Raman (633 nm excitation) spectrum. bν: stretch; δ: symmetric deformation; δas: asymmetric deformation; ω: wagging; β: in-plane bending; τ: twisting.

Table C9: Raman Frequencies (cm⁻¹) and Assignments of Crystalline D-glutaminea. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PEDb (≥5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1204</td>
<td>1215</td>
<td>-τCβH₂ (21), -νCαCβ (18), -τC₂H₂ (16), -ρ'NH₃ (13), δNCαC(OO) (5)</td>
<td>τCH₂/νC–C</td>
</tr>
<tr>
<td>1165</td>
<td>1153</td>
<td>-ρC₂H₂ (20), τC₂H₂ (17), -ρ'NH₃ (13), τCβH₂ (12), -νCαCβ (8), τC₂H₂ (6)</td>
<td>ρCH/τCH₂</td>
</tr>
<tr>
<td>1134</td>
<td>1122</td>
<td>νCβCγ (34), ρN₂H₂ (17), -νCαCβ (7), νCαN (6), -βN₂C₂Cγ (5)</td>
<td>νC–C/ρNH₂</td>
</tr>
<tr>
<td>1105</td>
<td>1109</td>
<td>ρNH₂ (27), -τC₂H₂ (10), -ρN₂H₂ (10), -δNCαC(OO) (9), -ρC₂H₂ (7), νCαN (7)</td>
<td>ρNH₂</td>
</tr>
<tr>
<td>1096</td>
<td>1097</td>
<td>νCβCγ (26), -ρN₂H₂ (26), -τC₂N (13), -ρNH₃ (8)</td>
<td>AmIIP</td>
</tr>
<tr>
<td>1086</td>
<td>1038</td>
<td>νCαN (36), -νCαCβ (9), ρC₂H₂ (8), ρ'NH₃ (6), ρC₂H₂ (5), ρ'NH₂ (5)</td>
<td>νC–N</td>
</tr>
<tr>
<td>1051</td>
<td>1003</td>
<td>ρNH₂ (25), -τC₂N (19), ρC₂H₂ (14), ρC₂H₂ (14), νCαCβ (7), -τC₂H₂ (5)</td>
<td>ρNH₂/νC–N</td>
</tr>
<tr>
<td>999</td>
<td>974</td>
<td>-ρ'NH₃ (38), νCαCβ (25), -τC₂C (8), -σC₂Cβγ (7), νCαN (6)</td>
<td>ρNH₂/νC–C</td>
</tr>
</tbody>
</table>

aFrequencies correspond to visible Raman (633 nm excitation) spectrum. bν: stretch; δ: symmetric deformation; δas: asymmetric deformation; ω: wagging; β: in-plane bending; τ: twisting.

Table C10: Raman Frequencies (cm⁻¹) and Assignments of Crystalline N-Acetyl-L-glutaminea. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PEDb (≥5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1192</td>
<td>1192</td>
<td>-τCβH₂ (32), -τC₂H₂ (24), -νCαCβ (12), ρC₂H₂ (11)</td>
<td>τCH₂</td>
</tr>
<tr>
<td>1180</td>
<td>1166</td>
<td>νCαN (46), -δNCαCβCOO (8), ωC₂H₂ (7), -σC₂Cβγ (6), νCβCγ (5)</td>
<td>νC–N</td>
</tr>
<tr>
<td>1138</td>
<td>1119</td>
<td>νCαCβ (22), -νCαN (12), νCβCγ (10), ρN₂H₂ (8), -τC₂H₂ (5)</td>
<td>νC–C/νC–N</td>
</tr>
<tr>
<td>1111</td>
<td>1110</td>
<td>ρN₂H₂ (20), -τC₂Cβγ (15), νCβCγ (14), ρCβH₂ (7), ρC₂H₂ (6), ρC₂N (5), -ωC₂H₂ (5)</td>
<td>ρNH₂/νC–C</td>
</tr>
<tr>
<td>1097</td>
<td>1093</td>
<td>-τCβCγ (42), ρN₂H₂ (24), νC₂N (13)</td>
<td>AmIIP</td>
</tr>
<tr>
<td>1061</td>
<td>1058</td>
<td>ρC₂H₃ (63), -δNCO/CH₂ (19), -δas/CH₃ (9), -ρ'CH₃ (7)</td>
<td>ρCH₃</td>
</tr>
<tr>
<td>1022</td>
<td>1043</td>
<td>ρC₂H₃ (25), ρC₂H₃ (15), ρC₂H₂ (14), -τC₂C (8), νCαCβ (8), δCCO/ν (5)</td>
<td>ρCH₃/νCH₃</td>
</tr>
<tr>
<td>1013</td>
<td>1013</td>
<td>ρCH₃ (58), -τC₂CH₂ (12), ρCH₃ (5), -δas/CH₃ (5)</td>
<td>ρCH₃</td>
</tr>
<tr>
<td>997</td>
<td>971</td>
<td>ωC₂CH₃ (28), νC₂N (16), νCO/CH₂ (11), -νCαCβ (10), -δC₂N (8)</td>
<td>νC–CH₃/νC–N</td>
</tr>
</tbody>
</table>

aFrequencies correspond to visible Raman (633 nm excitation) spectrum. bν: stretch; δ: symmetric deformation; δas: asymmetric deformation; ω: wagging; β: in-plane bending; τ: twisting.
Table C11: Raman Frequencies (cm$^{-1}$) and Assignments of Crystalline L-glutamine t-butyl ester HCl$^a$. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PED$^b$ ($\geq$5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1195</td>
<td>1195</td>
<td>νCO (17), ρCH$_3$ (10), −ρCH$_3$ (10), −ρC$_2$H$_5$ (9), −νCO (6), −δC$^\prime$O'CCC (6), −δC$^\prime$O'CCC (6), ρCH$_3$ (5), ρCH$_3$ (5)</td>
<td>νCO/ρCH$_3$</td>
</tr>
<tr>
<td>1151</td>
<td>1161</td>
<td>ρC$<em>\alpha$H (23), ρNH$<em>3$ (19), τC$\alpha$N$<em>2$ (11), −νC$</em>\alpha$C$</em>\beta$ (10), ρC$</em>\beta$H$_2$ (8)</td>
<td>ρCH/ρNH$_3$</td>
</tr>
<tr>
<td>1117</td>
<td>1126</td>
<td>−ρNH$<em>3$ (19), −νC$</em>\beta$C$<em>\gamma$ (16), −νC$</em>\alpha$N (14), δNC$<em>\alpha$C$</em>\beta$C (9), −ρN$_2$H$_2$ (5)</td>
<td>ρNH$_3$/νC-C</td>
</tr>
<tr>
<td>1105</td>
<td>1119</td>
<td>−ρNH$<em>3$ (15), ρN$<em>2$H$<em>2$ (14), ρC$</em>\beta$H$<em>2$ (10), νC$</em>\beta$C$</em>\gamma$ (9), ρC$</em>\alpha$H$<em>2$ (7), ρC$</em>\alpha$H (6), δNC$<em>\alpha$C$</em>\beta$C (6)</td>
<td>ρNH$_3$/ρNH$_2$</td>
</tr>
<tr>
<td>1082</td>
<td>1100</td>
<td>ρN$<em>2$H$<em>2$ (34), −νC$</em>\beta$C$</em>\gamma$ (24), νC$_2$N$_2$ (14), νC$_3$O$_2$ (5)</td>
<td>AmH$^p$</td>
</tr>
<tr>
<td>1043</td>
<td>1068</td>
<td>−νC$_\alpha$N (17), ρCH$<em>3$ (13), −νC$</em>\gamma$H$<em>3$ (11), −ρCH$<em>3$ (11), νC$</em>\beta$C$</em>\gamma$ (10), ρCH$<em>3$ (5), νC$</em>\alpha$C (5)</td>
<td>νC-N/ρCH$_3$</td>
</tr>
<tr>
<td>1030</td>
<td>1058</td>
<td>ρCH$<em>3$ (27), −νC$</em>\gamma$H$<em>3$ (26), −νC$</em>\gamma$H$_3$ (10), ρCH$_3$ (10), −νCCH$_3$ (6)</td>
<td>ρCH$_3$</td>
</tr>
<tr>
<td>1057</td>
<td>1057</td>
<td>−νC$_\alpha$N (24), νC$<em>3$C$</em>\gamma$ (11), ρCH$_3$ (9), ρCH$_3$ (8), −ρCH$_3$ (7), −ρC$_3$H$_2$ (3)</td>
<td>νC-N/νC-C</td>
</tr>
<tr>
<td>1014</td>
<td>1014</td>
<td>−ρNH$<em>3$ (24), −ρC$</em>\beta$H$<em>2$ (18), −ρC$</em>\gamma$H$<em>2$ (16), ρC$</em>\alpha$N (14), −νC$<em>\alpha$C$</em>\beta$ (5), τC$_\beta$H$<em>2$ (5), −δ$</em>\alpha$N$_3$C$_3$(O)$(C)$ (5)</td>
<td>ρNH$_3$/ρCH$_2$</td>
</tr>
<tr>
<td>998</td>
<td>989</td>
<td>ρNH$<em>3$ (35), νC$</em>\alpha$C$<em>\beta$ (25), −νC$</em>\alpha$C (11)</td>
<td>ρNH$_3$/νC-C</td>
</tr>
</tbody>
</table>

$^a$Frequencies correspond to visible Raman (633 nm excitation) spectrum. $^b$: stretch; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.

Table C12: Raman Frequencies (cm$^{-1}$) and Assignments of Crystalline L-glycyl-L-glutamine$^a$. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PED$^b$ ($\geq$5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1201</td>
<td>1201</td>
<td>τC$<em>\beta$H$<em>2$ (33), −ρC$</em>\alpha$H (24), τC$</em>\gamma$H$<em>2$ (15), τC$</em>\gamma$H$<em>2$ (6), νC$</em>\alpha$C$_\beta$ (5)</td>
<td>τCH$_3$/ρCH</td>
</tr>
<tr>
<td>1171</td>
<td>1147</td>
<td>−νC$<em>\alpha$N (37), δNC$</em>\alpha$C (7), −ρN$_2$H$<em>2$ (7), −νC$</em>\alpha$N (7)</td>
<td>νC-N</td>
</tr>
<tr>
<td>1139</td>
<td>1138</td>
<td>νC$<em>\alpha$C$</em>\beta$ (26), −νC$<em>\beta$C$</em>\gamma$ (19), −ρC$<em>\beta$H$<em>2$ (8), −ρN$<em>2$H$<em>2$ (6), −δC$</em>\alpha$C$</em>\beta$C$</em>\beta$ (6), −νC$</em>\alpha$N (6), νC-$\beta$</td>
<td></td>
</tr>
<tr>
<td>1124</td>
<td>1128</td>
<td>νC$<em>\alpha$N (37), τC$</em>\alpha$H$<em>2$ (18), −ρC$</em>\alpha$H$<em>2$ (13), −δC$</em>\alpha$C$_\beta$C(OO) (5)</td>
<td>νC-C</td>
</tr>
<tr>
<td>1105</td>
<td>1105</td>
<td>ρN$_2$H$_2$ (29), νN$_2$H$<em>2$ (12), −τC$</em>\alpha$H$<em>2$ (7), −νC$</em>\alpha$H$<em>2$ (6), −τC$</em>\alpha$H$<em>2$ (6), −νC$</em>\alpha$N (6), −ρN$<em>2$C$</em>\alpha$C (5)</td>
<td>ρNH$_3$</td>
</tr>
<tr>
<td>1093</td>
<td>1090</td>
<td>ρN$_2$H$_2$ (33, 13, −ρN$<em>2$H$<em>2$ (12), −ρC$</em>\alpha$H$<em>2$ (13), −ρC$</em>\alpha$C$</em>\beta$ (5)</td>
<td>AmH$^p$</td>
</tr>
<tr>
<td>1082</td>
<td>1081</td>
<td>−νC$<em>\alpha$N (14, 13), νC$</em>\beta$C$<em>\gamma$ (13), νC$</em>\alpha$C$<em>\alpha$ (8), ρC$</em>\alpha$H$<em>2$ (7), νC$</em>\alpha$C$<em>\beta$ (7), ρC$</em>\alpha$H (6), −νC$_\alpha$N (5)</td>
<td>νC-N/νC-C</td>
</tr>
<tr>
<td>1048</td>
<td>1048</td>
<td>−τC$<em>\alpha$N$<em>2$ (56), νC$</em>\alpha$N (9), −τC$</em>\alpha$C$<em>\beta$ (7), −τC$</em>\alpha$C$_\gamma$ (5)</td>
<td>νC-C</td>
</tr>
<tr>
<td>1027</td>
<td>1013</td>
<td>ρC$<em>\beta$H$<em>2$ (28), δC$</em>\alpha$C$</em>\beta$C$<em>\gamma$ (9), τC$</em>\gamma$H$<em>2$ (8), −ρC$</em>\beta$C$<em>\gamma$ (7), ρC$</em>\alpha$H$<em>2$ (5), δNC$</em>\alpha$C$_\beta$ (5)</td>
<td>ρCH$_3$</td>
</tr>
<tr>
<td>997</td>
<td>946</td>
<td>νC$^\prime$C$<em>\alpha$ (19), −ρC$</em>\alpha$H$<em>2$ (10), νC$^\prime$C$</em>\alpha$ (9), −ρC$^\prime$C$<em>\alpha$C$</em>\alpha$ (6), −βC$^\prime$C$<em>\alpha$C$</em>\alpha$N (6), −ρN$_2$H$_2$ (5)</td>
<td>νC-C/νC$_\gamma$</td>
</tr>
</tbody>
</table>

$^a$Frequencies correspond to visible Raman (633 nm excitation) spectrum. $^b$: stretch; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.
Table C13: Raman Frequencies (cm\(^{-1}\)) and Assignments of Crystalline L-seryl-L-asparagine\(^a\). Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Calc.</th>
<th>PED(^b) (≥5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1188</td>
<td>1184</td>
<td>−ρ(<em>{C\beta'}H_2) (26), ρ(</em>{\prime N' H_3}) (14), −δNC(<em>\alpha'C') (9), −τC(</em>\beta'H_2) (8), βC(<em>\beta'O\gamma'H) (7), ρ(</em>{\prime C\alpha'H}) (6), νC(<em>\alpha'N) (5), ρ(</em>{N' H_3}) (5)</td>
<td>ρCH(_2/\rho NH_3)</td>
</tr>
<tr>
<td>1159</td>
<td>1164</td>
<td>−νC(<em>\alpha N) (37), ρ(</em>{\prime N' H_3}) (7), νC(<em>\alpha C\beta) (7), δCC(</em>\alpha N) (5), ρ(_{C\alpha' H}) (5)</td>
<td>νC–N</td>
</tr>
<tr>
<td>1121</td>
<td>1119</td>
<td>−νC(<em>\beta'O\gamma'H) (17), ρ(</em>{\prime N' H_2}) (13), νC(<em>\alpha C\beta) (13), νC(</em>\alpha'N') (8), −ρ(<em>{\prime C\alpha' H}) (6), ρ(</em>{N' H_3}) (5), −ρ(_{\prime N' H_3}) (5)</td>
<td>νC–OH/ρNH(_2)</td>
</tr>
<tr>
<td>1108</td>
<td>1116</td>
<td>νC(<em>\beta'O\gamma'H) (66), ρ(</em>{N' H_2}) (9)</td>
<td>νC–OH</td>
</tr>
<tr>
<td>1083</td>
<td>1107</td>
<td>−ρ(<em>{N_4 H_2}) (16), νC(</em>\alpha'N') (10), −ρ(<em>{C\alpha'H}) (10), νC(</em>\beta'O\gamma'H) (8), ρ(<em>{N' H_3}) (8), −ρ(</em>{C\alpha' N}) (6), ρ(<em>{\prime N' H_2}) (9), −νC(</em>\alpha N) (5)</td>
<td>ρNH(_2/\nu C–N)</td>
</tr>
<tr>
<td>1065</td>
<td>1083</td>
<td>−ρ(<em>{N' H_4}) (20), νC(</em>\alpha'N') (10), −νC(<em>\gamma' C\alpha') (9), νC(</em>\alpha C\beta) (8), ρ(<em>{C\alpha' H}) (7), −νC(</em>\beta'O\gamma'H) (5), ρ(_{C\alpha' H}) (5)</td>
<td>νC–OH/ρC–N</td>
</tr>
<tr>
<td>1051</td>
<td>1069</td>
<td>νC(<em>\alpha C\beta) (23), −νC(</em>\alpha'N') (11), νC(<em>\alpha N) (9), −δCC(</em>\alpha N) (7), −τC(<em>\alpha C\beta C\gamma) (5), −ρ(</em>{N_4 H_2}) (5)</td>
<td>AmIII(^P)</td>
</tr>
<tr>
<td>1019</td>
<td>1023</td>
<td>−ρ(<em>{C\beta'H_2}) (23), δC(</em>\alpha' C\gamma) (12), δC(<em>\beta' C\gamma) (14), −ρ(</em>{C\alpha' C}) (7), δ(<em>\delta C\alpha COO) (6), βC(</em>\alpha NC') (6), ρ(_{C\alpha C\gamma}) (5)</td>
<td>ρCH(_2/\delta CCC)</td>
</tr>
<tr>
<td>1004</td>
<td>981</td>
<td>ρ(<em>{C\beta'H_2}) (28), ρ(</em>{\prime N' H_3}) (25), −νC(<em>\alpha C\beta') (14), ρ(</em>{N' H_3}) (10)</td>
<td>ρCH(_2/\rho NH_3)</td>
</tr>
</tbody>
</table>

\(^{a}\) Frequencies correspond to visible Raman (633 nm excitation) spectrum. \(^{b}\) ν: stretch; δ\(_s\): symmetric deformation; δ: deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.
Table C14: UVRR Frequencies (cm\(^{-1}\)) and Assignments of Butyramide in Acetonitrile and Water. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.

<table>
<thead>
<tr>
<th>H\textsubscript{2}O</th>
<th>CH\textsubscript{3}CN</th>
<th>Calc.</th>
<th>PED(^b) (≥5% contribution)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1132</td>
<td>1125</td>
<td>1143</td>
<td>(\nu)C(<em>\alpha)C(</em>\beta) (27), (\rho)C(<em>\alpha)H(<em>3) (19), (\rho)N(</em>\epsilon)H(<em>2) (14), (-\sigma)C(</em>\alpha)C(</em>\beta)C(<em>\gamma) (8), (-\sigma)C(</em>\beta)C(<em>\gamma)C(</em>\delta) (6), (\omega)C(_\beta)H(_2) (6)</td>
<td>(\nu)C–C/(\rho)CH(_3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>–C(<em>\beta)C(</em>\gamma) (5)</td>
<td></td>
</tr>
<tr>
<td>1126</td>
<td></td>
<td></td>
<td>(-\rho)C(<em>\alpha)H(<em>3) (17), (\rho)C(</em>\beta)H(<em>2) (17), (\rho)C(</em>\gamma)H(<em>2) (17), (\tau)C(</em>\gamma)H(<em>2) (16), (-\delta)N(</em>\epsilon)C(</em>\delta)C(<em>\gamma) (7), (\tau)C(</em>\beta)H(_2) (7)</td>
<td>(\rho)CH(_3)/(\rho)CH(_2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-\nu)C(<em>\beta)C(</em>\gamma) (6)</td>
<td></td>
</tr>
<tr>
<td>1076</td>
<td>1064</td>
<td>1092</td>
<td>(\rho)N(<em>\epsilon)H(<em>2) (37), (\nu)C(</em>\delta)N(</em>\epsilon) (17), (-\nu)C(<em>\delta)C(</em>\gamma) (14), (-\tau)C(_\gamma)H(_2) (6)</td>
<td>AmIII(^P)</td>
</tr>
<tr>
<td>1050</td>
<td>1040</td>
<td>1067</td>
<td>(\nu)C(<em>\alpha)C(</em>\beta) (59), (-\nu)C(<em>\beta)C(</em>\gamma) (10), (\rho)C(_\alpha)H(<em>3) (9), (-\rho)N(</em>\epsilon)H(<em>2) (6), (\omega)C(</em>\gamma)H(_2) (6)</td>
<td>(\nu)C–CH(_3)</td>
</tr>
</tbody>
</table>

\(^a\)Frequencies correspond to UVRR (204 nm excitation) spectra. \(^b\) \(\nu\): stretch; \(\delta\): symmetric deformation; \(\delta\): deformation; \(\sigma\): scissoring; \(\rho\): rocking; \(\omega\): wagging; \(\beta\): in-plane bending; \(\tau\): twisting.
Figure C1: AmIII$^P$ frequency dependence on the OCCC dihedral angle of butyramide from DFT calculations. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C2: Crystal structures of Gln and Gln derivatives examined in this study. The atomic labeling schemes shown are used for Table C8–Table C13. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C3: OTREP diagram of GlnTBE with atomic labeling scheme used in Table C4–Table C7. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C4: Visible Raman (633 nm excitation) spectra of compounds examined in this study. (a) L-Gln; (b) D-Gln; (c) NAcGln; (d) GlnTBE; (e) Gly-Gln; and (f) Ser-Asn. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C5: UVRR (229 nm excitation) spectra of compounds examined in this study. (a) L-Gln; (b) D-Gln; (c) NAcGln; (d) GlnTBE; and (e) Ser-Asn. CAdapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C6: UVRR (204 nm excitation) spectra of butyramide in (a) water and (b) acetonitrile. The AmIII$^P$ of butyramide downshifts 12 cm$^{-1}$ as the solvent transfers from acetonitrile to water. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
Figure C7: Temperature dependent Circular Dichroism (CD) spectra of Q₃. Adapted with permission from [180]. Copyright © (2015), American Chemical Society.
SUPPORTING INFORMATION FOR CHAPTER 5.0

D.1 EXPERIMENTAL SECTION

D.1.1 UVRR spectroscopy of fibril films

Solutions of NDQ10 and DQ10 fibril aggregates were aliquoted onto the grooves of brass cylindrical cells. The solutions were dried for \( \sim 5 \) h, after which a film could be observed. The UVRR instrumentation used to collect spectra of NDQ10 and DQ10 fibril films was the same as described in the main text. The laser light was focused onto the grooves of the brass cells, which were spun in order to prevent accumulation of photochemical or thermal degradation products. UVRR spectra were collected using both \( \sim 197 \) nm and \( \sim 204 \) nm excitation.

D.1.2 UVRR Spectral Processing

All UVRR spectra were processed using home-written MATLAB scripts in order to remove cosmic rays, average and calibrate spectra, as well as subtract the spectral contributions of water (\( \text{e.g.} \ H_2O, D_2O, \) and HDO) and Suprasil quartz from NMR tubes. The spectra were calibrated using the 801.3 cm\(^{-1}\), 1028.3 cm\(^{-1}\), 1157.6 cm\(^{-1}\), 1266.4 cm\(^{-1}\), and 1444.4 cm\(^{-1}\) bands of cyclohexane. The spectral contributions of water and quartz were removed using a method similar to that described by Banerjee and coworkers [222, 223].

To subtract the contributions of water and quartz, we first calculated the first-derivatives of the spectra. The relative contributions of water and quartz in the raw spectra were found
via a classical multiple linear least-squares regression such that:

\[ S' = KS'_p \]  

(D.1)

where \( S' \) is the \((n \times 1)\) row vector that represents the first-derivative (denoted by \( ' \)) of the experimentally measured raw spectrum, and \( S'_p \) is the \((n \times m)\) matrix composed of row vector elements that contain the first-derivative spectra of the pure water and quartz spectra. The \((m \times 1)\) row vector, \( K \), contains the least-squares scaling coefficients for each of the different background spectral components. The least-squares solution of \( K \) for eq. D.1 is [224]:

\[ K = (S'_p^T S'_p)^{-1} S'_p^T S' \]  

(D.2)

where \((S'_p^T S'_p)^{-1} S'_p^T\) is the pseudo-inverse matrix of \( S'_p \). Banerjee and coworkers [222, 223] discuss the advantages of using the first-derivative spectra to determine relative contributions of different spectral components. Using the first-derivative spectra is most advantageous for spectra that contain multiple, overlapping spectral components.

After determining \( K \), eq. D.3 was utilized to subtract the contributions of water and quartz to obtain the spectrum of the analyte of interest:

\[ S_{\text{analyte}} = S - KS_p \]  

(D.3)

where \( S_{\text{analyte}} \), \( S \), and \( S_p \) are the zeroth-derivative analyte, raw, and water/quartz UVRR spectra, respectively.

### D.1.3 UVRR Spectral Peak Fitting

The GRAMS AI software suite (ver. 8.0, Thermo Fisher Scientific) was used to peak fit the UVRR spectra. The spectra, \( S(\nu) \), were parsimoniously fit as the sum of pure Gaussian and Lorentzian bands. \textit{i.e.}

\[ S(\nu) = \sum_i \left[ f_i H_i e^{-\left(\frac{\nu - \nu_i}{w_i}\right)^2 (4 \ln(2))} + (1 - f_i) \frac{H_i}{4\left(\frac{\nu - \nu_i}{w_i}\right)^2 + 1} \right] \]  

(D.4)

where \( f_i = 1 \) if the \( i^{th} \) band is a Gaussian, or 0 if the \( i^{th} \) band is a Lorentzian. The parameters \( H_i, \nu_i, \) and \( w_i \) are the heights, center frequencies, and widths, respectively, of the \( i^{th} \) band.
D.1.4 Calculation of the $\Psi$ and $\chi_3$ Angle Distributions

The distributions of $\Psi$ and $\chi_3$ dihedral angles, shown in Figure 5.7 and Figure 5.11 in the main text, were calculated using methodologies previously described in detail [96, 97, 105, 180]. Briefly, we assume that the inhomogeneously broadened, experimentally measured AmIII$_3^S$ and AmIII$_3^P$ band profiles, $B(\nu)$, can be modeled as the sum of $M$ Lorentzian bands:

$$B(\nu) = \frac{1}{\pi} \sum_i p_i \Gamma^2 \left( \frac{\nu^2}{\Gamma^2} + (\nu - \nu_i)^2 \right)$$  \hspace{1cm} (D.5)

where $p_i$ is the probability for the $i^{th}$ band to occur at center frequency $\nu_i$. The band width parameter, $\Gamma$, is the homogeneous linewidth of the AmIII$_3^S$ or AmIII$_3^P$ vibrations. We previously [105, 180] estimated from peptide crystals that $\Gamma \approx 7.5 \text{ cm}^{-1}$ for the AmIII$_3^S$ and $\sim 6.6 \text{ cm}^{-1}$ AmIII$_3^P$. After decomposing the band profiles into Lorentzians, we then correlate the different $i^{th}$ frequencies of the AmIII$_3^S$ and AmIII$_3^P$ band envelopes to their respective $\Psi$ or $\chi_3$ dihedral angles.

D.1.4.1 Correlating the AmIII$_3^S$ Frequencies to $\Psi$ Angles

We used the following equation to correlate the AmIII$_3^S$ band frequencies to $\Psi$ angles for the Figure 5.7 black distributions shown in the main text:

$$\nu_i(\Psi) = 1239 \text{ (cm}^{-1}) - 54 \text{ (cm}^{-1}) \sin(\Psi + 26^\circ)$$  \hspace{1cm} (D.6)

Eq. D.6 was derived by Mikhonin et al. [97] for situations when there is strong peptide-peptide hydrogen bonding, such as in the case for fibril peptide bonds.

The $\Psi$ angles for the Figure 5.7 distributions shown in blue were obtained by using:

$$\nu_i(\Psi, T) = 1250 \text{ (cm}^{-1}) - 54 \text{ (cm}^{-1}) \sin(\Psi + 26^\circ) + 0.06 \text{ (cm}^{-1}/^\circ C)(T - T_0)$$  \hspace{1cm} (D.7)

where $T = 22^\circ C$ is the experimental temperature and $T_0 = 0^\circ C$. Eq. D.7 was derived by Mikhonin et al. [97] for situations when the hydrogen bonding state of the peptide bond N–H groups is unknown. This situation occurs, for example, in the case of solvent accessible fibril peptide bonds. It is unknown if these peptide bonds are exchangeable in solvent because
they are located on the surface of fibrils, or whether they are more in more disordered regions of the aggregates.

**D.1.4.2 Correlating the AmIII\textsuperscript{P} Frequencies to $\chi_3$ Angles** In the case of the AmIII\textsuperscript{P}, we used the following equation to correlate the band frequencies to $\chi_3$ angles for the distributions shown in Figure 5.11

$$
\nu_i(\chi_3) = 1076 \text{ (cm}^{-1}) + 29 \text{ (cm}^{-1}) \cos(2\chi_3) + 9 \text{ (cm}^{-1}) \cos(\chi_3 + 99^\circ)
$$

where $\nu_i$ is the $i^{th}$ AmIII\textsuperscript{P} frequency. Eq. D.8 was derived by Punihaole et al. [180] as an “average” expression to be used in situations when the hydrogen bonding and dielectric environments of Gln side chains are unknown. We previously showed [146] that stronger (weaker) hydrogen bonding and higher (lower) dielectric environments upshift (downshift) the AmIII\textsuperscript{P} frequency. In the case of Q10 fibrils, the inter-amide hydrogen bonding of the Gln side chains is strong, but the dielectric constant of the environment is also low. Thus, we utilized eq. D.8 since it averages these two competing effects. We are presently investigating which of these effects dominates the AmIII\textsuperscript{P} frequency in polyQ fibrils.

It should be noted that in using eq. D.8, each AmIII\textsuperscript{P} frequency can be correlated to as many as four possible $\chi_3$ dihedral angles. However, as discussed in detail by Punihaole et al. [180], $\chi_3$ dihedral angles that are greater than $+90^\circ$ and less than $-90^\circ$ are nearly forbidden for Gln and Asn. Thus, in using eq. D.8, we only considered the physically relevant $\chi_3$ angles that occur between $-90^\circ$ and $+90^\circ$.

**D.2 COMPUTATIONAL SECTION**

**D.2.1 Density Functional Theory (DFT) Calculations**

DFT calculations [135] were performed on the zwitterion form of the Gln amino acid (Figure D1) using the GAUSSIAN 09 program [136]. The M06-2X density functional was employed using the 6-311++g** basis set [137]. Water was modeled implicitly by placing the
Gln molecule in an ellipsoidal cavity surrounded by a polarizable continuum dielectric model. The calculated frequencies were not scaled. The potential energy distribution (PED) of each vibration was obtained from the *GAUSSIAN 09* output files by employing a MATLAB script that has been previously published [180].

### D.2.2 RMSD Metric

The extent of model fibril dissociation was quantified with a root mean square deviation (RMSD) metric. The equation for RMSD is shown and this metric corresponds to the spatial deviation of atoms.

\[
RMSD = \frac{1}{N} \sqrt{\sum_{i=1}^{N} [(x_i - x_{\text{ref}})^2 + (y_i - y_{\text{ref}})^2 + (z_i - z_{\text{ref}})^2]}
\]  

where \( N \) is the number of atoms used in the RMSD calculation, \( x_i, y_i, z_i \) are the current coordinate positions of atom \( i \), and \( x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}} \) are the coordinate positions of atom \( i \) in the reference structure. Backbone atoms used were \( C_\alpha \), the carbonyl carbon, the carbonyl oxygen, and the peptide backbone nitrogen. For our peptide system, there are 320 atoms: four backbone atoms per residue, ten Gln residues per peptide, and eight peptides per fibril model.

Before the RMSD measurement was taken for each step, the model fibril was superimposed on the initial reference structure to eliminate the effect of fibril translation and rotation on the RMSD value. The interpretation of this metric was that a rising RMSD indicates fibril dissociation and lack of stability, whereas a constant RMSD signifies a stable fibril structure. We assigned a RMSD ceiling of 3 Å, and when a fibril model’s RMSD increased above this value it was judged to be dissociated. Figure D2 shows that the antiparallel and parallel \( \beta \)-sheet structures stayed well below the RMSD limit, while the \( \beta \)-hairpin model dissociated at \( \sim 58 \) ns.
D.3 RESULTS AND DISCUSSION

D.3.1 UVRR of NDQ10 and DQ10 Fibril Films

Figure D3 shows the 197 nm – 204 nm UVRR difference spectra of dried NDQ10 and DQ10 fibril films. The AmI$^P$ and AmII$^P$ bands are located at $\sim$1660 cm$^{-1}$ and $\sim$1612 cm$^{-1}$, respectively, for both NDQ10 and DQ10. These bands negligibly shift compared to the AmI$^P$ and AmII$^P$ bands of fibrils in solution, as shown in the Figure 5.3 197 nm – 204 nm UVRR difference spectra in the main text. This indicates that the hydrogen bonding environments of the Gln side chain primary amides are not significantly perturbed upon dehydrating the fibrils. This occurs because there is strong side chain inter-amide hydrogen bonding in NDQ10 and DQ10 fibrils.

The CH$_2$ wagging band is located at $\sim$1413 cm$^{-1}$ for DQ10 fibril films and at $\sim$1410 cm$^{-1}$ for NDQ10 fibril films. Compared to the Figure 5.3 197 nm – 204 nm UVRR difference spectra in the main text, the CH$_2$ wagging band downshifts $\sim$5 cm$^{-1}$ for NDQ10 and $\sim$17 cm$^{-1}$ for DQ10 fibrils upon dehydration. We attribute this $\sim$17 cm$^{-1}$ downshift of the CH$_2$ wagging band in DQ10 fibrils upon dehdyration to a local dielectric environment change of the methylene groups around of the Gln side chains.

D.3.2 Band Assignments of Mono-deuterated Primary Amides

We employed DFT calculations to aid in our band assignments of the Gln UVRR spectrum in 50%/50% H$_2$O/D$_2$O (Figure 5.5 in the main text). In assigning the Figure 5.5 spectrum, we assume that vibrations containing significant contributions of C$_\delta$–N$_\epsilon^2$ stretching show resonance enhancement in the the UVRR spectrum because the electronic excited state is expected to be expanded along this coordinate [147]. Table D1 and Table D2 show the potential energy distributions (PEDs) obtained from the DFT calculations for the “cis-N$_\epsilon^2$HD” and “trans-N$_\epsilon^2$HD” species of the mono-deuterated primary amide side chains of Gln. Our band assignments of the Figure 5.5 spectrum from the main text are shown in Table D3.
D.3.2.1 Assignment of Amide Vibrations  The DFT calculations show that partial deuteration of the Gln side chains results in a reorganization of the eigenvector composition of the primary amide vibrations compared to fully protonated side chains. This results in the decoupling of N–H and N–D motions and the appearance of vibrations that resemble canonical secondary amide modes. Our findings agree with Saito and coworkers’ normal mode analyses [211, 212] of partially deuterated acetamide.

We assign the AmI vibration to a band located at $\sim$1660 cm$^{-1}$. Our normal mode analysis indicates that the PED of this vibration consists mostly of $C_\delta=O_\epsilon$ stretching ($\sim$77%), but also contains significant contributions of $C_\delta-N_\epsilon$ stretching and $N_\epsilon C_\delta C_\gamma$ bending. The PED of this vibration is essentially the same as that of the AmIF$^S$ and AmIP$^S$ vibrations [146, 180].

The DFT calculations also predict the appearance of AmII$^S$- and AmIII$^S$-like vibrations. Both peptide backbone C–N stretching and N–H in-plane bending motions are important in defining the PEDs of the canonical AmII$^S$ and AmIII$^S$ vibrations. Therefore, in the case of the mono-deuterated primary amides, we searched for vibrations that contain significant contributions of $C_\delta-N_\epsilon$ stretching and $N_\epsilon$CHD deformations. As shown in Table D1 and Table D2, there are several vibrations that contain significant contributions of $C_\delta-N_\epsilon$ stretching, $N_\epsilon$HD scissoring, and $N_\epsilon$HD rocking. We assign these vibrations to AmII$^S$-like, AmIII$^S$-like, or AmIII$^{S'}$-like vibrations.

There are two AmII$^S$-like vibrations predicted by the DFT calculations to be at 1524 cm$^{-1}$ for the trans-$N_\epsilon$HD species and $\sim$1479 cm$^{-1}$ for the cis-$N_\epsilon$HD species. Both of these vibrations contain significant $C_\delta-N_\epsilon$ stretching and $N_\epsilon$HD scissoring, although the 1524 cm$^{-1}$ mode also contains $N_\epsilon$HD rocking. The predicted $\sim$1524 cm$^{-1}$ mode is experimentally observed at $\sim$1547 cm$^{-1}$, while the predicted $\sim$1479 cm$^{-1}$ vibration is observed at $\sim$1476 cm$^{-1}$.

The DFT calculations indicate that two AmIII$^S$-like vibrations for the trans-$N_\epsilon$HD species are predicted to occur at $\sim$1247 cm$^{-1}$ and $\sim$1329 cm$^{-1}$. Both vibrations contain significant contributions of $C_\delta-N_\epsilon$ stretching and $N_\epsilon$HD scissoring. However, as with the canonical AmIII$^S$ modes observed in peptides, these vibrations are significantly coupled since they contain significant contributions of CH$_2$ wagging and twisting, as well as $C_\alpha$–H rocking [225]. We assign these AmIII$^S$-like vibrations to bands observed at $\sim$1247 cm$^{-1}$ and $\sim$1308 cm$^{-1}$.
There are also two AmIII* S-like vibrations predicted to be at $\sim 1055 \text{ cm}^{-1}$ for the trans-$N_{\epsilon 2}$HD species and $\sim 953 \text{ cm}^{-1}$ for the cis-$N_{\epsilon 2}$HD species. These vibrations resemble AmIII* S-like modes since they both contain large contributions of $N_{\epsilon 2}$HD rocking. This is analogous to the canonical AmIII* S, which is mostly N-D in-plane bending.

### D.3.3 Bennett Acceptance Ratio Method

The Bennett acceptance ratio (BAR) is used here to estimate the free energy difference between two states. The full equation for the BAR is

$$\sum_{i=1}^{n_i} \frac{1}{1 + \exp(\ln(n_i/n_j) + \beta \Delta U_{ij} - \beta \Delta G)} - \sum_{i=1}^{n_j} \frac{1}{1 + \exp(\ln(n_j/n_i) + \beta \Delta U_{ji} - \beta \Delta G)} = 0$$

where $\Delta G$ is the Gibbs free energy difference between states i and j (here antiparallel and parallel $\beta$-sheet fibrils), $n_i, n_j$ are the number of samples used for states i and j, $\beta = k_B T$ (where $k_B$ is the Boltzmann constant and $T$ is the simulation temperature), and $\Delta U_{i,j}, U_{j,i}$ are the potential energy differences between states i and j. This equation is solved numerically using an iterative method.

Bennett clearly states that the best estimates of the free energy differences between states occurs when the extent of the energy overlap is greatest and when the density-of-states as a function of the energy difference is smoothest [226]. Here, we employ pymbar, which utilizes a multistate Bennett acceptance ratio method that can handle two or more states. Since we are working with two states, fibril models $a$ and $b$, the multistate method is identical to the traditional BAR method derived for two states [207]. Figure D4 demonstrates the overlapping potential energy distributions necessary for a converged BAR calculation. To ensure comparable energetics, identical atom counts as well as system dimensions were used for all fibril simulations (see main text and NAMD configuration files for details).

### D.3.4 Hydrogen Bonding Analysis

Figure D5 show the number of hydrogen bonds of fibril models $a$ and $b$ (from Figure 1.1 in the main text), which were obtained from the MD trajectories. Table D4 lists the average
number of the different categories of hydrogen bonds formed during the MD simulations for fibril models a and b. As discussed in the main text, antiparallel β-sheet model a forms, on average, more peptide bond-peptide bond hydrogen bonds than the parallel β-sheet model b. In particular, the antiparallel β-sheet forms significantly more peptide bond-peptide bond hydrogen bonds than the parallel β-sheet. It is interesting to note that, if we assume a peptide bond-peptide bond hydrogen bond energy of 5 kJ mol$^{-1}$, this hydrogen bonding difference would account for the majority of the 160.5 kJ mol$^{-1}$ free energy difference between the antiparallel and parallel β-sheet fibril models.
Table D1: DFT Calculated Frequencies (cm\(^{-1}\)) and Assignments of cis-Glutamine-N\(_{2}\)HD. Adapted with permission from [196].

Copyright © (2016), American Chemical Society.

<table>
<thead>
<tr>
<th>Assign.</th>
<th>Calc.</th>
<th>PED(^a,b) (≥ 5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\nu_{\alpha\alpha}) COO</td>
<td>1715</td>
<td>(-\nu C\text{'}O (53), \nu C\text{'}O (33), \rho C\alpha C\text{'} (7))</td>
</tr>
<tr>
<td>(\delta_{\alpha\alpha}) NH(_3)</td>
<td>1666</td>
<td>(-\delta_{\alpha\alpha}) NH(<em>3) (48), (\delta</em>{\alpha\alpha}) NH(_3) (44), (\rho) NH(_3) (5)</td>
</tr>
<tr>
<td>(\delta_{\alpha\alpha}) NH(_3)</td>
<td>1621</td>
<td>(-\delta_{\alpha\alpha}) NH(<em>3) (40), (-\delta</em>{\alpha\alpha}) NH(<em>3) (40), (-\delta</em>{3}) NH(_3) (10)</td>
</tr>
<tr>
<td>(\sigma) CH(_2)</td>
<td>1494</td>
<td>(\sigma C\beta H_2 (87))</td>
</tr>
<tr>
<td>AmII-like</td>
<td>1479</td>
<td>(-\sigma N_{2}) HD (37), (-\nu C\beta N_{2} (28), \nu C\gamma C\delta (10), (-\omega C\gamma H_2 (9), \rho C\delta O\epsilon_1 (6))</td>
</tr>
<tr>
<td>(\delta_{\alpha}) NH(_3)</td>
<td>1464</td>
<td>(\delta_{\alpha}) NH(_3) (50), (-\sigma C\gamma H_2 (20))</td>
</tr>
<tr>
<td>(\sigma) CH(_2)</td>
<td>1460</td>
<td>(\sigma C\alpha H_2 (61), \delta_{3}) NH(_3) (21)</td>
</tr>
<tr>
<td>(\omega) CH(_2)</td>
<td>1437</td>
<td>(\omega C\beta H_2 (19), \sigma N_{2}) HD (14), (-\omega C\gamma H_2 (11), \nu C\text{'}O (10), \rho C\alpha H (8), \nu C\beta C\gamma (5), (-\delta NH_3 (5), \sigma C\gamma H_2 (5))</td>
</tr>
<tr>
<td>(\nu_{\alpha}) COO</td>
<td>1414</td>
<td>(\nu C\text{'}O (22), \omega C\gamma H_2 (13), \sigma N_{2}) HD (11), (-\nu C\alpha C\gamma (9), (-\rho C\alpha H (7), \nu C\gamma C\delta (7), (-\beta C\text{'}O_2 (6), \nu C\text{'}O (6))</td>
</tr>
<tr>
<td>(\rho) CH</td>
<td>1387</td>
<td>(\rho C\alpha H (28), \omega C\beta H_2 (11), \rho C\beta H (9), \nu C\text{'}O (8), \nu C\alpha C\beta (7), (-\tau C\beta H_2 (6))</td>
</tr>
<tr>
<td>(\rho) CH</td>
<td>1357</td>
<td>(\rho C\alpha H (27), \tau C\beta H_2 (19), \omega C\beta H_2 (13), (-\tau C\beta H_2 (11))</td>
</tr>
<tr>
<td>(\omega) CH(_3)+(\nu) CC</td>
<td>1347</td>
<td>(\omega C\beta H_2 (21), \tau C\beta H_2 (21), \rho C\alpha H (11), \nu C\beta N_{2} (11), (-\rho C\delta O\epsilon_1 (5))</td>
</tr>
<tr>
<td>(\tau) CH(_2)</td>
<td>1309</td>
<td>(\tau C\alpha H_2 (35), \rho C\gamma H (26), (-\rho C\beta H_2 (8), \rho C\alpha H (6))</td>
</tr>
<tr>
<td>(\omega) CH(_2)</td>
<td>1272</td>
<td>(\omega C\gamma H_2 (43), \omega C\beta H_2 (20), (-\nu C\beta N_{2} (13))</td>
</tr>
<tr>
<td>(\tau) CH(_2)+(\nu) CC</td>
<td>1214</td>
<td>(\tau C\beta H_2 (21), \nu C\alpha C\beta (18), \tau C\gamma H_2 (10), \rho NH_3 (13), (-\delta NC\alpha C\text{'}O_2 (5))</td>
</tr>
<tr>
<td>(\rho) CH+(\tau) CH(_2)</td>
<td>1153</td>
<td>(-\rho C\alpha H (26), \tau C\beta H_2 (16), (-\rho NH_3 (14), \tau C\beta H_2 (12), (-\nu C\beta C\beta (8), \rho C\gamma H_2 (6))</td>
</tr>
<tr>
<td>(\nu) CC</td>
<td>1117</td>
<td>(\nu C\beta C\gamma (42), \nu C\alpha N (10), \rho NH_3 (8), (-\nu C\gamma C\delta (7), (-\delta NC\alpha C\text{'}O_2 (5), (-\rho C\alpha C\beta C\gamma (5))</td>
</tr>
<tr>
<td>(\rho) NH(_3)</td>
<td>1105</td>
<td>(\rho NH_3 (28), (-\nu C\beta C\gamma (12), (-\rho C\alpha H (9), (-\delta NC\alpha C\text{'}O_2 (8), (-\rho C\alpha H (8), (-\rho C\gamma H_2 (7))</td>
</tr>
<tr>
<td>(\nu) CN</td>
<td>1040</td>
<td>(\nu C\beta N (35), (-\nu C\gamma H (11), \rho C\beta H_2 (9), \rho C\gamma H_2 (6), \rho C\alpha H (5))</td>
</tr>
<tr>
<td>(\rho) NH(_3)+(\nu) CN</td>
<td>1005</td>
<td>(\rho NH_3 (22), (-\nu C\beta N (20), \rho C\gamma H_2 (12), \rho C\beta H_2 (12), \nu C\alpha C\beta (10))</td>
</tr>
<tr>
<td>(\rho) NH(_4)</td>
<td>983</td>
<td>(\rho NH_3 (29), \nu C\beta C\beta (13), (-\sigma C\alpha C\gamma C\gamma (6), \rho N_{2}) HD (6), (-\sigma C\beta C\gamma C\delta (5), \nu C\gamma C\delta (5))</td>
</tr>
<tr>
<td>AmIII(^8)-like</td>
<td>953</td>
<td>(\rho N_{2}) HD (15), (\nu C\beta N_{2} (14), (-\nu C\beta C\beta (13), \rho NH_3 (11), \nu C\gamma C\delta (9), \nu C\alpha C\gamma (8), (-\sigma N_{2}) HD (7)</td>
</tr>
</tbody>
</table>

\(^a\)C\beta\)-N\(_{2}\) stretching and N\(_{2}\) HD scissoring components in PED are in bold.

\(^b\)\(\nu\): stretch; \(\delta_{\alpha\alpha}\): asymmetric deformation; \(\delta_{\alpha}\): symmetric deformation; \(\delta\): deformation; \(\sigma\): scissoring; \(\rho\): rocking; \(\omega\): wagging; \(\beta\): in-plane bending; \(\tau\): twisting.
Table D2: DFT Calculated Frequencies (cm\(^{-1}\)) and Assignments of trans-Glutamine-N\(_2\)HD. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

<table>
<thead>
<tr>
<th>Assign.</th>
<th>Calc.</th>
<th>PED(^{a,b}) (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AmII-like</td>
<td>1742</td>
<td>(-\nu CN_\beta C_\gamma (77), -\beta NC_\gamma C_\delta (7), \nu C_\beta N_2 (7))</td>
</tr>
<tr>
<td>AmII-like</td>
<td>1715</td>
<td>(-\nu C'O (53), \nu C'O (33), \rho C_\gamma C_\delta (7))</td>
</tr>
<tr>
<td>AmII-like</td>
<td>1666</td>
<td>(\delta_\alpha' NH_3 (48), -\delta_\alpha' NH_3 (44), -\rho NH_3 (5))</td>
</tr>
<tr>
<td>AmII-like</td>
<td>1621</td>
<td>(-\delta_\alpha' NH_3 (40), -\delta_\alpha' NH_3 (40), -\delta_\alpha' NH_3 (10))</td>
</tr>
<tr>
<td>AmII-like</td>
<td>1524</td>
<td>(\sigma C_\gamma N_2 HD (49), \nu C_\beta N_2 (30), -\nu C_\gamma C_\delta (6), -\rho N_2 HD (6), -\rho C_\delta O_\alpha (5))</td>
</tr>
<tr>
<td>(\sigma CH_2)</td>
<td>(1494)</td>
<td>(\sigma C_\gamma H_2 (88))</td>
</tr>
<tr>
<td>(\delta_\alpha NH_3 + \sigma CH_2)</td>
<td>(1465)</td>
<td>(-\delta_\alpha NH_3 (36), \sigma C_\gamma H_2 (33), \omega C_\gamma H_2 (5))</td>
</tr>
<tr>
<td>(\sigma CH_2)</td>
<td>(1460)</td>
<td>(-\sigma C_\gamma H_2 (47), -\delta_\alpha NH_3 (34), \nu C'O (3))</td>
</tr>
<tr>
<td>(\omega CH_2)</td>
<td>(1435)</td>
<td>(-\omega C_\beta H_2 (21), \nu C'O (13), -\rho C_\gamma H_2 (11), \omega C_\gamma H_2 (10), -\sigma C_\gamma H_2 (7), -\nu C_\gamma C_\delta (6), -\nu C_\beta C_\gamma (5), -\delta_\alpha' NH_3 (5))</td>
</tr>
<tr>
<td>(\nu s COO)</td>
<td>(1407)</td>
<td>(\nu C'O (22), -\omega C_\gamma H_2 (20), -\nu C_\gamma C_\delta (9), \nu C_\gamma C_\delta (7), -\beta C'O_\alpha (6), \nu C'O (6), \sigma N_2 HD (5))</td>
</tr>
<tr>
<td>(\rho CH)</td>
<td>(1384)</td>
<td>(\rho C_\gamma H_2 (40), -\rho C_\gamma H_2 (11), -\tau C_\gamma H_2 (6), \nu C_\gamma C_\delta (6), -\nu C'O (5))</td>
</tr>
<tr>
<td>(\tau CH_2)</td>
<td>(1356)</td>
<td>(\tau C_\gamma H_2 (35), \rho C_\gamma H_2 (16), -\tau C_\gamma H_2 (15), -\rho C_\gamma H_2 (6))</td>
</tr>
<tr>
<td>AmII(^S)-like</td>
<td>(1329)</td>
<td>(\omega C_\gamma H_2 (39), -\sigma C_\gamma H_2 (9), -\sigma N_2 HD (9), \nu C_\beta N_2 (8), -\nu C_\gamma C_\delta (5))</td>
</tr>
<tr>
<td>(\tau CH_2)</td>
<td>(1307)</td>
<td>(-\tau C_\gamma H_2 (34), -\rho C_\gamma H_2 (22), -\rho C_\gamma H_2 (8), -\rho C_\gamma H_2 (8))</td>
</tr>
<tr>
<td>AmII(^S)-like</td>
<td>(1247)</td>
<td>(-\omega C_\gamma H_2 (25), \nu C_\beta N_2 (20), -\sigma N_2 HD (9), -\rho NH_3 (7), -\rho C_\gamma O_\delta (6), \nu C_\gamma C_\delta (6), -\omega C_\gamma H_2 (5), \tau C_\gamma H_2 (5))</td>
</tr>
<tr>
<td>(\tau CH_2)</td>
<td>(1207)</td>
<td>(-\tau C_\gamma H_2 (18), -\tau C_\gamma H_2 (15), -\tau C_\gamma H_2 (15), -\tau C_\gamma H_2 (15), -\rho C_\gamma H_2 (11), -\nu C_\gamma C_\delta (9), -\rho C_\gamma H_2 (6))</td>
</tr>
<tr>
<td>(\rho CH + \tau CH_2)</td>
<td>(1153)</td>
<td>(\rho C_\gamma H_2 (19), -\tau C_\gamma H_2 (16), \rho C_\gamma H_2 (14), -\tau C_\gamma H_2 (11), \nu C_\gamma C_\delta (9), -\rho C_\gamma H_2 (6))</td>
</tr>
<tr>
<td>(\nu CN)</td>
<td>(1116)</td>
<td>(\nu C_\delta N_2 (39), -\nu C_\delta N_2 (10), -\rho NH_3 (10), \nu C_\gamma C_\delta (6), -\delta N_\gamma C_\gamma O_\delta (6), \rho C_\gamma H_2 (5), \sigma C_\gamma C_\delta C_\gamma (5))</td>
</tr>
<tr>
<td>(\rho NH_3)</td>
<td>(1104)</td>
<td>(-\rho NH_3 (27), -\nu C_\gamma C_\delta (15), -\nu C_\gamma H_2 (8), -\nu C_\gamma H_2 (7), -\delta N_\gamma C_\gamma O_\delta (6))</td>
</tr>
<tr>
<td>AmII(^S)-like</td>
<td>(1055)</td>
<td>(-\rho N_2 HD (27), -\nu C_\gamma C_\delta (14), -\nu C_\beta N_2 (9), -\nu C_\gamma O_\delta (8), -\nu C_\gamma O_\delta (5), -\rho C_\gamma O_\delta (5))</td>
</tr>
<tr>
<td>(\nu CN)</td>
<td>(1032)</td>
<td>(\nu C_\gamma N_2 (31), -\rho N_2 HD (10), -\rho NH_3 (7), -\nu C_\gamma C_\delta (7), -\rho C_\gamma H_2 (6))</td>
</tr>
<tr>
<td>(\rho NH_3)</td>
<td>(1003)</td>
<td>(\rho NH_3 (25), -\nu C_\gamma N_2 (19), -\rho C_\gamma H_2 (14), -\rho C_\gamma H_2 (14), -\nu C_\gamma C_\delta (7), -\tau C_\gamma H_2 (5))</td>
</tr>
<tr>
<td>(\rho NH_3)</td>
<td>(973)</td>
<td>(-\rho NH_3 (37), -\nu C_\gamma C_\delta (25), -\nu C_\gamma C_\delta (9), -\sigma C_\gamma C_\delta C_\gamma (7), -\nu C_\gamma N_2 (6))</td>
</tr>
</tbody>
</table>

\(^a\)C\(_\delta\)-N\(_2\) stretching and N\(_2\)HD scissoring components in PED are in bold.

\(^b\)ν: stretch; δ: asymmetric deformation; δ\(_s\): symmetric deformation; δ\(_s\): deformation; σ: scissoring; ρ: rocking; ω: wagging; β: in-plane bending; τ: twisting.
Table D3: UVRR Band Frequencies (cm$^{-1}$) and Assignments of $trans$- and $cis$-Glutamine-$N\varepsilon_2$HD. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$cis$-$N\varepsilon_2$HD</td>
<td>1628</td>
<td>1715</td>
<td>$\nu_{as}$COO</td>
<td>1658</td>
<td>1742</td>
<td>AmI-like</td>
</tr>
<tr>
<td></td>
<td>1666</td>
<td>$\delta_{as}$NH$_3$</td>
<td>1666</td>
<td>$\delta_{as}$NH$_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1621</td>
<td>$\delta_{as}$NH$_3$</td>
<td>1621</td>
<td>$\delta_{as}$NH$_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1494</td>
<td>$\sigma$CH$_2$</td>
<td>1494</td>
<td>$\sigma$CH$_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1476</td>
<td>1479</td>
<td>AmII-like</td>
<td>1465</td>
<td>$\delta_{s}$NH$_3$+$\sigma$CH$_2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1449</td>
<td>1460</td>
<td>$\sigma$CH$_2$</td>
<td>1449</td>
<td>1460</td>
<td>$\sigma$CH$_2$</td>
</tr>
<tr>
<td></td>
<td>1420</td>
<td>1437</td>
<td>$\omega$CH$_2$</td>
<td>1420</td>
<td>1435</td>
<td>$\omega$CH$_2$</td>
</tr>
<tr>
<td></td>
<td>1420</td>
<td>1414</td>
<td>$\nu_{as}$COO</td>
<td>1420</td>
<td>1407</td>
<td>$\nu_{as}$COO</td>
</tr>
<tr>
<td></td>
<td>1398</td>
<td>1387</td>
<td>$\rho$CH</td>
<td>1398</td>
<td>1384</td>
<td>$\rho$CH</td>
</tr>
<tr>
<td></td>
<td>1360</td>
<td>1357</td>
<td>$\rho$CH</td>
<td>1360</td>
<td>1356</td>
<td>$\tau$CH$_2$</td>
</tr>
<tr>
<td></td>
<td>1335</td>
<td>1347</td>
<td>$\omega$CH$_2$+$\tau$CH$_2$</td>
<td>1308</td>
<td>1329</td>
<td>AmIII$^S$-like</td>
</tr>
<tr>
<td></td>
<td>1309</td>
<td>$\tau$CH$_2$</td>
<td>1307</td>
<td>$\tau$CH$_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1278</td>
<td>1272</td>
<td>$\omega$CH$_2$</td>
<td>1247</td>
<td>1247</td>
<td>AmIII$^S$-like</td>
</tr>
<tr>
<td></td>
<td>1214</td>
<td>$\tau$CH$_2$+$\nu$CC</td>
<td>1207</td>
<td>$\tau$CH$_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1153</td>
<td>$\rho$CH+$\tau$CH$_2$</td>
<td>1153</td>
<td>$\rho$CH+$\tau$CH$_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1117</td>
<td>$\nu$CC</td>
<td>1116</td>
<td>$\nu$CC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1105</td>
<td>$\rho$NH$_3$</td>
<td>1104</td>
<td>$\rho$NH$_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1040</td>
<td>$\nu$CN</td>
<td>1055</td>
<td>AmIII$^S$-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1005</td>
<td>$\rho$NH$_3$+$\nu$CN</td>
<td>1003</td>
<td>$\rho$NH$_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>983</td>
<td>$\rho$NH$_3$</td>
<td>973</td>
<td>$\rho$NH$_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>964</td>
<td>953</td>
<td>AmIII$^S$-like</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$$\nu$: stretch; $\delta_{as}$: asymmetric deformation; $\delta$: symmetric deformation; $\delta$: deformation; $\sigma$: scissoring; $\rho$: rocking; $\omega$: wagging; $\beta$: in-plane bending; $\tau$: twisting.
Table D4: Average Number of Hydrogen Bonds for Antiparallel and Parallel β-sheet Fibril Models. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.

<table>
<thead>
<tr>
<th>H-bonding type</th>
<th>Parallel β-strand fibril</th>
<th>Antiparallel β-strand fibril</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide-peptide</td>
<td>106.7</td>
<td>118.9</td>
</tr>
<tr>
<td>Peptide-solvent</td>
<td>281.2</td>
<td>256.5</td>
</tr>
<tr>
<td>Peptide backbone-side chain</td>
<td>36.0</td>
<td>16.8</td>
</tr>
<tr>
<td>Peptide backbone-peptide backbone</td>
<td>39.2</td>
<td>66.8</td>
</tr>
<tr>
<td>Side chain-side chain</td>
<td>31.4</td>
<td>35.3</td>
</tr>
</tbody>
</table>

Figure D1: Atomic labeling scheme for glutamine used in DFT calculations. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
Figure D2: Plot of the evolution of backbone RMSD metric for three fibril models with respect to their initial structure. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
Figure D3: UVRR 197 nm – 204 nm difference spectra of fibril films prepared from (a) NDQ10 and (b) DQ10. The asterisk indicates an artifact of subtracting the intense $O_2$ stretching band in the difference spectrum. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
Figure D4: Potential energy distributions for the Figure 5.1 simulated model a antiparallel and model b parallel $\beta$-sheet fibril systems shown as red and blue histograms, respectively. These distributions represent the probability that a particular potential energy was sampled during the simulation for the antiparallel $\beta$-sheet fibril system (red) and the parallel $\beta$-sheet fibril system (blue). The significant overlap indicates that potential energies are sufficiently converged for use of the BAR method. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
Figure D5: Histograms of the number of hydrogen bonds for models a and b in Figure 1.1. (a) peptide-peptide hydrogen bonds, (b) peptide-solvent hydrogen bonds, (c) peptide backbone-side chain hydrogen bonds, (d) peptide backbone-peptide backbone, and (e) side chain-side chain hydrogen bonds. Data was taken throughout the 200 ns trajectory for model a (blue) and model b (red). The overlap of the two distributions is shown in purple. Hydrogen bonds were defined as having a distance between heavy atoms of less than 3.0 Å and an angle cutoff of 30°. Adapted with permission from [196]. Copyright © (2016), American Chemical Society.
REFERENCES


[10] Joan S. Steffan, Laszlo Bodai, Judit Pallos, Marnix Poelman, Alexander McCam- 
bell, Barbara L. Apostol, Alexsey Kazantsev, Emily Schmidt, Ya-Zhen Zhu, Marilee 
Greenwald, et al. Histone Deacetylase Inhibitors Arrest Polyglutamine-Dependent 

Wellington, Michael R. Hayden, Russell L. Margolis, Christopher A. Ross, Jean Daus-
set, Robert J. Ferrante, et al. The Gln-Ala Repeat Transcriptional Activator CA150 
Interacts with Huntingtin: Neuropathologic and Genetic Evidence for a Role in Hunt-
ington’s Disease Pathogenesis. Proceedings of the National Academy of Sciences, 

[12] Frederick C. Nucifora, Masayuki Sasaki, Matthew F. Peters, Hui Huang, Jillian K. 
Cooper, Mitsunori Yamada, Hitoshi Takahashi, Shoji Tsuji, Juan Troncoso, Valina L. 
Dawson, et al. Interference by Huntingtin and Atrophin-1 with CBP-Mediated Tran-

Polyglutamine Length-Dependent Interaction of Hsp40 and Hsp70 Family Chaperones 
with Truncated N-terminal Huntingtin: their Role in Suppression of Aggregation and 

[14] Yaohui Chai, Stacia L. Koppenhafer, Sarah J. Shoesmith, Matthew K. Perez, and 
Henry L. Paulson. Evidence for Proteasome Involvement in Polyglutamine Disease: 
Localization to Nuclear Inclusions in SCA3/MJD and Suppression of Polyglutamine 

orders: Aggregation Revisited. Human Molecular Genetics, 12(suppl 2):R173–R186, 
2003.

Hevelone, and Steven M. Hersch. Complexity and Heterogeneity: What Drives the 
Ever-changing Brain in Huntington’s disease? Annals of the New York Academy of 

Polyglutamine Peptides Delivered to Nuclei are Toxic to Mammalian Cells. Human 

[18] Toshiya Sato, Masami Miura, Mitsunori Yamada, Takayuki Yoshida, Jonathan D. 
Wood, Ikuru Yazawa, Masao Masuda, Takeo Suzuki, Ryong-Moon Shin, Hau-Jie Yau, 
et al. Severe Neurological Phenotypes of Q129 DRPLA Transgenic Mice Serendipi-
tously Created by En Masse Expansion of CAG Repeats in Q76 DRPLA Mice. Human 


[28] Kei Watase, Edwin J. Weeber, Bisong Xu, Barbara Antalffy, Lisa Yuva-Paylor, Kouichi Hashimoto, Masanobu Kano, Richard Atkinson, Yaling Sun, Dawna L. Armstrong,


