Glutamine and Asparagine Side Chain Hyperconjugation-Induced Structurally Sensitive Vibrations

David Punhaole, Zhenmin Hong, Ryan S. Jakubek, Elizabeth M. Dahlburg, Steven Geib, and Sanford A. Asher*

Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Chevron Science Center, Pittsburgh, Pennsylvania 15260, United States

Supporting Information

ABSTRACT: We identified vibrational spectral marker bands that sensitively report on the side chain structures of glutamine (Gln) and asparagine (Asn). Density functional theory (DFT) calculations indicate that the Amide III* (AmIII*) vibrations of Gln and Asn depend cosinusoidally on their side chain OCCC dihedral angles (the $\chi_1$ and $\chi_2$ angles of Gln and Asn, respectively). We use UV resonance Raman (UVR) and visible Raman spectroscopy to experimentally correlate the AmIII* Raman band frequency to the primary amide OCCC dihedral angle. The AmIII* 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 AmIII* 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 AmIII* frequency. The $C_{\beta}$-$C_{\gamma}$ ($C_{\alpha}$-$C_{\beta}$) bond length dependence on the $\chi_1$ ($\chi_2$) dihedral angle results from hyperconjugation between the $C_{\beta}$=O, ($C_{\gamma}$=O) irrespective and $C_{\gamma}$-C$_{\alpha}$ ($C_{\alpha}$-$C_{\beta}$) $\sigma$ orbitals. Using a Protein Data Bank library, we show that the $\chi_1$ and $\chi_2$ dihedral angles of Gln and Asn depend on the peptide backbone Ramachandran angles. We demonstrate that the inhomogeneously broadened AmIII* band line shapes can be used to calculate the $\chi_1$ 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.

INTRODUCTION

Amyloid-like fibril protein aggregates and prion proteins often contain stretches of glutamine (Gln) and asparagine (Asn) residues. For example, polyglutamine (polyGln)-rich fibrils are the pathological hallmarks of several “CAG” codon repeat diseases.1–9 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.10

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 studies11–13 suggest that there are at least two different populations of Gln side chain conformers in polyGln fibrils. Sharma et al.14 claim on the basis of low-resolution X-ray fiber and powder diffraction data that the side chains in polyGln 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 studies15,16 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.

UV resonance Raman (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.17–22 Deep UV excitation (~200 nm) selectively resonance enhances secondary and primary amide vibrations.23–30 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.31–34 For example, Asher and co-workers35–37 quantitatively correlated the Amide III, $\chi_1$, and $\chi_2$ angles with the vibrational spectral marker bands that sensitively report on the side chain structures of glutamine (Gln) and asparagine (Asn). Density functional theory (DFT) calculations indicate that the Amide III* (AmIII*) vibrations of Gln and Asn depend cosinusoidally on their side chain OCCC dihedral angles (the $\chi_1$ and $\chi_2$ angles of Gln and Asn, respectively). We use UV resonance Raman (UVR) and visible Raman spectroscopy to experimentally correlate the AmIII* Raman band frequency to the primary amide OCCC dihedral angle. The AmIII* 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 AmIII* 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 AmIII* frequency. The $C_{\beta}$-$C_{\gamma}$ ($C_{\alpha}$-$C_{\beta}$) bond length dependence on the $\chi_1$ ($\chi_2$) dihedral angle results from hyperconjugation between the $C_{\beta}$=O, ($C_{\gamma}$=O) $\pi$ and $C_{\gamma}$-C$_{\alpha}$ ($C_{\alpha}$-$C_{\beta}$) $\sigma$ orbitals. Using a Protein Data Bank library, we show that the $\chi_1$ and $\chi_2$ dihedral angles of Gln and Asn depend on the peptide backbone Ramachandran angles. We demonstrate that the inhomogeneously broadened AmIII* band line shapes can be used to calculate the $\chi_1$ 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.

Received: August 6, 2015
Revised: September 17, 2015
Published: September 21, 2015
We demonstrate that the C=O bond length dependence on the stretching motions. We fit Ramachandran angles. Applying this new insight and the β and Asn residues that adopt PPII, the structural sensitivity of the Amide IIIP (AmIIIP) vibration derives from coupling between the dihedral angle. They determined that the structural sensitivity of the Amide IIIP vibration contains significant contributions of Cβ−Cγ (Cα−Cβ) stretching, N3H2 (N3H2+) rocking, and Cα−Nε (Cβ−Nε) stretching motions. We find that the structural sensitivity of the AmIII mode originates mainly from the Cβ−Cγ (Cα−Cβ) bond length dependence on the X1 (X2) dihedral angle. We demonstrate that the Cβ−Cγ (Cα−Cβ) bond length correlation on the X2 (X1) dihedral angle derives from hyperconjugation between the Cβ−Cγ (Cα−Cβ) σ orbital and the Cε═Oσ (C═Oπ) π* orbital.

We compare our results with the Gln and Asn entries of the potential energy distribution (PED) of this vibration in Gln (Asn) contains significant contributions of Cβ−Cγ (Cα−Cβ) stretching, N3H2 (N3H2+) rocking, and Cα−Nε (Cβ−Nε) stretching motions. We observe distinct X1 and X2 dihedral angle preferences for Gln and Asn residues that adopt PPII, β-sheet and α-helix Ramachandran angles. Applying this new insight and the dependence of the X1 dihedral angle on the AmIII vibrational frequency, we determine the X3 angle spatial distribution of Gln3 and Asp2-Gln10-Lys2 peptides in aqueous solution. We find that Gln3 and Asp2-Gln10-Lys2 favor X1 dihedral angles similar to those of Gln in solution. This result is consistent with Gln3 and Asp2-Gln10-Lys2 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.

### EXPERIMENTAL DETAILS

#### Materials

- L-Glutamine (L-Gln, ≥99% purity), L-glutamine t-butyler ester hydrochloride (Gln-TBE, ≥98% purity), and glycol-L-glutamine (Gly-Gln, ≥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-ser t-lys-asparagine (Ser-Asn, ≥99% purity) was purchased from Bachem. Optima-grade H2O was purchased from Fisher Scientific, and D2O (99.9% atom D purity) was purchased from Cambridge Isotope Laboratories, Inc. Gln3 was purchased from Pierce Biotechnology at 95% purity.

- Sample Preparation. Gly-Gln and Ser-Asn were obtained as crystalline powders and used without further purification or recrystallization. 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 recrystallization in D2O. Samples of Gln3 were prepared at 0.5 mg·mL−1 in HPLC-grade water containing 0.05 M 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.

#### X-ray Diffraction

X-ray diffraction of crystals was performed using a Bruker X8 Prospector Ultra equipped with a copper microfocus tube (λ = 1.54178 Å). The crystals were mounted and placed in a cold stream of N2 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. The Supporting Information discusses, in detail, the methods used to determine the unit cells and crystal structures of the compounds examined.

#### 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 ∼2 cm−1. The 918 and 1376 cm−1 bands of acetone35 were used to calibrate the spectral frequencies.

#### UV Resonance Raman Spectroscopy

UV resonance Raman (UVRR) spectra of crystals were collected using CW 229 nm light generated by an Innova 300C FreD frequency doubled Ar+ laser.46 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 ∼1.5 to 2 mW. The 801, 1028, 2852, 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 ∼204 nm excitation. The UV light was generated by Raman shifting the third harmonic of a Nd:YAG laser (Coherent, Inc.) with H2 gas (∼30 psi) and selecting the fifth anti-Stokes line. A thermostated (20°) 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 configuration77 and detected with a Spec-10 CCD camera.

### COMPUTATIONAL DETAILS

#### Density Functional Theory (DFT) Calculations

The DFT calculations were carried out using the GAUSSIAN 09 package.49 The geometry optimizations and frequency calculations were performed using the M06-2X functional 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 (see Supporting Information). Figure 1 shows the DFT calculated minimum energy structure of L-Gln and the atomic
labeled scheme. To study the conformational dependence of the Raman bands, we fixed the $\chi_3$ dihedral angle of L-Gln, reoptimized the geometry, and calculated the harmonic vibrational frequencies for a series of conformers with $\chi_3$ angles of $-16^\circ$, $0^\circ$, $4^\circ$, $\pm 30^\circ$, $\pm 60^\circ$, $\pm 120^\circ$, $\pm 150^\circ$, and $\pm 180^\circ$.

## RESULTS AND DISCUSSION

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

Figure 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 co-workers for both solid-state crystalline samples and in the solution state. Ramirez and co-workers 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 co-workers. These assignments build off of our previous, detailed assignment of propanamide, 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 Tables 1 and 2, respectively.

### Table 1. UVRR Frequencies (cm$^{-1}$) and Assignments of L-Gln in H$_2$O

<table>
<thead>
<tr>
<th>expt</th>
<th>calc</th>
<th>potential energy distribution (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1679</td>
<td>1745</td>
<td>$\nu$C$_3$O$_5$ (75), $\nu$C$_3$N$_6$ (8), $\nu$C$_3$C$_3$ (7)</td>
</tr>
<tr>
<td>1652</td>
<td>1666</td>
<td>$\delta$$_s$NH$_4$ (48), $\delta$$_s$NH$_4$ (44), $\nu$NH$_3$ (5)</td>
</tr>
<tr>
<td>1620</td>
<td>1623</td>
<td>$-\sigma$N$_2$H$_2$ (86), $-\sigma$C$_3$N$_6$ (10)</td>
</tr>
<tr>
<td>1621</td>
<td>1658</td>
<td>$\delta$$_s$NH$_4$ (40), $\delta$$_s$NH$_4$ (48), $\delta$$_s$NH$_4$ (10)</td>
</tr>
<tr>
<td>1585</td>
<td>1715</td>
<td>$\nu$CO (33), $\nu$C$_3$N$_6$ (33), $\nu$C$_3$C$_3$ (7)</td>
</tr>
<tr>
<td>1464</td>
<td>1494</td>
<td>$\nu$C$_3$H$_2$ (88)</td>
</tr>
<tr>
<td>1466</td>
<td>1466</td>
<td>$\delta$$_s$NH$_4$ (33), $\sigma$C$_3$H$_2$ (30), $-\sigma$C$_3$H$_2$ (7), $\sigma$C$_3$C$_3$ (5), $\nu$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1447</td>
<td>1460</td>
<td>$-\sigma$C$_3$H$_2$ (44), $-\delta$$_s$NH$_4$ (36)</td>
</tr>
<tr>
<td>1427</td>
<td>1445</td>
<td>$-\sigma$C$_3$H$_2$ (17), $\nu$C$_3$C$_3$ (13), $-\sigma$C$_3$H$_2$ (13), $-\sigma$C$_3$N$_6$ (5), $-\sigma$C$_3$C$_3$ (5), $\nu$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1411</td>
<td>1419</td>
<td>$\nu$CO (26), $\nu$C$_3$H$_2$ (12), $\nu$C$_3$C$_3$ (11), $-\sigma$C$_3$H$_2$ (8), $\sigma$C$_3$H$_2$ (7), $\nu$C$_3$C$_3$ (7), $\nu$C$_3$C$_3$ (7)</td>
</tr>
<tr>
<td>1365</td>
<td>1388</td>
<td>$\nu$C$_3$H$_2$ (25), $\nu$C$_3$H$_2$ (13), $\nu$CO (10), $-\nu$C$_3$H$_2$ (9), $\nu$C$_3$C$_3$ (8), $-\nu$C$_3$H$_2$ (6), $-\nu$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1351</td>
<td>1358</td>
<td>$-\nu$C$_3$H$_2$ (30), $-\sigma$C$_3$H$_2$ (16), $-\sigma$C$_3$H$_2$ (15), $\nu$C$_3$C$_3$ (9), $\nu$C$_3$C$_3$ (6)</td>
</tr>
<tr>
<td>1328</td>
<td>1348</td>
<td>$-\nu$C$_3$H$_2$ (26), $-\nu$C$_3$H$_2$ (18), $-\nu$C$_3$H$_2$ (13), $-\nu$C$_3$N$_6$ (11)</td>
</tr>
<tr>
<td>1293</td>
<td>1309</td>
<td>$-\nu$C$_3$H$_2$ (35), $-\nu$C$_3$H$_2$ (26), $\nu$C$_3$H$_2$ (8), $\nu$C$_3$H$_2$ (6)</td>
</tr>
<tr>
<td>1264</td>
<td>1272</td>
<td>$-\nu$C$_3$H$_2$ (43), $-\nu$C$_3$H$_2$ (30), $\nu$C$_3$N$_6$ (13)</td>
</tr>
<tr>
<td>1206</td>
<td>1215</td>
<td>$-\nu$C$_3$H$_2$ (21), $-\nu$C$_3$C$_3$ (18), $-\sigma$C$_3$H$_2$ (16), $-\nu$C$_3$H$_2$ (13), $\sigma$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1158</td>
<td>1153</td>
<td>$-\nu$C$_3$H$_2$ (20), $\nu$C$_3$H$_2$ (17), $-\nu$C$_3$H$_2$ (13), $\nu$C$_3$H$_2$ (12), $-\nu$C$_3$C$_3$ (8), $-\sigma$C$_3$H$_2$ (6)</td>
</tr>
<tr>
<td>1130</td>
<td>1122</td>
<td>$\nu$C$_3$C$_3$ (34), $\sigma$C$_3$N$_6$ (17), $-\sigma$C$_3$C$_3$ (7), $\nu$C$_3$N$_6$ (6), $-\nu$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1109</td>
<td>1109</td>
<td>$\nu$NH$_3$ (27), $-\nu$C$_3$C$_3$ (10), $-\nu$C$_3$C$_3$ (10), $-\nu$C$_3$C$_3$ (10), $\sigma$C$_3$C$_3$ (10), $\sigma$C$_3$N$_6$ (10)</td>
</tr>
<tr>
<td>1107</td>
<td>1097</td>
<td>$\nu$C$_3$C$_3$ (26), $-\nu$C$_3$C$_3$ (26), $-\sigma$C$_3$N$_6$ (13), $\nu$C$_3$N$_6$ (8)</td>
</tr>
<tr>
<td>1038</td>
<td>1038</td>
<td>$\nu$C$_3$C$_3$ (36), $-\nu$C$_3$C$_3$ (9), $\nu$C$_3$C$_3$ (8), $-\nu$C$_3$N$_6$ (6), $\nu$C$_3$C$_3$ (5), $\nu$C$_3$C$_3$ (5)</td>
</tr>
<tr>
<td>1078</td>
<td>1003</td>
<td>$\nu$NH$_3$ (25), $-\sigma$C$_3$N$_6$ (19), $\nu$C$_3$H$_2$ (14), $\nu$C$_3$H$_2$ (14), $\nu$C$_3$C$_3$ (7), $-\nu$C$_3$H$_2$ (5)</td>
</tr>
<tr>
<td>1006</td>
<td>974</td>
<td>$-\nu$C$_3$H$_2$ (38), $\nu$C$_3$C$_3$ (25), $-\nu$C$_3$C$_3$ (8), $-\sigma$C$_3$C$_3$ (7), $\nu$C$_3$N$_6$ (6)</td>
</tr>
</tbody>
</table>

*ν, stretch; δ$_s$, asymmetric deformation; δ$_s$, symmetric deformation; δ, deformation; σ, scissoring; ρ, rocking; ω, wagging; β, in-plane bending; γ, twisting.

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 CO–N$_2$ stretching because the electronic excited state is expanded along this coordinate.
Table 2. UVRR Frequencies (cm⁻¹) and Assignments of t-Gln in D₂O

<table>
<thead>
<tr>
<th>exp</th>
<th>calc</th>
<th>potential energy distribution (≥5% contribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1650</td>
<td>1739</td>
<td>ν(C=O), (78), ν(N-C), (67), ν(N=O) (7)</td>
</tr>
<tr>
<td>1637</td>
<td>1708</td>
<td>μ(C=O) (55), ν(C=O) (34), ν(=C) (7)</td>
</tr>
<tr>
<td>1645</td>
<td>1494</td>
<td>μ(C=O) (88)</td>
</tr>
<tr>
<td>1645</td>
<td>1453</td>
<td>ν(C=O), (48), 1604-1507 cm⁻¹, ν(C=O), (8)</td>
</tr>
<tr>
<td>1644</td>
<td>1454</td>
<td>ν(C=O), (38), ν(C=C), (19), ν(C=O), (9)</td>
</tr>
<tr>
<td>1412</td>
<td>1424</td>
<td>ν(C=O), (29), ν(C=O), (12), ν(C=O) (10), 1604-1507 cm⁻¹, ν(C=O) (8)</td>
</tr>
<tr>
<td>1368</td>
<td>1393</td>
<td>ρ(C=O), (21), ρ(C=O), (15), ν(C=O), (11), ν(C=O) (8), 1604-1507 cm⁻¹, ρ(C=O) (6), 1604-1507 cm⁻¹, ν(C=O) (5)</td>
</tr>
<tr>
<td>1344</td>
<td>1362</td>
<td>ν(C=O), (37), ν(C=O), (13), ν(C=O), (12), ν(C=O) (11), ν(C=O) (8), 1604-1507 cm⁻¹, ρ(C=O) (7), 1604-1507 cm⁻¹, ν(C=O) (5)</td>
</tr>
<tr>
<td>1349</td>
<td>1349</td>
<td>ν(C=O), (44), ν(C=O), (20), ρ(C=O), (15), ρ(C=O) (6)</td>
</tr>
<tr>
<td>1304</td>
<td>1304</td>
<td>ν(C=O), (32), ν(C=O), (29), ν(C=O) (8), ν(C=O) (7)</td>
</tr>
<tr>
<td>1282</td>
<td>1282</td>
<td>ν(C=O), (37), ν(C=O), (33), ν(C=O) (8), ν(C=O) (7), ν(C=O) (6)</td>
</tr>
<tr>
<td>1202</td>
<td>1202</td>
<td>δ(N-D), (23), δ(N-D), (20), ν(C=O), (12), ν(C=O) (8)</td>
</tr>
<tr>
<td>1198</td>
<td>1198</td>
<td>δ(N-D), (27), ν(C=O), (12), ν(C=O) (11), ν(C=O) (9), ν(C=O) (8)</td>
</tr>
<tr>
<td>1186</td>
<td>1186</td>
<td>δ(N-D), (32), δ(N-D), (27), ν(C=O), (8), δ(N-D), (6)</td>
</tr>
<tr>
<td>1161</td>
<td>1161</td>
<td>s(N-D), (56), ρ(C=O), (11), ν(C=O), (9), ν(C=O) (8)</td>
</tr>
<tr>
<td>1111</td>
<td>1111</td>
<td>ν(C=O), (21), δ(N-D), (16), δ(N-D), (14), ν(C=O), (13), ν(C=O), (9)</td>
</tr>
<tr>
<td>1119</td>
<td>1119</td>
<td>δ(N-D), (17), δ(N-D), (10), δ(N-D), (9), ν(C=O), (8), δ(N-D), (8), δ(N-D), (7), δ(N-D), (5)</td>
</tr>
<tr>
<td>1105</td>
<td>1105</td>
<td>ν(C=O), (44), δ(N-D), (15), δ(N-D), (8), δ(N-D), (6), ν(C=O), (5)</td>
</tr>
<tr>
<td>1058</td>
<td>1058</td>
<td>ρ(C=O), (21), ρ(C=O), (16), ν(C=O), (12), ν(C=O), (9), ν(C=O), (8), ν(C=O), (7), ν(C=O), (6), ν(C=O), (5)</td>
</tr>
<tr>
<td>992</td>
<td>992</td>
<td>ν(C=O), (29), δ(N-D), (15), δ(N-D), (10), δ(N-D), (10), δ(N-D), (9), δ(N-D), (5)</td>
</tr>
<tr>
<td>960</td>
<td>960</td>
<td>ρ(N-D), (22), ν(C=O), (14), ν(C=O), (14), δ(N-D), (9), 1604-1507 cm⁻¹, δ(N-D), (7), δ(N-D), (5), δ(N-D), (3), ν(C=O), (6), ν(C=O), (5)</td>
</tr>
<tr>
<td>929</td>
<td>929</td>
<td>ν(C=O), (17), ρ(N-D), (16), δ(N-D), (16), ν(C=O), (8), ν(C=O), (7), δ(N-D), (6), δ(N-D), (5)</td>
</tr>
</tbody>
</table>

The spectral region between 1600 and 1700 cm⁻¹ is dominated by two primary amide vibrations, the Amide I° (Amil°) and Amide II° (Amil°) bands. The superscript ° denotes the primary amide to distinguish these vibrations from the widely known vibrations of secondary amides found in proteins. The Amil° band is located at ~1680 cm⁻¹ and derives mainly from C=O stretching. In D₂O, the Amil° band (called the Amil°) downshifts to ~1650 cm⁻¹. The Amil° band at ~1620 cm⁻¹ derives from a vibration whose PED contains mostly NH₂ rocking (~86%) and C=O=O stretching (~10%). Upon N-deuteration, the C=O=O stretching and ND₂ scissoring motions decouple. This causes the Amil° band to disappear, and a new band, which derives from ND₂ scissoring, appears at ~1160 cm⁻¹.

The most intense features of the Gln spectra in Figure 2 occur in the region between 1400 and 1500 cm⁻¹. Most of the bands found in this region derive from CH₂ scissoring or wagging modes. However, we assign the most intense band, located at ~1430 cm⁻¹, to a vibration that contains significant contributions of CH₂ wagging, C=O=O stretching, CH₂ scissoring, and C=O=O stretching in its PED. This assignment is based on our previous work with propanamide, which shows a similar intense band at ~1430 cm⁻¹.

The region between 1200 and 1400 cm⁻¹ contains bands that derive mostly from CH wagging, CH₂ wagging, and CH₂ twisting modes. We assign the ~1365 and ~1350 cm⁻¹ bands in the Figure 2a spectrum to CH₂ wagging modes. We assign the strong bands located at ~1330 cm⁻¹, ~1290 cm⁻¹, and the very weak ~1205 cm⁻¹ bands to CH₂ twisting modes. The ~1265 cm⁻¹ feature is assigned to a CH₂ wagging vibration. Only two bands, at ~1370 and ~1345 cm⁻¹, appear in D₂O. We assign the ~1370 cm⁻¹ band to a CH₂ wagging mode and the ~1345 cm⁻¹ band to a C=H rocking mode. We conclude that these vibrations appear strongly in the UVR spectrum in Figure 2b because they contain significant C=O=O=O stretching.

The region between 1000 and 1200 cm⁻¹ contains bands that derive from vibrations with large C=O stretching, N≡H₂ rocking, or NH₃ rocking contributions. Most of the vibrations in this region are complex. We assign the ~1160 cm⁻¹ band to a coupled C=H 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 and ~1005 cm⁻¹ bands to NH₃ rocking vibrations.

Table 2. UVRR Frequencies (cm⁻¹) and Assignments of t-Gln in D₂O

The remaining two bands in the 1000–1200 cm⁻¹ region are located at ~1130 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=O=O=O stretching and NH₂ rocking. The ~1110 cm⁻¹ band is assigned to a vibration that consists of an out-of-phase combination of C=O=O=O stretching and NH₂ rocking. This vibration also contains a significant C=O=O=O stretching component (~13%), which is in phase with NH₂ rocking.

The in-phase combination of NH₂ rocking and C=O=O=O stretching of the ~1110 cm⁻¹ vibration is reminiscent of the Amill mode of secondary amides. While complex, the Amill vibrational spectrum shows significant contributions of in-phase C=O=O=O stretching and N–H in-plane bending motions of the secondary amide group. We propose to call the ~1110 cm⁻¹ mode the Amille III° (Amill°) because the eigenvector composition of this vibration is analogous to that of the canonical Amill of secondary amides. As discussed in detail below, the Amill° vibration is sensitive to the X(5) and X(2) dihedral angles of Gln and Asn.

Conformational Dependence of the Amill° Band. We performed DFT calculations on t-Gln molecules with X(5) dihedral angles fixed at different values (see Computational section for details) in order to identify spectroscopic markers that are diagnostic of the side chain X(5) and X(2) dihedral angles of Gln and Asn, respectively. We examined the frequency dependence of different primary amide vibrations and found that the Amill° vibrational frequency and normal mode depends strongly on the COCC dihedral angle.

Figure 3a shows the calculated cosinusoidal dependence of the Amill° vibrational frequency on the X(5) dihedral angle. The maximum frequency of the vibration occurs at X(5) ~ 0°, while minima occur near X(5) ~ ± 90°. The Gln Amill° band frequency dependence on the X(5) angle follows a cosinusoidal relationship:

\[ \nu(X_5) = \nu_0 + A \cos(2X_5) + B \cos(X_5 - C) \]
dependence of the AmIIIß 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 (Supporting Information Table S14), and propanamide,16 we conclude that N=H rocking, C=O=N ß stretching, and C=O=O ß stretching define the AmIIIß vibration. However, depending on the OCCC dihedral angle, other motions such as C=H twisting and C=O ß stretching can contribute to this vibration.

Therefore, we examined how the Gln Cß−Nß, Cß−N and Cß−Cß bond lengths change as a function of the χ3 dihedral angle in order to understand the origin of the conformational sensitivity of the AmIIIß vibration. Changes in these bond lengths impact the AmIIIß frequency by affecting the vibrational mode bond force constants. As seen in Figure 3b–d, all the bond lengths show a dependence on the χ3 dihedral angle. However, as seen in Figure 3b, the Cß−Cß bond length shows the largest dependence on the χ3 dihedral angle. The AmIIIß vibrational frequency has a strong correlation with the Cß−Cß bond length, as shown in Figure 3e. The AmIIIß vibrational frequency increases as the Cß−Cß bond length decreases and vice versa.

The Cß−Cß bond length dependence on the χ3 dihedral angle appears to be due to hyperconjugation between the Cß−Cß σ and the Cß=O ß ß ß orbitals (Figure 4). This interaction is strongest when these orbitals maximally overlap in the absence of significant phase cancellation due to the ß ß ß orbital antisymmetry. When hyperconjugation occurs, the ß ß ß orbital donates electron density to the ß ß ß orbital, which decreases the Cß−Cß bond order and increases its bond length. This decreases the Cß−Cß stretching force constant, which downshifts the AmIIIß frequency.

We tested this hypothesis with natural bond orbital (NBO) analysis, which allows the DFT calculated electron densities to be displayed in terms of approximate σ and ß ß ß molecular orbitals. According to our hypothesis, the Cß−Cß bond length should be largest when hyperconjugation is maximized and smallest when there is no hyperconjugation. Indeed, as seen in Figure 4b, there is significant overlap of the Cß−Cß σ and Cß=O ß ß NBO molecular orbitals at ±90°, where the Cß−Cß bond length is largest. In contrast, at χ3 ∼ 0°, where the Cß−Cß bond length is shortest, the orbital overlap cancels due to the antisymmetry of the ß ß ß orbital. Figure 5 shows the NBO charge on the Cß atom. As expected from our hyperconjugation hypothesis, the NBO Cß atom charge is less negative at χ3 ∼ ±90° compared to χ3 ∼ 0°. The NBO Cß atom charge becomes even more negative at χ3 ∼ ±150° and χ3 ∼ ±180°, even without additional hyperconjugation of the Cß−Cß σ and Cß=O ß ß ß orbitals. This result is likely an artifact because these extreme χ3 dihedral angles are associated with physically impossible high energy structures that will be subject to other electron density alterations.

Our model accounts for the AmIIIß frequency downshift as the dihedral angles approach χ3 ∼ ±90°, where hyperconjugation is strongest. This behavior is the reverse of the Bohlmann effect,54–57 where a “negative” hyperconjugation transfers electron density from a lone pair orbital to an optimally positioned C−H ß ß ß orbital. This decreases the C−H bond order and substantially downshifts the C−H stretching frequencies.

Experimental Dependence of AmIIIß Band Frequency on OCCC Dihedral Angle. We experimentally examined the dependence of the AmIIIß band frequency on the primary

Figure 3. Calculated AmIIIß frequency and bond length dependence on the χ3 dihedral angle of the Gln side chain. (a) AmIIIß frequency dependence, (b) Cß−Cß bond length, (c) Cß−Nß bond length, (d) Cß−Cß bond length, and (e) the dependence of the AmIIIß frequency on the Cß−Cß bond length.

where \( v_0 = 1084 \text{ cm}^{-1}, A = 10 \text{ cm}^{-1}, B = 3 \text{ cm}^{-1}, \) and \( C = -31° \). These parameters were calculated from a least-squares fit of eq 1 to the frequency dependence on the χ3 angle in Figure 3a.

Figure 3a shows that the AmIIIß frequency dependence on the χ3 dihedral angle is asymmetric about \( χ3 ∼ 0° \). 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 χ3 frequency dependence of eq 1. This is evident when we compare the l-Gln χ3 dependence on the AmIIIß frequency with that of butyramide (shown in Supporting Information Figure S1). In the case of butyramide, which is achiral, there is no asymmetry about 0°. As a result, the AmIIIß frequency dependence on the OCCC dihedral angle of butyramide can be satisfactorily modeled with just one cosine term (Supporting Information eq S1).

Origin of the OCCC Dihedral Angle Dependence of the AmIIIß Vibration. Understanding the conformational

DOI: 10.1021/acs.jpcl.0b07651

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 AmIIIP 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 the Supporting Information.

**Dependence of AmIIIP Band Frequency in Crystals.** Figure 6 shows the AmIIIP 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 1, obtaining the following relationship:

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

which is shown in the Figure 6 black curve. To obtain the eq 2 parameters, we fixed the A/B ratio to ~3 as found in eq 1 and performed a least-squares minimization of the experimental data. Equation 2 provides an excellent fit of the experimental data and captures the chiral asymmetry that occurs near $\chi_3 \sim \pm 90^\circ$.

**Dependence of AmIIIP Band Frequency for Fully Hydrated Primary Amides.** The AmIIIP band frequency also depends on the local hydrogen bonding and dielectric environment of the primary amide group. In water, the AmIIIP band of l-Gln is located at $\sim 1110\text{ cm}^{-1}$, as compared with $\sim 1097\text{ cm}^{-1}$ in the solid-state. On the basis of Rhys et al.’s neutron diffraction study, the solution-state equilibrium structure of l-Gln in water does not appear to differ significantly from the single known l-Gln crystal structure. From their solution-state

Figure 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, $\pi^*$ NBO molecular orbitals when the $\chi_3$ dihedral angle is (a) 0°, (b) +90°, and (c) ±180°.

Figure 5. NBO charge of C$_\beta$ in l-Gln as a function of the $\chi_3$ dihedral angle.

Figure 6. Experimental correlation of the AmIIIP frequency to the $\chi_3$ dihedral angle. The average frequency (from the 633 and 229 nm Raman spectra) of the AmIIIP band was plotted as a function of the OCCC dihedral angle: 1 = l-Gln, 2 = Gly-Gln, 3 = d-Gln, 4 = GltTBE, 5 = NAcGln, and 6 = Ser-Asn. The data were fit with eq 2 (black line, $r_{adj}^2 = 0.83$). The blue curve corresponds to eq 3. The red curve corresponds to eq 4. The yellow curve corresponds to eq 5 and is an average of the red and blue curves.
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 \((-\sim -13.54^\circ\)) from the L-Gln crystal examined in this study. Thus, by setting the AmIIIP frequency to 1110 cm\(^{-1}\), \( \chi_3 \) to \(-\sim -13.54^\circ\), and solving for \( \nu_0 \), we obtain eq 3:

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

which is shown by the Figure 6 blue curve. This equation correlates the AmIII\(^P\) band frequency to OCCC dihedral angles for situations in which the primary amide group is fully exposed to water, such as in polyproline II-like (PPII-like) structures, 2.51-helices,\(^60\) and extended \( \beta \)-strand-like peptide conformations dissolved in water.

**Dependence of AmIII\(^P\) Band Frequency for Low Dielectric Constant and Weak Hydrogen Bonding Environments.** The AmIII\(^P\) frequency downshifts \(~15\text{ cm}^{-1}\) in the low dielectric and hydrogen bonding environment of acetonitrile compared to that in water (see Supporting Information and Figure S6). This downshift derives from the different water versus acetonitrile stabilizations of the ground state \( \text{O} = \text{C} = \text{N} = \text{H}_2 \) and \( \text{O} = \text{C} = \text{N} = \text{H}_2^* \) resonance structures of the primary amide group.\(^28\) In both solvents, the \( \text{O} = \text{C} = \text{N} = \text{H}_2 \) resonance structure dominates; however, in acetonitrile, the \( \text{O} = \text{C} = \text{N} = \text{H}_2^* \) resonance structure contributes less than in water. Thus, the \( \text{C}_9 = \text{N}_8 \) bond length is larger in acetonitrile compared to water due to the lesser favorability of the \( \text{O} = \text{C} = \text{N} = \text{H}_2^* \) resonance structure. Consequently, there is a smaller \( \text{C}_9 = \text{N}_8 \) stretching force constant in acetonitrile compared to water, which results in a downshift of the AmIII\(^P\) frequency.

Equation 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 cm\(^{-1}\) downshift in \( \nu_0 \) from eq 3 to determine eq 4:

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

which is shown in red in Figure 6.

**Dependence of AmIII\(^P\) Band Frequency for Unknown Dielectric and Hydrogen Bonding Environments.** We suggest the use of eq 5, which is the average of eqs 3 and 4, for cases where the hydrogen bonding and dielectric environment of the primary amide group is unknown:

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

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. Equation 5 is shown by the yellow curve in Figure 6.

**Predicting Side Chain \( \chi_3 \) and \( \chi_2 \) Dihedral Angles in Gln and Asn as a Function of Ramachandran (\( \Phi, \Psi \)) Angles.**
Table 3. Predicted AmIII\(^\circ\) Frequencies and OCCC Dihedral Angles for Gln and Asn Residues with Different Ramachandran Angles

<table>
<thead>
<tr>
<th></th>
<th>(\Phi) (deg)</th>
<th>(\Psi) (deg)</th>
<th>(\chi_1) (deg)</th>
<th>AmIII(^\circ) freq (cm(^{-1}))</th>
<th>(\chi_2) (deg)</th>
<th>AmIII(^\circ) freq (cm(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPII</td>
<td>−65</td>
<td>145</td>
<td>−8 (−22, −32)</td>
<td>1111 (1106, 1099) 64</td>
<td>−36</td>
<td>1096</td>
</tr>
<tr>
<td>(\beta)-sheet</td>
<td>115</td>
<td>130</td>
<td>−44, 41</td>
<td>1089, 1080</td>
<td>−6, 56</td>
<td>1075, 1064</td>
</tr>
<tr>
<td>(\alpha)-helix</td>
<td>−63</td>
<td>−43</td>
<td>−34, 45</td>
<td>1098, 1076</td>
<td>−49, −19, 62</td>
<td>1085, 1107, 1058</td>
</tr>
</tbody>
</table>

*Values in parentheses were measured experimentally for Gln\(_3\) and Asp\(_2\)-Gln\(_{10}\)-Lys\(_2\).*

Shapovalov and Dunbrack\(^{44}\) recently developed a new peptide backbone dependent rotamer library, which includes the nonrotameric Gln and Asn side chain \(\chi_1\) and \(\chi_2\) dihedral angles. Their database was compiled by analyzing high resolution crystal structures from the Protein Data Bank (PDB) and consists of \(\sim\)30000 entries for Asn and \(\sim\)20000 entries for Gln. Parts a and b of Figure 7 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 (\(\Phi, \Psi\)) 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 \(\chi_1\) and \(\chi_2\) dihedral angle preferences of Gln and Asn residues that possess canonical PPII, \(\beta\)-sheet, or \(\alpha\)-helix Ramachandran angles values. On the basis of work by Richardson\(^61\) and Kaplu\(^{62}\), we assume (\(\Phi, \Psi\)) angles centered around (−65°, 145°) for canonical PPII structures, (−115°, 130°) for canonical \(\beta\)-sheets and (−63°, −43°) for canonical \(\alpha\)-helices. Parts c–h of Figure 7 depict histograms of the \(\chi_1\) 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 \(\chi_1\) and \(\chi_2\) dihedral angles clearly depend upon the peptide bond \(\Phi, \Psi\) angles. This correlation could result from a preference for particular \(\chi_1\) 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 \(\chi_2\) or \(\chi_2\) dihedral angles for the (\(\Phi, \Psi\)) angle values of their individual peptide bonds.

The \(\chi_1\) 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 7c,d. The distribution of \(\chi_1\) 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 \(\chi_1\) ~ −8° and Asn showing a peak near \(\chi_1\) ~ −36°. It should be noted that the bias due to the l-amine acid chirality gives rise to a clear preference for negative \(\chi_2\) dihedral angles for the shorter side chain Asn residues.

The \(\chi_1\) and \(\chi_2\) dihedral angle histograms of Gln and Asn with \(\beta\)-sheet (\(\Phi, \Psi\)) angles in Figure 7e,f differ dramatically from another one. The population of Gln \(\chi_2\) dihedral angles (Figure 7e) is nearly symmetric about \(\chi_2\) ~ 0°. The histogram is bimodal, with two peaks located near \(\chi_2\) angles of ~44° and ~41°. In contrast, the population of Asn residues (Figure 7f) predominately adopts negative dihedral angles and is peaked around \(\chi_2\) ~ −61°. A minor peak also occurs around \(\chi_2\) ~ 56°.

Parts g and h of Figure 7 show histograms of the \(\chi_1\) and \(\chi_2\) dihedral angles of Gln and Asn residues that adopt canonical \(\alpha\)-helical Ramachandran angles. As in Figure 7e, the Figure 7g Gln \(\chi_1\) dihedral angle population is roughly bimodal and nearly symmetric about \(\chi_1\) ~ 0°. It is peaked at \(\chi_1\) angles of ~34° and ~45°. In contrast, in Figure 7h, the population of Asn \(\chi_2\) dihedral angles is narrower and sharply peaked at \(\chi_2\) ~ −19° with two minor peaks at \(\chi_2\) ~ −49° and ~62°.

The \(\chi_1\) 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 AmIII\(^\circ\) frequencies of Gln and Asn residues that adopt canonical PPII, \(\beta\)-sheet, and \(\alpha\)-helix (\(\Phi, \Psi\)) angles (shown in Table 3). For example, using eq 3, we calculate that Gln and Asn side chains with PPII (\(\Phi, \Psi\)) angles will have a maximum probability of showing AmIII\(^\circ\) bands centered at ~1111 cm\(^{-1}\) and ~1096 cm\(^{-1}\), respectively. Similarly, we calculate that the AmIII\(^\circ\) bands of Gln residues with \(\beta\)-sheet Ramachandran angles will have the greatest probability of being located at ~1080 cm\(^{-1}\) and/or ~1089 cm\(^{-1}\). In contrast, the AmIII\(^\circ\) bands for Asn residues with \(\beta\)-sheet (\(\Phi, \Psi\)) angles will have the largest probability of being located at ~1064 cm\(^{-1}\) and/or ~1075 cm\(^{-1}\). For \(\alpha\)-helical Ramachandran angles, we calculate that the probability maxima for AmIII\(^\circ\) bands will be at ~1076 cm\(^{-1}\) and/or ~1098 cm\(^{-1}\) for Gln and ~1058, ~1085, and/or ~1107 cm\(^{-1}\) for Asn residues.

We can calculate the expected Raman spectral AmIII\(^\circ\) band shapes from the Gln \(\chi_3\) and Asn \(\chi_2\) dihedral angle histograms in Figure 7 using the AmIII\(^\circ\) Raman band frequency dependencies of eqs 2–5. 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 AmIII\(^\circ\) band is much broader than the homogeneous line width of an AmIII\(^\circ\) band expected for a single Gln and Asn residue in a typical PPII, \(\beta\)-sheet, or \(\alpha\)-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.

**Experimentally Determined Gln PPII-like Structure Peptide \(\chi_3\) Dihedral Angles. UVRR Spectra of Gln Peptides in PPII-like Structures.** We examined the UVRR spectra of two peptides, Gln\(_3\) and Asp\(_2\)-Gln\(_{10}\)-Lys\(_2\), in order to determine their solution-state \(\chi_3\) angles. Xiong et al.\(^{65}\) previously showed that Asp\(_2\)-Gln\(_{10}\)-Lys\(_2\) exists in predominately PPII-like and 2.5\(\beta\)-helix-like conformations when prepared using a "disaggregation" protocol developed by Wetzel and co-workers.\(^{63}\) In this protocol, the Asp\(_2\)-Gln\(_{10}\)-Lys\(_2\) 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 Gln$_3$ has predominately PPII-like peptide bond conformations. Figure 8a shows the peak fitted ~204 nm excitation UVRR spectrum of Gln$_3$ in the region between 1050 and 1500 cm$^{-1}$. The AmIII$^P$ region, between ~1200 and 1280 cm$^{-1}$, is most sensitive to the secondary structure of the peptide because its frequency depends on the Ramachandran $\Psi$ angle. This region is well fit by two Gaussian bands located at ~1210 and ~1260 cm$^{-1}$. Using the methodology of Mikhonin et al., we correlated the band positions to their $\Psi$ angles. We used their eq 6A to correlate the 1210 cm$^{-1}$ frequency of the AmIII$^P$ band to a $\Psi$ angle of 103$^\circ$ ± 3$^\circ$ and the 1260 cm$^{-1}$ frequency to a $\Psi$ angle of 157$^\circ$ ± 2$^\circ$. The $\Psi$ angle of ~157$^\circ$ derives from peptide bonds situated in PPII-like conformations, while the $\Psi$ angle of ~103$^\circ$ derives from peptide bond situated in $\beta$-strand-like conformations. Assuming identical Ramon cross sections for the dihedral angles of Gln, Gln$_3$, and Asp$_2$-Gln$_{10}$-Lys$_2$, we estimate that the AmIII$^P$ bandwidths of solution-state Gln, Gln$_3$, and Asp$_2$-Gln$_{10}$-Lys$_2$ are much broader than those measured in our crystals suggesting that there is a distribution of hydrogen bonding states and $\chi_3$ angles in these compounds.

Given the estimated homogeneous line width, we can roughly calculate the distribution of $\chi_3$ angles of Gln, Gln$_3$, and Asp$_2$-Gln$_{10}$-Lys$_2$ by using a methodology that is similar to that of Asher et al. To do this, we assume that the inhomogeneously broadened AmIII$^P$ bands derive from a distribution of different $\chi_3$ dihedral angles, which can be represented as the sum of $M$ Lorentzian bands:

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

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

We can apply eq 3 to correlate the $\nu_i$ AmIII$^P$ frequencies of the $M$ Lorentzian bands to their corresponding $\chi_3$ dihedral angles. As shown in Figure 6, a single AmIII$^P$ 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° and less than −90° are nearly forbidden (Figure 7). Thus, we consider only the two \( \chi_3 \) dihedral angle solutions that are found in the region between −90° and +90°, as shown in Figure 9.

![Figure 9](image)

Figure 9. \( \chi_3 \) dihedral angle histograms calculated by decomposing AmIII bands into a sum of Lorentzians for (a) Gln, (b) Gln3, and (c) Asp2-Gln10-Lys2. The Gln, Gln3, and Asp2-Gln10-Lys2 results all show one Gaussian centered at negative \( \chi_3 \) angles, \( \chi_{3,1} \) and \( \chi_{3,2} \):

\[
I(\chi_3) = A e^{-\frac{\chi_3 - \chi_{3,1}}{w}} + A e^{-\frac{\chi_3 - \chi_{3,2}}{w}}
\]

(7)

The Gln, Gln3, and Asp2-Gln10-Lys2 results all show one Gaussian centered at negative \( \chi_3 \) angles and another Gaussian centered at positive \( \chi_3 \) angles (Figure 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.\(^58\) For Gln3 and Asp2-Gln10-Lys2, we conclude that the Gaussians centered at negative \( \chi_3 \) angles correspond to the physically relevant solutions to eq 3 because they fall within the range of \( \chi_3 \) dihedral angles most commonly adopted by Gln residues that populate PPII (\( \Phi, \Psi \)) angles (Figure 7c).

Figure 10 shows the resulting \( \chi_3 \) dihedral angle distributions for Gln, Gln3, and Asp2-Gln10-Lys2 by assuming the physically relevant solutions to eq 3. The distributions of Gln3 and Asp2-Gln10-Lys2 populate \( \chi_3 \) angles similar to that of Gln. This suggests that primary amides of Gln3 and Asp2-Gln10-Lys2 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.\(^64\)

**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 AmIII 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 \( \chi_3 \) angle in the \( \chi_3 \) 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 8:

\[
\Delta G(\chi_{3,i}) = -RT \ln \left( \frac{p(\chi_{3,i})}{p(\chi_{3,0})} \right)
\]

(9)

where \( p(\chi_{3,i})/p(\chi_{3,0}) \) is the ratio of populations with \( \chi_{3,i} \) angles \( \chi_{3,j} \) 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 8 that each \( \chi_{3,i} \) dihedral angle state has a degeneracy of one.
torsional force constants along the conjugation between these two orbitals increases the \( \chi \) energy landscapes in Figure 11 to eq 10 to determine the AmIII\(^p\) 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 AmIII\(^p\) frequency and the \( \chi_3 \) and \( \chi_3 \) 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 AmIII\(^p\) spectroscopic marker band enables us to experimentally probe conformations of the Gln side chains of polyGln fibrils in order to obtain new, molecular-level insights into fibril structures.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b07651.

Description of X-ray crystallographic methods to determine crystal structures, circular dichroism (CD) methodology (PDF)

MATLAB programs used to calculate PEDs from DFT calculations (ZIP)

1-glutamine t-butyl ester HCl (CIF)

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: asher@pitt.edu.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

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 supercomputing resources provided. We thank Dr. Sergei V. Bykov for useful discussions and Prof. Roland Dunbrack, Jr., for providing his rotamer database.

## REFERENCES


