

# **Polyglutamine Fibrils: New Insights into Antiparallel $\beta$ -sheet Conformational Preference and Side Chain Structure**

## **Supporting Information**

David Punihaole<sup>†</sup>, Riley J. Workman<sup>‡</sup>, Zhenmin Hong<sup>†</sup>, Jeffry D. Madura<sup>\*‡</sup>,  
and Sanford A. Asher<sup>\*†</sup>

<sup>†</sup>*Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

<sup>‡</sup>*Department of Chemistry and Biochemistry, Center for Computational Sciences, Duquesne  
University, Pittsburgh, Pennsylvania 15282*

<sup>\*</sup>Email: asher@pitt.edu; madura@duq.edu

# Experimental Section

## UVR 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 UVR 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. UVR spectra were collected using both  $\sim 197$  nm and  $\sim 204$  nm excitation.

## UVR Spectral Processing

All UVR 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 (e.g.  $\text{H}_2\text{O}$ ,  $\text{D}_2\text{O}$ , and  $\text{HDO}$ ) and Suprasil quartz from NMR tubes. The spectra were calibrated using the  $801.3\text{ cm}^{-1}$ ,  $1028.3\text{ cm}^{-1}$ ,  $1157.6\text{ cm}^{-1}$ ,  $1266.4\text{ cm}^{-1}$ , and  $1444.4\text{ cm}^{-1}$  bands of cyclohexane. The spectral contributions of water and quartz were removed using a method similar to that described by Banerjee and coworkers.<sup>1,2</sup>

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:

$$\mathbf{S}' = \mathbf{K}\mathbf{S}'_{\mathbf{P}} \tag{S1}$$

where  $\mathbf{S}'$  is the  $(n \times 1)$  row vector that represents the first-derivative (denoted by  $'$ ) of the experimentally measured raw spectrum, and  $\mathbf{S}'_{\mathbf{P}}$  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,  $\mathbf{K}$ , contains the least-squares scaling coefficients for each of the different background spectral components. The least-squares solution of  $\mathbf{K}$  for eq. S1 is:<sup>3</sup>

$$\mathbf{K} = (\mathbf{S}'_{\mathbf{P}}{}^T \mathbf{S}'_{\mathbf{P}})^{-1} \mathbf{S}'_{\mathbf{P}}{}^T \mathbf{S}' \quad (\text{S2})$$

where  $(\mathbf{S}'_{\mathbf{P}}{}^T \mathbf{S}'_{\mathbf{P}})^{-1} \mathbf{S}'_{\mathbf{P}}{}^T$  is the pseudo-inverse matrix of  $\mathbf{S}'_{\mathbf{P}}$ . Banerjee and coworkers<sup>1,2</sup> 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  $\mathbf{K}$ , eq. S3 was utilized to subtract the contributions of water and quartz to obtain the spectrum of the analyte of interest:

$$\mathbf{S}_{\text{analyte}} = \mathbf{S} - \mathbf{K} \mathbf{S}_{\mathbf{P}} \quad (\text{S3})$$

where  $\mathbf{S}_{\text{analyte}}$ ,  $\mathbf{S}$ , and  $\mathbf{S}_{\mathbf{P}}$  are the zeroth-derivative analyte, raw, and water/quartz UVRR spectra, respectively.

## 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. i.e.

$$S(\nu) = \sum_i \left[ f_i H_i e^{-\left(\frac{\nu - \nu_i}{w_i}\right)^2 (4 \ln(2))} + (1 - f_i) \frac{H_i}{4 \left(\frac{\nu - \nu_i}{w_i}\right)^2 + 1} \right] \quad (\text{S4})$$

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.

## Calculation of the $\Psi$ and $\chi_3$ Angle Distributions

The distributions of  $\Psi$  and  $\chi_3$  dihedral angles, shown in Figures 7 and 10 in the main text, were calculated using methodologies previously described in detail.<sup>4-7</sup> Briefly, we assume that the inhomogenously broadened, experimentally measured AmIII<sub>3</sub><sup>S</sup> and AmIII<sup>P</sup> band profiles,  $B(\nu)$ , can be modeled as the sum of  $M$  Lorentzian bands:

$$B(\nu) = \frac{1}{\pi} \sum_i^M \frac{p_i \Gamma^2}{\Gamma^2 + (\nu - \nu_i)^2} \quad (\text{S5})$$

where  $p_i$  is the probability for the  $i^{\text{th}}$  band to occur at center frequency  $\nu_i$ . The band width parameter,  $\Gamma$ , is the homogeneous linewidth of the AmIII<sub>3</sub><sup>S</sup> or AmIII<sup>P</sup> vibrations. We previously<sup>5,7</sup> estimated from peptide crystals that  $\Gamma$  is  $\sim 7.5 \text{ cm}^{-1}$  for the AmIII<sub>3</sub><sup>S</sup> and  $\sim 6.6 \text{ cm}^{-1}$  AmIII<sup>P</sup>. After decomposing the band profiles into Lorentzians, we then correlate the different  $i^{\text{th}}$  frequencies of the AmIII<sub>3</sub><sup>S</sup> and AmIII<sup>P</sup> band envelopes to their respective  $\Psi$  or  $\chi_3$  dihedral angles.

### Correlating the AmIII<sub>3</sub><sup>S</sup> Frequencies to $\Psi$ Angles

We used the following equation to correlate the AmIII<sub>3</sub><sup>S</sup> band frequencies to  $\Psi$  angles for the Figure 7 black distributions shown in the main text:

$$\nu_i(\Psi) = 1239 \text{ (cm}^{-1}\text{)} - 54 \text{ (cm}^{-1}\text{)} \sin(\Psi + 26^\circ) \quad (\text{S6})$$

Eq. S6 was derived by Mikhonin et al.<sup>6</sup> 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 7 distributions shown in blue were obtained by using:

$$\nu_i(\Psi, T) = 1250 \text{ (cm}^{-1}\text{)} - 54 \text{ (cm}^{-1}\text{)} \sin(\Psi + 26^\circ) + 0.06 \text{ (cm}^{-1}/^\circ\text{C)}(T - T_0) \quad (\text{S7})$$



where  $T = 22^\circ\text{C}$  is the experimental temperature and  $T_0 = 0^\circ\text{C}$ . Eq. S7 was derived by Mikhonin et al.<sup>6</sup> 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.

### Correlating the AmIII<sup>P</sup> Frequencies to $\chi_3$ Angles

In the case of the AmIII<sup>P</sup>, we used the following equation to correlate the band frequencies to  $\chi_3$  angles for the distributions shown in Figure 10:

$$\nu_i(\chi_3) = 1076 \text{ (cm}^{-1}\text{)} + 29 \text{ (cm}^{-1}\text{)} \cos(2\chi_3) + 9 \text{ (cm}^{-1}\text{)} \cos(\chi_3 + 99^\circ) \quad (\text{S8})$$

where  $\nu_i$  is the  $i^{\text{th}}$  AmIII<sup>P</sup> frequency. Eq. S8 was derived by Punihaole et al.<sup>7</sup> 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<sup>8</sup> that stronger (weaker) hydrogen bonding and higher (lower) dielectric environments upshift (downshift) the AmIII<sup>P</sup> 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. S8 since it averages these two competing effects. We are presently investigating which of these effects dominates the AmIII<sup>P</sup> frequency in polyQ fibrils.

It should be noted that in using eq. S8, each AmIII<sup>P</sup> frequency can be correlated to as many as four possible  $\chi_3$  dihedral angles. However, as discussed in detail by Punihaole et al.,<sup>7</sup>  $\chi_3$  dihedral angles that are greater than  $+90^\circ$  and less than  $-90^\circ$  are nearly forbidden for Gln and Asn. Thus, in using eq. S8, we only considered the physically relevant  $\chi_3$  angles that occur between  $-90^\circ$  and  $+90^\circ$ .

# Computational Section

## Density Functional Theory (DFT) Calculations

DFT calculations<sup>9</sup> were performed on the zwitterion form of the Gln amino acid (Figure S1) using the *GAUSSIAN 09* program.<sup>10</sup> The M06-2X density functional was employed using the 6-311++g\*\* basis set.<sup>11</sup> 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.<sup>7</sup>

## 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.

$$RMSD = \frac{1}{N} \sqrt{\sum_{i=1}^N [(x_i - x_{ref})^2 + (y_i - y_{ref})^2 + (z_i - z_{ref})^2]} \quad (S9)$$

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  $x_{ref}, y_{ref}, z_{ref}$  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 glutamine 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 S2 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.

## Results and Discussion

### UVRR of NDQ10 and DQ10 Fibril Films

Figure S3 shows the 197 nm – 204 nm UVRR difference spectra of dried NDQ10 and DQ10 fibril films. The AmI<sup>P</sup> and AmII<sup>P</sup> bands are located at  $\sim 1660\text{ cm}^{-1}$  and  $\sim 1612\text{ cm}^{-1}$ , respectively, for both NDQ10 and DQ10. These bands negligibly shift compared to the AmI<sup>P</sup> and AmII<sup>P</sup> bands of fibrils in solution, as shown in the Figure 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<sub>2</sub> wagging band is located at  $\sim 1413\text{ cm}^{-1}$  for DQ10 fibril films and at  $\sim 1410\text{ cm}^{-1}$  for NDQ10 fibril films. Compared to the Figure 3 197 nm – 204 nm UVRR difference spectra in the main text, the CH<sub>2</sub> wagging band downshifts  $\sim 5\text{ cm}^{-1}$  for NDQ10 and  $\sim 17\text{ cm}^{-1}$  for DQ10 fibrils upon dehydration. We attribute this  $\sim 17\text{ cm}^{-1}$  downshift of the CH<sub>2</sub> wagging band in DQ10 fibrils upon dehydration to a local dielectric environment change of the methylene groups around of the Gln side chains.

## 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<sub>2</sub>O/D<sub>2</sub>O (Figure 5 in the main text). In assigning the Figure 5 spectrum, we assume that vibrations containing significant contributions of C<sub>δ</sub>–N<sub>ε2</sub> stretching show resonance enhancement in the the UVRR spectrum because the electronic excited state is expected to be expanded along this coordinate.<sup>12</sup> Tables S1 and S2 show the potential energy distributions (PEDs) obtained from the DFT calculations for the “*cis*-N<sub>ε2</sub>HD” and “*trans*-N<sub>ε2</sub>HD” species of the mono-deuterated primary amide side chains of Gln. Our band assignments of the Figure 5 spectrum from the main text are shown in Table S3.

### 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<sup>13,14</sup> of partially deuterated acetamide.

We assign the AmI vibration to a band located at  $\sim 1660\text{ cm}^{-1}$ . Our normal mode analysis indicates that the PED of this vibration consists mostly of C<sub>δ</sub>=O<sub>ε1</sub> stretching ( $\sim 77\%$ ), but also contains significant contributions of C<sub>δ</sub>–N<sub>ε2</sub> stretching and N<sub>ε1</sub>C<sub>δ</sub>C<sub>γ</sub> bending. The PED of this vibration is essentially the same as that of the AmI<sup>S</sup> and AmI<sup>P</sup> vibrations.<sup>7,8</sup>

The DFT calculations also predict the appearance of AmII<sup>S</sup>- and AmIII<sup>S</sup>-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<sup>S</sup> and AmIII<sup>S</sup> vibrations. Therefore, in the case of the mono-deuterated primary amides, we searched for vibrations that contain significant contributions of C<sub>δ</sub>–N<sub>ε2</sub> stretching and N<sub>ε2</sub>HD deformations. As shown in Table S1 and S2, there are several vibrations that contain significant contributions of C<sub>δ</sub>–N<sub>ε2</sub> stretching,

$\text{N}_{\epsilon_2}\text{HD}$  scissoring, and  $\text{N}_{\epsilon_2}\text{HD}$  rocking. We assign these vibrations to  $\text{AmII}^{\text{S}}$ -like,  $\text{AmIII}^{\text{S}}$ -like, or  $\text{AmIII}'^{\text{S}}$ -like vibrations.

There are two  $\text{AmII}^{\text{S}}$ -like vibrations predicted by the DFT calculations to be at  $\sim 1524 \text{ cm}^{-1}$  for the *trans*- $\text{N}_{\epsilon_2}\text{HD}$  species and  $\sim 1479 \text{ cm}^{-1}$  for the *cis*- $\text{N}_{\epsilon_2}\text{HD}$  species. Both of these vibrations contain significant  $\text{C}_{\delta}\text{-N}_{\epsilon_2}$  stretching and  $\text{N}_{\epsilon_2}\text{HD}$  scissoring, although the  $\sim 1524 \text{ cm}^{-1}$  mode also contains  $\text{N}_{\epsilon_2}\text{HD}$  rocking. The predicted  $\sim 1524 \text{ cm}^{-1}$  mode is experimentally observed at  $\sim 1547 \text{ cm}^{-1}$ , while the predicted  $\sim 1479 \text{ cm}^{-1}$  vibration is observed at  $\sim 1476 \text{ cm}^{-1}$ .

The DFT calculations indicate that two  $\text{AmIII}^{\text{S}}$ -like vibrations for the *trans*- $\text{N}_{\epsilon_2}\text{HD}$  species are predicted to occur at  $\sim 1247 \text{ cm}^{-1}$  and  $\sim 1329 \text{ cm}^{-1}$ . Both vibrations contain significant contributions of  $\text{C}_{\delta}\text{-N}_{\epsilon_2}$  stretching and  $\text{N}_{\epsilon_2}\text{HD}$  scissoring. However, as with the canonical  $\text{AmIII}^{\text{S}}$  modes observed in peptides, these vibrations are significantly coupled since they contain significant contributions of  $\text{CH}_2$  wagging and twisting, as well as  $\text{C}_{\alpha}\text{-H}$  rocking.<sup>15</sup> We assign these  $\text{AmIII}^{\text{S}}$ -like vibrations to bands observed at  $\sim 1247 \text{ cm}^{-1}$  and  $\sim 1308 \text{ cm}^{-1}$ .

There are also two  $\text{AmIII}'^{\text{S}}$ -like vibrations predicted to be at  $\sim 1055 \text{ cm}^{-1}$  for the *trans*- $\text{N}_{\epsilon_2}\text{HD}$  species and  $\sim 953 \text{ cm}^{-1}$  for the *cis*- $\text{N}_{\epsilon_2}\text{HD}$  species. These vibrations resemble  $\text{AmIII}'^{\text{S}}$ -like modes since they both contain large contributions of  $\text{N}_{\epsilon_2}\text{HD}$  rocking. This is analogous to the canonical  $\text{AmIII}'^{\text{S}}$ , which is mostly N-D in-plane bending.

## 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

$$\sum_{i=1}^{n_i} \frac{1}{1 + \exp(\ln(n_i/n_j) + \beta\Delta U_{ij} - \beta\Delta G)} - \sum_{i=1}^{n_j} \frac{1}{1 + \exp(\ln(n_j/n_i) + \beta\Delta U_{ji} - \beta\Delta G)} = 0 \quad (\text{S10})$$

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.<sup>16</sup> 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.<sup>17</sup> Figure S4 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).

## Hydrogen Bonding Analysis

Figure S5 show the number of hydrogen bonds of fibril models  $a$  and  $b$  (from Figure 1 in the main text), which were obtained from the MD trajectories. Table S4 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  $\beta$ -sheet model  $a$  forms, on average, more peptide bond-peptide bond hydrogen bonds than the parallel  $\beta$ -sheet model  $b$ . In particular, the antiparallel  $\beta$ -sheet forms significantly more peptide bond-peptide bond hydrogen bonds than the parallel  $\beta$ -sheet. It is interesting to note that, if we assume a peptide bond-peptide bond hydrogen bond energy of  $5 \text{ kJ mol}^{-1}$ , this hydrogen bonding difference would account for the majority of the  $160.5 \text{ kJ mol}^{-1}$  free energy difference between the antiparallel and parallel  $\beta$ -sheet fibril models.

## References

- [1] Banerjee, S.; Li, D. *Appl. Spectrosc.* **1991**, *45*, 1047–1049.
- [2] Friese, M. A.; Banerjee, S. *Appl. Spectrosc.* **1992**, *46*, 246–248.
- [3] Chi, Z.; Chen, X. G.; Holtz, J. S. W.; ; Asher, S. A. *Biochemistry* **1998**, *37*, 2854–2864.
- [4] Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775–11781.
- [5] Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433–8440.
- [6] Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. *J. Phys. Chem. B* **2006**, *110*, 1928–1943.
- [7] Punihaole, D.; Hong, Z.; Jakubek, R. S.; Dahlburg, E. M.; Geib, S.; Asher, S. A. *J. Phys. Chem. B* **2015**, *119*, 13039–13051.
- [8] Punihaole, D.; Jakubek, R. S.; Dahlburg, E. M.; Hong, Z.; Myshakina, N. S.; Geib, S.; Asher, S. A. *J. Phys. Chem. B* **2015**, *119*, 3931–3939.
- [9] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [10] Frisch, M. J. et al. Gaussian09 Revision D.01. Gaussian Inc. Wallingford CT 2009.
- [11] Zhao, Y.; Truhlar, D. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- [12] Clark, L. B. *J. Am. Chem. Soc.* **1995**, *117*, 7974–7986.
- [13] Uni, T.; Machida, K.; Saito, Y. *Bull. Chem. Soc. Jpn* **1969**, *42*, 897–904.
- [14] Uno, T.; Machida, K.; Saito, Y. *Spectrochim. Acta, Part A* **1971**, *27*, 833 – 844.
- [15] Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2002**, *106*, 3391–3394.

- [16] Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245 – 268.
- [17] Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 1–13.



**Table S1: DFT Calculated Frequencies (cm<sup>-1</sup>) and Assignments of *cis*-Glutamine-N<sub>ε2</sub>HD**

| Assign.                             | Calc. | PED <sup>a,b</sup> (≥5% contribution)   |
|-------------------------------------|-------|---|
| $\nu_{as}\text{COO}$                | 1715  | $-\nu\text{C}'\text{O}$ (53), $\nu\text{C}'\text{O}$ (33), $\rho\text{C}_{\alpha}\text{C}'$ (7)   |
| $\delta_{as}\text{NH}_3$            | 1666  | $-\delta_{as}'\text{NH}_3$ (48), $\delta_{as}\text{NH}_3$ (44), $\rho\text{NH}_3$ (5)   |
| $\delta_{as}\text{NH}_3$            | 1621  | $-\delta_{as}\text{NH}_3$ (40), $-\delta_{as}'\text{NH}_3$ (40), $-\delta_s\text{NH}_3$ (10)  |
| $\sigma\text{CH}_2$                 | 1494  | $\sigma\text{C}_{\beta}\text{H}_2$ (87)   |
| AmII-like                           | 1479  | $-\sigma\text{N}_{\epsilon2}\text{HD}$ (37), $-\nu\text{C}_{\delta}\text{N}_{\epsilon2}$ (28), $\nu\text{C}_{\gamma}\text{C}_{\delta}$ (10), $-\omega\text{C}_{\gamma}\text{H}_2$ (9), $\rho\text{C}_{\delta}\text{O}_{\epsilon1}$ (6)  |
| $\delta_s\text{NH}_3$               | 1464  | $\delta_s\text{NH}_3$ (50), $-\sigma\text{C}_{\gamma}\text{H}_2$ (20)   |
| $\sigma\text{CH}_2$                 | 1460  | $\sigma\text{C}_{\gamma}\text{H}_2$ (61), $\delta_s\text{NH}_3$ (21)  |
| $\omega\text{CH}_2$                 | 1437  | $\omega\text{C}_{\beta}\text{H}_2$ (19), $\sigma\text{N}_{\epsilon2}\text{HD}$ (14), $-\omega\text{C}_{\gamma}\text{H}_2$ (11), $-\nu\text{C}'\text{O}$ (10), $\rho\text{C}_{\alpha}\text{H}$ (8), $-\nu\text{C}_{\beta}\text{C}_{\gamma}$ (5), $-\delta_s\text{NH}_3$ (5), $\sigma\text{C}_{\gamma}\text{H}_2$ (5) |
| $\nu_s\text{COO}$                   | 1414  | $\nu\text{C}'\text{O}$ (22), $-\omega\text{C}_{\gamma}\text{H}_2$ (13), $\sigma\text{N}_{\epsilon2}\text{HD}$ (11), $-\nu\text{C}_{\alpha}\text{C}'$ (9), $-\rho\text{C}_{\alpha}\text{H}$ (7), $\nu\text{C}_{\gamma}\text{C}_{\delta}$ (7), $-\beta\text{C}'\text{O}_2^-$ (6), $\nu\text{C}'\text{O}$ (6)          |
| $\rho\text{CH}$                     | 1387  | $\rho\text{C}_{\alpha}\text{H}$ (28), $\omega\text{C}_{\beta}\text{H}_2$ (11), $-\rho'\text{C}_{\alpha}\text{H}$ (9), $\nu\text{C}'\text{O}$ (8), $\nu\text{C}_{\alpha}\text{C}_{\beta}$ (7), $-\tau\text{C}_{\beta}\text{H}_2$ (6)   |
| $\rho\text{CH}$                     | 1357  | $\rho\text{C}_{\alpha}\text{H}$ (27), $\tau\text{C}_{\beta}\text{H}_2$ (19), $-\omega\text{C}_{\beta}\text{H}_2$ (13), $-\tau\text{C}_{\gamma}\text{H}_2$ (11)  |
| $\omega\text{CH}_2+\tau\text{CH}_2$ | 1347  | $\omega\text{C}_{\beta}\text{H}_2$ (21), $\tau\text{C}_{\beta}\text{H}_2$ (21), $\rho'\text{C}_{\alpha}\text{H}$ (11), $\nu\text{C}_{\delta}\text{N}_{\epsilon2}$ (11), $-\rho\text{C}_{\delta}\text{O}_{\epsilon1}$ (5)  |
| $\tau\text{CH}_2$                   | 1309  | $\tau\text{C}_{\gamma}\text{H}_2$ (35), $\rho'\text{C}_{\alpha}\text{H}$ (26), $-\rho\text{C}_{\beta}\text{H}_2$ (8), $\rho\text{C}_{\alpha}\text{H}$ (6)   |
| $\omega\text{CH}_2$                 | 1272  | $\omega\text{C}_{\gamma}\text{H}_2$ (43), $\omega\text{C}_{\beta}\text{H}_2$ (20), $-\nu\text{C}_{\delta}\text{N}_{\epsilon2}$ (13)   |
| $\tau\text{CH}_2+\nu\text{CC}$      | 1214  | $\tau\text{C}_{\beta}\text{H}_2$ (21), $\nu\text{C}_{\alpha}\text{C}_{\beta}$ (18), $\tau\text{C}_{\gamma}\text{H}_2$ (16), $\rho'\text{NH}_3$ (13), $-\delta\text{NC}_{\alpha}\text{C}'\text{O}_2^-$ (5)   |
| $\rho\text{CH}+\tau\text{CH}_2$     | 1153  | $-\rho'\text{C}_{\alpha}\text{H}$ (20), $\tau\text{C}_{\gamma}\text{H}_2$ (16), $-\rho'\text{NH}_3$ (14), $\tau\text{C}_{\beta}\text{H}_2$ (12), $-\nu\text{C}_{\alpha}\text{C}_{\beta}$ (8), $\rho\text{C}_{\gamma}\text{H}_2$ (6)   |
| $\nu\text{CC}$                      | 1117  | $\nu\text{C}_{\beta}\text{C}_{\gamma}$ (42), $\nu\text{C}_{\alpha}\text{N}$ (10), $\rho\text{NH}_3$ (8), $-\nu\text{C}_{\alpha}\text{C}_{\beta}$ (7), $-\delta'\text{NC}_{\alpha}\text{C}'\text{O}_2^-$ (5), $-\sigma\text{C}_{\alpha}\text{C}_{\beta}\text{C}_{\gamma}$ (5)  |
| $\rho\text{NH}_3$                   | 1105  | $\rho\text{NH}_3$ (28), $-\nu\text{C}_{\beta}\text{C}_{\gamma}$ (12), $-\rho'\text{C}_{\alpha}\text{H}$ (9), $-\delta'\text{NC}_{\alpha}\text{C}'\text{O}_2^-$ (8), $-\rho\text{C}_{\alpha}\text{H}$ (8), $-\rho\text{C}_{\beta}\text{H}_2$ (7)   |
| $\nu\text{CN}$                      | 1040  | $\nu\text{C}_{\alpha}\text{N}$ (35), $-\nu\text{C}_{\beta}\text{C}_{\gamma}$ (11), $\rho\text{C}_{\beta}\text{H}_2$ (9), $\rho\text{C}_{\gamma}\text{H}_2$ (6), $\rho'\text{C}_{\alpha}\text{H}$ (5)  |
| $\rho\text{NH}_3+\nu\text{CN}$      | 1005  | $\rho\text{NH}_3$ (22), $-\nu\text{C}_{\alpha}\text{N}$ (20), $\rho\text{C}_{\gamma}\text{H}_2$ (12), $\rho\text{C}_{\beta}\text{H}_2$ (12), $\nu\text{C}_{\alpha}\text{C}_{\beta}$ (10)  |
| $\rho\text{NH}_3$                   | 983   | $-\rho'\text{NH}_3$ (29), $\nu\text{C}_{\alpha}\text{C}_{\beta}$ (13), $-\sigma\text{C}_{\alpha}\text{C}_{\beta}\text{C}_{\gamma}$ (6), $\rho\text{N}_{\epsilon2}\text{HD}$ (6), $-\sigma\text{C}_{\beta}\text{C}_{\gamma}\text{C}_{\delta}$ (5), $\nu\text{C}_{\gamma}\text{C}_{\delta}$ (5)                       |
| AmIII <sup>S</sup> -like            | 953   | $\rho\text{N}_{\epsilon2}\text{HD}$ (15), $\nu\text{C}_{\delta}\text{N}_{\epsilon2}$ (14), $-\nu\text{C}_{\alpha}\text{C}_{\beta}$ (13), $\rho'\text{NH}_3$ (11), $\nu\text{C}_{\gamma}\text{C}_{\delta}$ (9), $\nu\text{C}_{\alpha}\text{C}'$ (8), $-\sigma\text{N}_{\epsilon2}\text{HD}$ (7)                      |

<sup>a</sup> $\text{C}_{\delta}$ -N<sub>ε2</sub> stretching and N<sub>ε2</sub>HD scissoring components in PED are in bold.

<sup>b</sup> $\nu$ : stretch;  $\delta_{as}$ : asymmetric deformation;  $\delta_s$ : symmetric deformation;  $\delta$ : deformation;  $\sigma$ : scissoring;  $\rho$ : rocking;  $\omega$ : wagging;  $\beta$ : in-plane bending;  $\tau$ : twisting.

Table S2: DFT Calculated ( $\text{cm}^{-1}$ ) and Assignments of *trans*-Glutamine- $\text{N}_{\epsilon 2}\text{HD}$

| Assign.                                 | Calc. | PED <sup>a,b</sup> ( $\geq 5\%$ contribution)   |
|---|-------|---|
| AmI-like                                | 1742  | $-\nu\text{C}_\delta\text{O}_{\epsilon 1}$ (77), $-\beta\text{NC}_\delta\text{C}_\gamma$ (7), $\nu\text{C}_\delta\text{N}_{\epsilon 2}$ ( <b>7</b> )  |
| $\nu_{as}\text{COO}$                    | 1715  | $-\nu\text{C}'\text{O}$ (53), $\nu\text{C}'\text{O}$ (33), $\rho\text{C}_\alpha\text{C}'$ (7)   |
| $\delta_{as}\text{NH}_3$                | 1666  | $\delta_{as}'\text{NH}_3$ (48), $-\delta_{as}\text{NH}_3$ (44), $-\rho\text{NH}_3$ (5)  |
| $\delta_{as}\text{NH}_3$                | 1621  | $-\delta_{as}\text{NH}_3$ (40), $-\delta_{as}'\text{NH}_3$ (40), $-\delta_s\text{NH}_3$ (10)  |
| AmII-like                               | 1524  | $\sigma\text{N}_{\epsilon 2}\text{HD}$ ( <b>49</b> ), $\nu\text{C}_\delta\text{N}_{\epsilon 2}$ ( <b>30</b> ), $-\nu\text{C}_\gamma\text{C}_\delta$ (6), $-\rho\text{N}_{\epsilon 2}\text{HD}$ ( <b>6</b> ), $-\rho\text{C}_\delta\text{O}_{\epsilon 1}$ (5)  |
| $\sigma\text{CH}_2$                     | 1494  | $\sigma\text{C}_\beta\text{H}_2$ (88)   |
| $\delta_s\text{NH}_3+\sigma\text{CH}_2$ | 1465  | $-\delta_s\text{NH}_3$ (36), $\sigma\text{C}_\gamma\text{H}_2$ (33), $\omega\text{C}_\gamma\text{H}_2$ (5)  |
| $\sigma\text{CH}_2$                     | 1460  | $-\sigma\text{C}_\gamma\text{H}_2$ (47), $-\delta_s\text{NH}_3$ (34), $\nu\text{C}'\text{O}$ (3)  |
| $\omega\text{CH}_2$                     | 1435  | $-\omega\text{C}_\beta\text{H}_2$ (21), $\nu\text{C}'\text{O}$ (13), $-\rho\text{C}_\alpha\text{H}$ (11), $\omega\text{C}_\gamma\text{H}_2$ (10), $-\sigma\text{C}_\gamma\text{H}_2$ (7), $-\nu\text{C}_\alpha\text{C}'$ (6), $\nu\text{C}_\beta\text{C}_\gamma$ (5), $\delta_s\text{NH}_3$ (5)   |
| $\nu_s\text{COO}$                       | 1407  | $\nu\text{C}'\text{O}$ (22), $-\omega\text{C}_\gamma\text{H}_2$ (20), $-\nu\text{C}_\alpha\text{C}'$ (9), $\nu\text{C}_\gamma\text{C}_\delta$ (7), $-\beta\text{C}'\text{O}_2^-$ (6), $\nu\text{C}'\text{O}$ (6), $\sigma\text{N}_{\epsilon 2}\text{HD}$ (5)  |
| $\rho\text{CH}$                         | 1384  | $\rho\text{C}_\alpha\text{H}$ (40), $-\rho'\text{C}_\alpha\text{H}$ (11), $-\tau\text{C}_\beta\text{H}_2$ (6), $\nu\text{C}_\alpha\text{C}_\beta$ (6), $\nu\text{C}'\text{O}$ (5)   |
| $\tau\text{CH}_2$                       | 1356  | $\tau\text{C}_\beta\text{H}_2$ (35), $\rho\text{C}_\alpha\text{H}$ (16), $-\tau\text{C}_\gamma\text{H}_2$ (15), $\rho'\text{C}_\alpha\text{H}$ (6)  |
| AmIIIS-like                             | 1329  | $\omega\text{C}_\beta\text{H}_2$ (39), $\rho'\text{C}_\alpha\text{H}$ (9), $-\sigma\text{N}_{\epsilon 2}\text{HD}$ ( <b>9</b> ), $\nu\text{C}_\delta\text{N}_{\epsilon 2}$ ( <b>8</b> ), $-\nu\text{C}_\gamma\text{C}_\delta$ (5)   |
| $\tau\text{CH}_2$                       | 1307  | $-\tau\text{C}_\gamma\text{H}_2$ (34), $-\rho'\text{C}_\alpha\text{H}$ (22), $\rho\text{C}_\beta\text{H}_2$ (8), $-\rho\text{C}_\alpha\text{H}$ (8)   |
| AmIIIS-like                             | 1247  | $-\omega\text{C}_\gamma\text{H}_2$ (25), $\nu\text{C}_\delta\text{N}_{\epsilon 2}$ (20), $-\sigma\text{N}_{\epsilon 2}\text{HD}$ ( <b>9</b> ), $\rho'\text{NH}_3$ (7), $-\rho\text{C}_\delta\text{O}_{\epsilon 1}$ (6), $\nu\text{C}_\alpha\text{C}_\beta$ (6), $-\omega\text{C}_\beta\text{H}_2$ (5), $\tau\text{C}_\beta\text{H}_2$ (5) |
| $\tau\text{CH}_2$                       | 1207  | $-\tau\text{C}_\beta\text{H}_2$ (18), $-\tau\text{C}_\gamma\text{H}_2$ (15), $-\nu\text{C}_\alpha\text{C}_\beta$ (13), $-\omega\text{C}_\gamma\text{H}_2$ (11), $-\rho'\text{NH}_3$ (10), $\nu\text{C}_\delta\text{N}_{\epsilon 2}$ (7)   |
| $\rho\text{CH}+\tau\text{CH}_2$         | 1153  | $\rho'\text{C}_\alpha\text{H}$ (19), $-\tau\text{C}_\gamma\text{H}_2$ (16), $\rho'\text{NH}_3$ (14), $-\tau\text{C}_\beta\text{H}_2$ (11), $\nu\text{C}_\alpha\text{C}_\beta$ (9), $-\rho\text{C}_\gamma\text{H}_2$ (6)   |
| $\nu\text{CN}$                          | 1116  | $-\nu\text{C}_\beta\text{C}_\gamma$ (39), $-\nu\text{C}_\alpha\text{N}$ (10), $-\rho\text{NH}_3$ (10), $\nu\text{C}_\alpha\text{C}_\beta$ (6), $\delta'\text{NC}_\alpha\text{C}'\text{O}_2^-$ (6), $\rho'\text{C}_\alpha\text{H}$ (5), $\sigma\text{C}_\alpha\text{C}_\beta\text{C}_\gamma$ (5)   |
| $\rho\text{NH}_3$                       | 1104  | $-\rho\text{NH}_3$ (27), $\nu\text{C}_\beta\text{C}_\gamma$ (15), $\rho'\text{C}_\alpha\text{H}$ (8), $\rho\text{C}_\alpha\text{H}$ (7), $\rho\text{C}_\beta\text{H}_2$ (7), $\delta'\text{NC}_\alpha\text{C}'\text{O}_2^-$ (7)   |
| AmIIIS-like                             | 1055  | $-\rho\text{N}_{\epsilon 2}\text{HD}$ (27), $-\nu\text{C}_\gamma\text{C}_\delta$ (14), $-\nu\text{C}_\delta\text{N}_{\epsilon 2}$ (9), $-\nu\text{C}_\alpha\text{N}$ (9), $-\nu\text{C}_\delta\text{O}_{\epsilon 1}$ (8), $\nu\text{C}_\beta\text{C}_\gamma$ (5), $-\rho\text{C}_\delta\text{O}_{\epsilon 1}$ (5)                         |
| $\nu\text{CN}$                          | 1032  | $\nu\text{C}_\alpha\text{N}$ (31), $-\rho\text{N}_{\epsilon 2}\text{HD}$ (10), $\rho'\text{NH}_3$ (7), $-\nu\text{C}_\beta\text{C}_\gamma$ (7), $\rho\text{C}_\beta\text{H}_2$ (6)  |
| $\rho\text{NH}_3$                       | 1003  | $\rho\text{NH}_3$ (25), $-\nu\text{C}_\alpha\text{N}$ (19), $\rho\text{C}_\gamma\text{H}_2$ (14), $\rho\text{C}_\beta\text{H}_2$ (14), $\nu\text{C}_\alpha\text{C}_\beta$ (7), $-\tau\text{C}_\beta\text{H}_2$ (5)  |
| $\rho\text{NH}_3$                       | 973   | $-\rho'\text{NH}_3$ (37), $\nu\text{C}_\alpha\text{C}_\beta$ (25), $-\nu\text{C}_\alpha\text{C}'$ (9), $-\sigma\text{C}_\alpha\text{C}_\beta\text{C}_\gamma$ (7), $\nu\text{C}_\alpha\text{N}$ (6)  |

<sup>a</sup> $\text{C}_\delta\text{-N}_{\epsilon 2}$  stretching and  $\text{N}_{\epsilon 2}\text{HD}$  scissoring components in PED are in bold.

<sup>b</sup> $\nu$ : stretch;  $\delta_{as}$ : asymmetric deformation;  $\delta_s$ : symmetric deformation;  $\delta$ : deformation;  $\sigma$ : scissoring;  $\rho$ : rocking;  $\omega$ : wagging;  $\beta$ : in-plane bending;  $\tau$ : twisting.

**Table S3: UVRR Band Frequencies ( $\text{cm}^{-1}$ ) and Assignments of *trans*- and *cis*-Glutamine- $\text{N}_{\epsilon 2}\text{HD}$**

| <i>cis</i> - $\text{N}_{\epsilon 2}\text{HD}$ |       |                                       | <i>trans</i> - $\text{N}_{\epsilon 2}\text{HD}$ |       |   |
|---|-------|---------------------------------------|---|-------|---|
| Expt.   | Calc. | Assign. <sup>a</sup>                  | Expt.   | Calc. | Assign. <sup>a</sup>                      |
| 1628  | 1715  | $\nu_{as}\text{COO}$                  | 1658  | 1742  | AmI-like                                  |
|   | 1666  | $\delta_{as}\text{NH}_3$              | 1628  | 1715  | $\nu_{as}\text{COO}$                      |
|   | 1621  | $\delta_{as}\text{NH}_3$              |   | 1666  | $\delta_{as}\text{NH}_3$                  |
| 1476  |       |                                       |   | 1621  | $\delta_{as}\text{NH}_3$                  |
|   | 1494  | $\sigma\text{CH}_2$                   | 1547  | 1524  | AmII-like                                 |
|   | 1479  | AmII-like                             |   | 1494  | $\sigma\text{CH}_2$                       |
|   | 1464  | $\delta_s\text{NH}_3$                 |   | 1465  | $\delta_s\text{NH}_3 + \sigma\text{CH}_2$ |
| 1449  | 1460  | $\sigma\text{CH}_2$                   | 1449  | 1460  | $\sigma\text{CH}_2$                       |
| 1420  | 1437  | $\omega\text{CH}_2$                   | 1420  | 1435  | $\omega\text{CH}_2$                       |
| 1420  | 1414  | $\nu_s\text{COO}$                     | 1420  | 1407  | $\nu_{as}\text{COO}$                      |
| 1398  | 1387  | $\rho\text{CH}$                       | 1398  | 1384  | $\rho\text{CH}$                           |
| 1360  | 1357  | $\rho\text{CH}$                       |   |       |   |
| 1335  |       |                                       | 1360  | 1356  | $\tau\text{CH}_2$                         |
|   | 1347  | $\omega\text{CH}_2 + \tau\text{CH}_2$ |   |       |   |
|   |       |                                       | 1308  | 1329  | AmIII <sup>S</sup> -like                  |
|   | 1309  | $\tau\text{CH}_2$                     |   | 1307  | $\tau\text{CH}_2$                         |
| 1278  | 1272  | $\omega\text{CH}_2$                   |   |       |   |
|   |       |                                       | 1247  | 1247  | AmIII <sup>S</sup> -like                  |
|   | 1214  | $\tau\text{CH}_2 + \nu\text{CC}$      |   | 1207  | $\tau\text{CH}_2$                         |
|   | 1153  | $\rho\text{CH} + \tau\text{CH}_2$     | 1153  | 1153  | $\rho\text{CH} + \tau\text{CH}_2$         |
|   | 1117  | $\nu\text{CC}$                        |   | 1116  | $\nu\text{CC}$                            |
|   | 1105  | $\rho\text{NH}_3$                     |   | 1104  | $\rho\text{NH}_3$                         |
|   |       |                                       | 1040  | 1055  | AmIII <sup>S</sup> -like                  |
|   | 1040  | $\nu\text{CN}$                        |   |       |   |
|   |       |                                       |   | 1032  | $\nu\text{CN}$                            |
|   | 1005  | $\rho\text{NH}_3 + \nu\text{CN}$      |   | 1003  | $\rho\text{NH}_3$                         |
|   | 983   |                                       |   |       | $\rho\text{NH}_3$                         |
|   |       |                                       |   | 973   | $\rho\text{NH}_3$                         |
| 964   | 953   | AmIII <sup>S</sup> -like              |   |       |   |

<sup>a</sup> $\nu$ : stretch;  $\delta_{as}$ : asymmetric deformation;  $\delta_s$ : symmetric deformation;  
 $\delta$ : deformation;  $\sigma$ : scissoring;  $\rho$ : rocking;  $\omega$ : wagging;  $\beta$ : in-plane bending;  
 $\tau$ : twisting.

**Table S4: Average number of hydrogen bonds for antiparallel and parallel  $\beta$ -sheet fibril models**

| H-bonding type                    | Parallel $\beta$ -strand fibril | Antiparallel $\beta$ -strand fibril |
|-----------------------------------|---------------------------------|-------------------------------------|
| Peptide-peptide                   | 106.7                           | 118.9                               |
| Peptide-solvent                   | 281.2                           | 256.5                               |
| Peptide backbone-side chain       | 36.0                            | 16.8                                |
| Peptide backbone-peptide backbone | 39.2                            | 66.8                                |
| Side chain-side chain             | 31.4                            | 35.3                                |

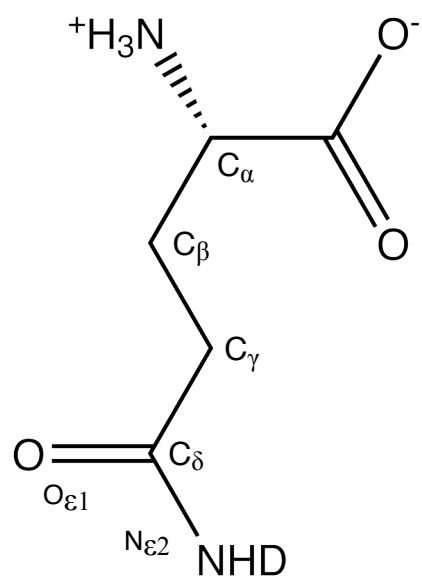


Figure S1: Atomic labeling scheme for glutamine used in DFT calculations.

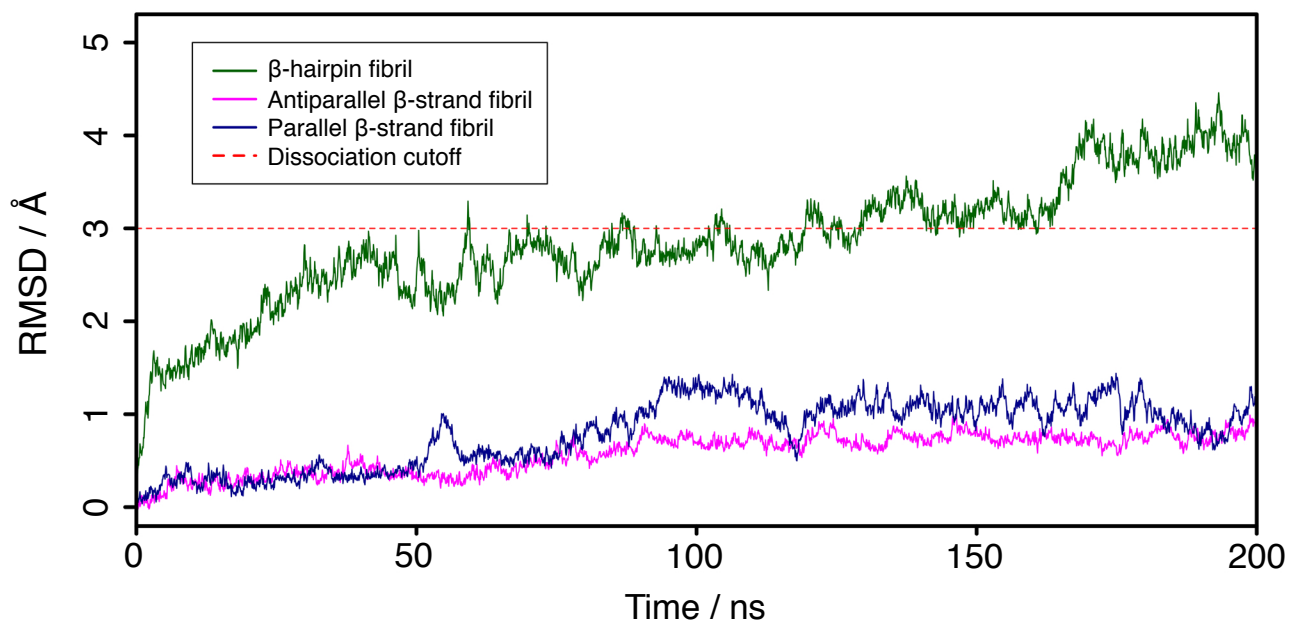


Figure S2: Plot of the evolution of backbone RMSD metric for three fibril models with respect to their initial structure.

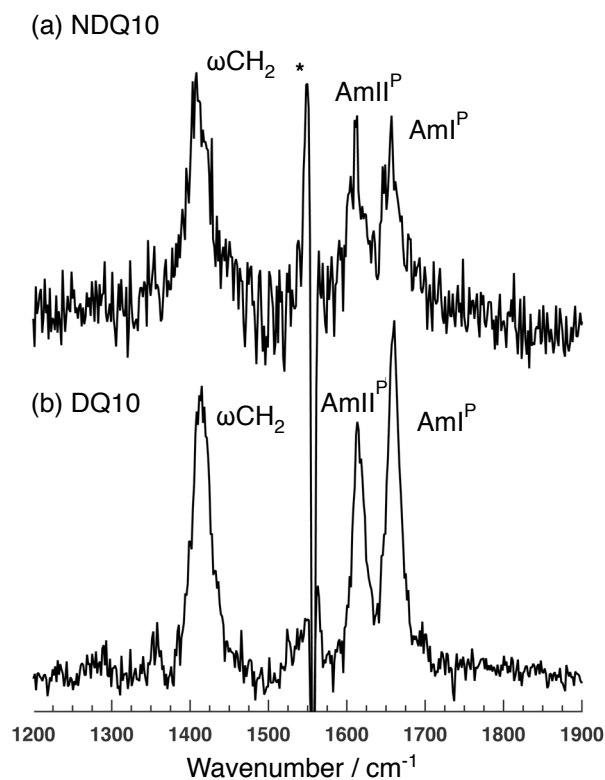


Figure S3: UVRR 197 nm – 204 nm difference spectra of fibril films prepared from (a) NDQ10 and (b) DQ10. The asterisk indicates an artifact of subtracting the intense  $O_2$  stretching band in the difference spectrum.

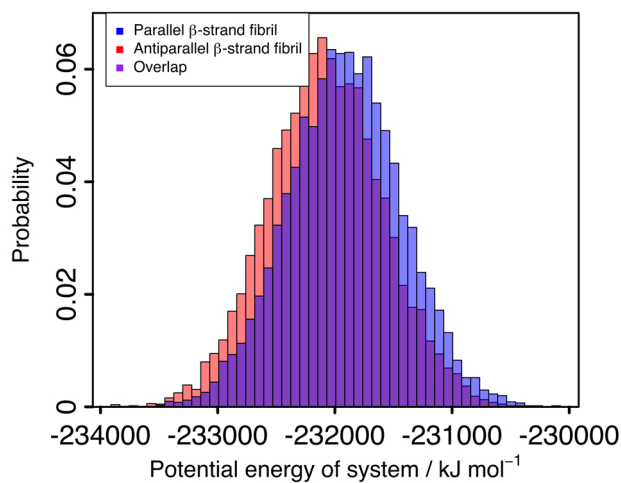


Figure S4: Potential energy distributions for the Figure 1 simulated model *a* antiparallel and model *b* parallel  $\beta$ -sheet fibril systems shown as red and blue histograms, respectively. These distributions represent the probability that a particular potential energy was sampled during the simulation 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.



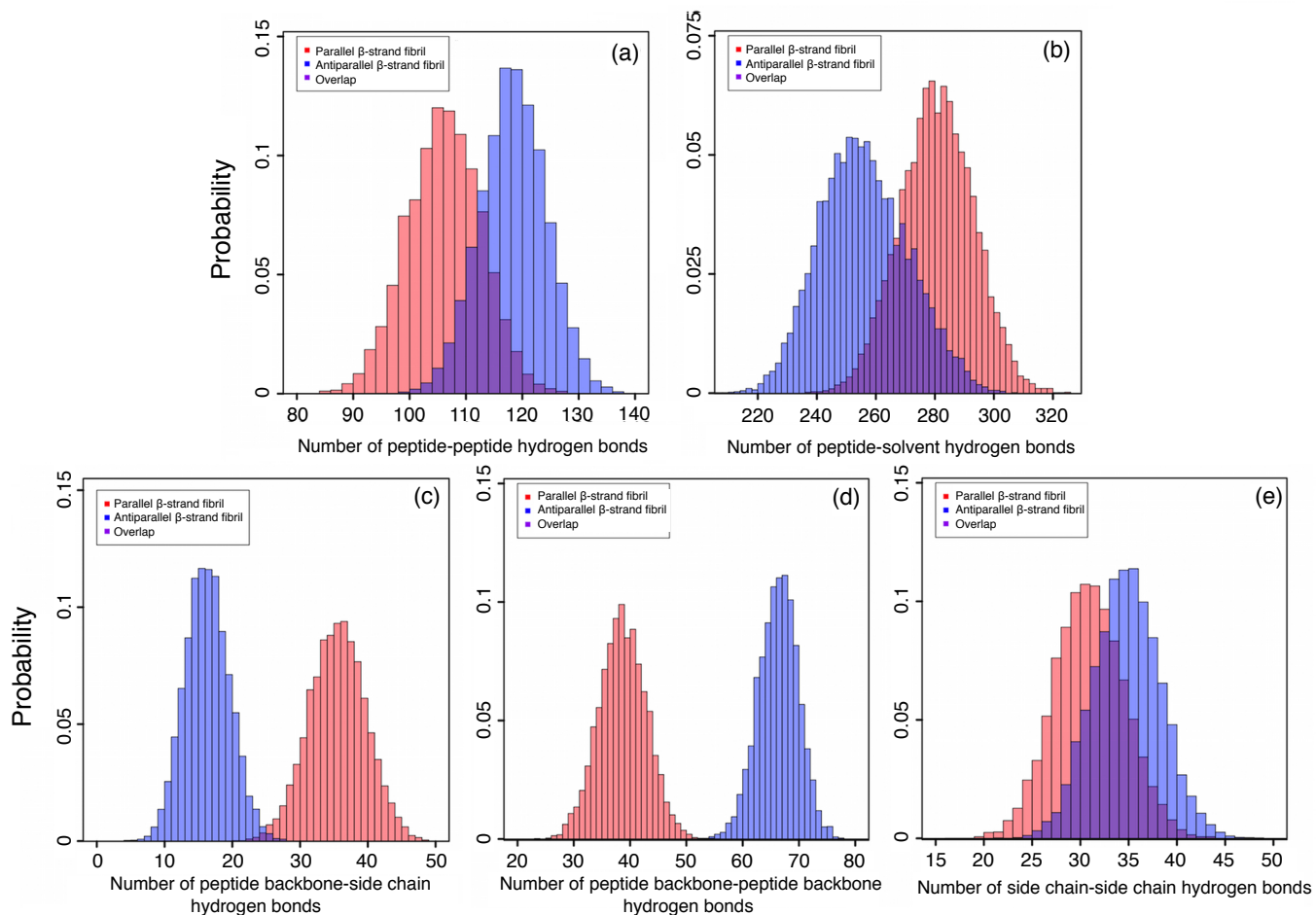


Figure S5: Histograms of the number of hydrogen bonds for models *a* and *b* in Figure 1. (a) peptide-peptide hydrogen bonds, (b) peptide-solvent hydrogen bonds, (c) peptide backbone-side chain hydrogen bonds, (d) peptide backbone-peptide backbone, and (e) side chain-side chain hydrogen bonds. Data was 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°.