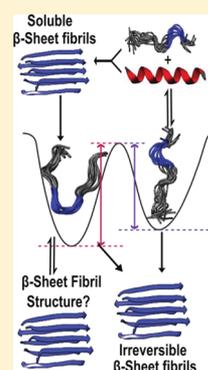


## Monomeric Polyglutamine Structures That Evolve into Fibrils

David Punihaole,<sup>†</sup> Ryan S. Jakubek,<sup>†</sup> Riley J. Workman,<sup>‡</sup> Lauren E. Marbella,<sup>†</sup> Patricia Campbell,<sup>§</sup> Jeffrey D. Madura,<sup>\*,‡</sup> and Sanford A. Asher<sup>\*,†</sup><sup>†</sup>Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States<sup>‡</sup>Department of Chemistry and Biochemistry, Center for Computational Sciences, Duquesne University, Pittsburgh, Pennsylvania 15282, United States<sup>§</sup>Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15260, United States

## Supporting Information

**ABSTRACT:** We investigate the solution and fibril conformations and structural transitions of the polyglutamine (polyQ) peptide, D<sub>2</sub>Q<sub>10</sub>K<sub>2</sub> (Q10), by synergistically using UV resonance Raman (UVRR) spectroscopy and molecular dynamics (MD) simulations. We show that Q10 adopts two distinct, monomeric solution conformational states: a collapsed  $\beta$ -strand and a PPII-like structure that do not readily interconvert. This clearly indicates a high activation barrier in solution that prevents equilibration between these structures. Using metadynamics, we explore the conformational energy landscape of Q10 to investigate the physical origins of this high activation barrier. We develop new insights into the conformations and hydrogen bonding environments of the glutamine side chains in the PPII and  $\beta$ -strand-like conformations in solution. We also use the secondary structure-inducing cosolvent, acetonitrile, to investigate the conformations present in low dielectric constant solutions with decreased solvent–peptide hydrogen bonding. As the mole fraction of acetonitrile increases, Q10 converts from PPII-like structures into  $\alpha$ -helix-like structures and  $\beta$ -sheet aggregates. Electron microscopy indicates that the aggregates prepared from these acetonitrile-rich solutions show morphologies similar to our previously observed polyQ fibrils. These aggregates redissolve upon the addition of water! These are the first examples of reversible fibril formation. Our monomeric Q10 peptides clearly sample broad regions of their available conformational energy landscape. The work here develops molecular-level insight into monomeric Q10 conformations and investigates the activation barriers between different monomer states and their evolution into fibrils.



## INTRODUCTION

Expansions of genomic “CAG” codon repeats are associated with at least 10 neurodegenerative disorders, including Huntington’s disease.<sup>1</sup> CAG repeats encode for expanded polyglutamine (polyQ) tracts in proteins. These expanded polyQ tracts cause proteins to aggregate into amyloid-like fibrils, which are the pathological hallmarks of CAG repeat diseases. The penetrance and severity of these diseases correlate with the length of the polyQ repeat expansion. In Huntington’s disease,<sup>2,3</sup> for example, repeats between 17 and 30 glutamines in the huntingtin protein are generally considered benign, whereas repeat lengths that exceed 36 residues typically result in fibrils and the manifestation of disease symptoms.

Although flanking sequences are known to influence aggregation kinetics and mechanisms,<sup>4,5</sup> the expanded polyQ protein segments are the only apparent commonality shared by CAG repeat diseases. For this reason, investigating the structures of polyQ peptides enables the molecular understanding of how glutamine repeats cause protein misfolding and aggregation. Indeed, previous studies show that aggregates prepared *in vitro* from model polyQ peptides share many characteristic features with pathologically relevant amyloid-like fibrils, including displaying filamentous morphologies, binding Thioflavin-T, and exhibiting  $\beta$ -sheet-rich structures.<sup>6,7</sup>

Many experimental studies suggest that soluble polyQ peptides are structurally disordered in aqueous solution.<sup>6–10</sup> Other studies, however, suggest that polyQ peptides contain small populations of “folded” structures, which are hypothesized to be putative cytotoxic agents.<sup>11</sup> Computational investigations<sup>12–15</sup> support the experimental finding that polyQ peptides, regardless of repeat length, are largely disordered, with only transient elements of regular secondary structures such as  $\alpha$ -helices,  $\beta$ -sheets, and turns.

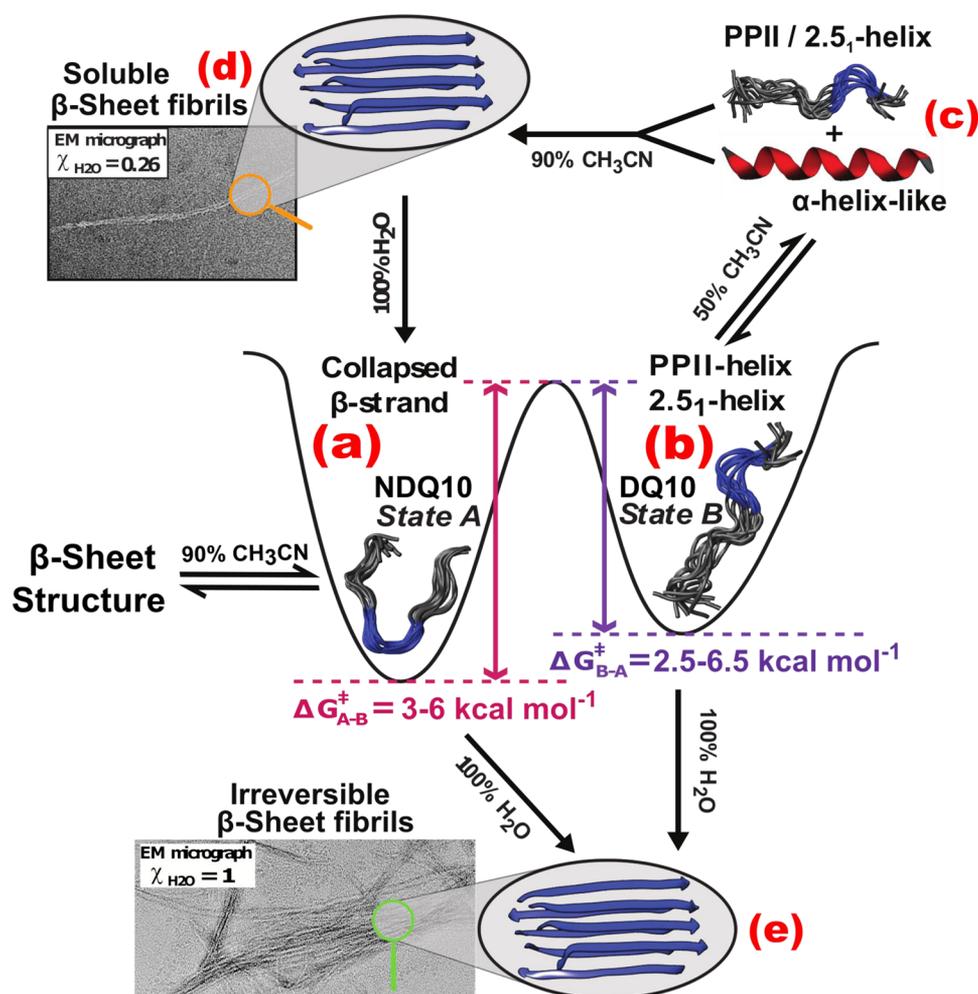
Although polyQ peptides are structurally disordered, they do not behave as true random coil polymers.<sup>2</sup> Several studies<sup>12,13,16–20</sup> suggest that the end-to-end distances of polyQ peptides deviate significantly from random coils, since they adopt relatively collapsed, globule-like structures. In addition, detailed analyses of circular dichroism (CD) and nuclear magnetic resonance (NMR) spectra<sup>9</sup> show that polyQ peptides possess a high propensity toward adopting polyproline II-like (PPII-like) secondary structures, although not necessarily forming long, continuous segments.

Despite numerous studies, several important aspects of polyQ peptide structure in solution remain poorly understood.

Received: May 1, 2017

Revised: May 15, 2017

Published: May 22, 2017



**Figure 1.** Summary of polyQ solution energy landscapes determined in this study. (a) and (b) show the potential wells for the aqueous Q10  $\beta$ -strand and PPII states observed. Conformation (c) depicts the  $\alpha$ -helix-like conformations observed in  $\sim 50\%$  acetonitrile/water solutions. Conformation (d) derives from Q10  $\beta$ -sheet fibrils formed in high acetonitrile concentrations. These fibrils redissolve in high water content acetonitrile/water solutions. Conformation (e) derives from irreversibly formed Q10  $\beta$ -sheet fibrils grown in water.

For example, little is known about the solution-state structural change(s) of polyQ peptides that lead to fibril nucleation and growth. One model, by Wetzel and co-workers,<sup>6</sup> proposes that fibril nucleation is initiated by an energetically unfavorable structural conversion from “random” coil to  $\beta$ -sheet. An alternative model, proposed by Pappu and co-workers,<sup>13</sup> argues that polyQ peptides adopt disordered globule structures, which nonspecifically aggregate into high molecular weight oligomers that subsequently convert into  $\beta$ -sheet-rich fibrils.

A detailed understanding of the conformational energy landscape that controls the solution-state structural propensities of polyQ peptides would yield valuable insights that could resolve some of the current debates in the field. Unfortunately, standard biophysical methods only achieve limited insight into the solution state structure of disordered proteins and peptides, such as polyQ.<sup>21,22</sup> Incisive progress requires the use of new experimental and computational biophysical tools.

In the work here, we investigate the ensemble of solution-state structures and the conformational energy landscape of the model peptide system,  $D_2Q_{10}K_2$  (Q10), by synergistically utilizing UVRR spectroscopy and MD simulations. This work builds upon our previous study by Xiong et al.<sup>23</sup> Figure 1 schematically summarizes the new insights that we develop in this work regarding the conformational energy landscape of

Q10. We find that Q10 can be poised to exist in two distinct, monomeric conformational states in aqueous solution. One of these states is in a collapsed  $\beta$ -strand-like structure (Figure 1a) that readily aggregates into amyloid-like fibrils (Figure 1e). The other state is a predominately PPII-like structure (Figure 1b) that resists aggregation.

The collapsed  $\beta$ -strand and PPII-like structures do not readily interconvert, which indicates that there is a large energy barrier that prevents these two states from equilibrating in solution. Using metadynamics, we simulate the conformational energy landscape of Q10 and investigate the origin of this high energy barrier. Overall, we find excellent, quantitative agreement between the simulation results and our UVRR data. This enables us to use our experimentally validated simulations to gain detailed, molecular-level structural insights into the conformational ensemble of Q10 in solution.

We also investigate the impact of low hydrogen bonding and dielectric environments on Q10 peptide structure. We measure the UVRR spectra of Q10 in different mixtures of acetonitrile and water. We find that at moderate concentrations of acetonitrile PPII-like conformations convert into structures composed of predominately  $\alpha$ -helix-like and turn structures (Figure 1c), while high concentrations of acetonitrile induce the formation of  $\beta$ -sheet fibrils (Figure 1d) and nonfibrillar

aggregates. Surprisingly, fibrils prepared from high acetonitrile concentrations redissolve in water-rich solutions.

## EXPERIMENTAL SECTION

**Materials.** The 14-residue D<sub>2</sub>Q<sub>10</sub>K<sub>2</sub> (Q10) peptide was purchased as a lyophilized powder from CPC Scientific at ≥95% purity. Trifluoroacetic acid (TFA) and 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) were purchased at ≥99% purity from Acros and Fluka, respectively. Far-UV grade acetonitrile (CH<sub>3</sub>CN) was purchased from Acros. Optima-grade H<sub>2</sub>O was purchased from Fisher Scientific, and D<sub>2</sub>O (99.9 atom % D) was purchased from Cambridge Isotope Laboratories, Inc.

**Sample Preparation.** Non-disaggregated Q10 (NDQ10) and disaggregated Q10 (DQ10) samples were prepared as described previously by Xiong et al.<sup>23</sup> NDQ10 solutions were prepared by directly dissolving the peptide in pure water. DQ10 solutions were prepared by utilizing a disaggregation protocol based on the one developed by Wetzel and co-workers.<sup>24</sup> Briefly, DQ10 samples were prepared by suspending the lyophilized peptide powder in a 1:1 (v/v) mixture of TFA and HFIP. The samples were sonicated for 20 min and incubated at room temperature for 2 h. The solvents were evaporated under a gentle stream of dry N<sub>2</sub> gas for ~30 min. The peptide film was dissolved in 1 mL of water. DQ10 peptide solutions were then diluted to their final concentrations with water. Raising the pH of monomeric NDQ10 increases its propensity to aggregate. To avoid this, we investigated both NDQ10 and DQ10 solutions at pH values that ranged between pH 2 and pH 3. The NDQ10 monomer solutions were also found to be very sensitive to impurities that nucleated aggregation into light scattering particles, which we presumed were dominated by β-sheet conformations. We scrupulously cleaned all vials used for sample preparation to remove nucleating centers.

**UV Resonance Raman (UVRR) Spectroscopy.** The UVRR instrumentation used is described in detail by Bykov et al.<sup>25</sup> The ~204 nm UV light was generated by Raman shifting the third harmonic of a Nd:YAG Infinity laser (Coherent, Inc.) in H<sub>2</sub> gas (~30 psi) and selecting the fifth anti-Stokes line. The laser light was focused onto a spinning Suprasil quartz NMR tube containing the samples. An ~165° backscattering geometry was used. The scattered light was dispersed using a home-built subtractive double monochromator and detected with a liquid N<sub>2</sub> cooled, back-thinned Spec10:400B CCD camera (Princeton Instruments) with a Lumogen E coating. The spectrometer resolution was ~5 cm<sup>-1</sup>. A detailed description of the spectral processing and fitting is given in the [Supporting Information](#).

**NMR Diffusion Measurements.** All <sup>1</sup>H diffusion NMR spectra were acquired with an 11.7 T magnet with a Bruker AVANCE III 500 console (Bruker Biopsin, Billerica, MA) at 298 K. NDQ10 and DQ10 solutions were prepared at 1 mg mL<sup>-1</sup> in 90% D<sub>2</sub>O. Spectra were recorded with a stimulated echo bipolar pulsed field gradient pulse sequence. The WATERGATE pulse sequence was used for water suppression.<sup>26</sup> The maximum strength of the gradient coil was 0.5 T m<sup>-1</sup> and was varied from 5 to 95% during the measurements. The decay of the NMR signal integrated area, *I*, as a function of the applied gradient strength, *G*, is described by the Stejskal–Tanner equation:<sup>27</sup>

$$I = I_0 e^{-[(\gamma G \delta)^2 (\Delta - (\tau/2) - (\delta/8)) D]} \quad (1)$$

where *I*<sub>0</sub> is the initial intensity,  $\gamma$  is the gyromagnetic ratio of <sup>1</sup>H,  $\delta$  is the length of the gradient pulse (2.5 ms),  $\Delta$  is the diffusion time (50 ms),  $\tau$  is the time between bipolar gradient pulses, and *D* is the apparent diffusion coefficient.

**Transmission Electron Microscopy (TEM).** Quantifoil 400 mesh copper grids were glow discharged using an EmiTech Glow Discharge unit for 70 s at 25 mA. Aliquots of 4 μL Q10 solutions were placed onto the freshly prepared glow discharged grids, rinsed with double distilled H<sub>2</sub>O, and stained with 4% (w/v) uranyl acetate. Samples were imaged in a FEI Tecnai 20F electron microscope. Images were collected using a Gatan Ultrascan 4K CCD camera at 30 000× and 50 000× magnification.

## COMPUTATIONAL METHODS

**General Simulation Details.** Metadynamics simulations were used to explore the Q10 conformational landscape. Constrained MD simulations were used to investigate the stabilities of the 2.5<sub>1</sub>-helix and PPII structures. The different peptide conformations were all constructed with the Molecular Operating Environment (MOE 2013.10) software suite.<sup>28</sup> The initial, solvated Q10 peptides were energy minimized for 10 000 steps using the conjugate gradient method. All peptide models were solvated with 6681 TIP3P water molecules in a periodic box with dimensions of 60 × 60 × 60 Å<sup>3</sup>. The NAMD software package (version 2.11)<sup>29</sup> was used for all MD simulations. The trajectory data were analyzed using VMD 1.9.2 and native Tcl scripting.<sup>70</sup> All MD simulations were run at the TACC supercomputer facility, while Q10 peptide models were prepared and energy minimized on a 4-core iMac desktop.

Potential energies were calculated with the CHARMM36 force field.<sup>30</sup> CHARMM36 contains the CMAP<sup>31</sup> dihedral corrections intended to decrease α-helix bias and stabilize β-strand secondary structure. The particle mesh Ewald algorithm<sup>32</sup> was used with a grid spacing of 1.0 Å to calculate full system electrostatics. An integration time step of 2 fs was used. All simulations were carried out under NPT conditions, with a Langevin thermostat and piston to regulate the temperature at 300 K and pressure at 1.013 25 bar.<sup>71</sup> The pair list distance was 14.0 Å, the pair interaction cutoff was 12.0 Å, and the switch distance was 10.0 Å.

**Metadynamics Simulation Details.** The initial Q10 monomer used in the metadynamics simulation was constructed by using a fully extended peptide with Ramachandran (Φ, Ψ) dihedral angles of 180°. The initial, solvated Q10 peptide systems were energy minimized for 10 000 steps using the conjugate gradient method followed by 50 ps of equilibration. The data used to construct the free energy landscape were collected over the next 1.0 μs of simulation time.

To specify the Q10 conformational space, we defined three collective variables that monitored the root-mean-square deviation (RMSD) of a Q10 structure relative to the idealized α-helix, β-hairpin, and PPII reference structures. The α-helix and PPII reference structures were created with MOE, whereas the β-hairpin reference structure was taken from preliminary metadynamics simulation results (coordinates provided in the [Supporting Information](#)). The RMSD was measured using only the peptide α-carbon positions. The maximum RMSD coordinate was set to 12.0 Å from the reference structure. This upper limit was maintained by applying a 1.0 kcal mol<sup>-1</sup> Å<sup>-1</sup> half-harmonic potential constraint. This three-dimensional

conformational energy landscape approach has been successfully employed by us previously.<sup>33–35</sup>

In our metadynamics simulations, artificial potentials in the form of Gaussian functions were added every 500 steps. The height of added Gaussian functions started at 1.0 kcal mol<sup>-1</sup>, and their width was 3.0 Å. Well-tempered metadynamics<sup>36</sup> was used to streamline sampling of the conformational energy landscape. This protocol is a variation of the original metadynamics algorithm developed by Laio and Parrinello<sup>37</sup> that utilized a gradual decrease in the height of added Gaussians throughout the simulation. This allowed shallow energy wells to be filled in quickly, while deeper energy wells filled more slowly. This method allows the deeper energy wells to be explored in finer detail during the simulation data collection run.

Coordinates for low-energy structures were extracted from the trajectory based on RMSD values and the Gaussian deposition history. Hydrogen bonding analysis of structures found in low-energy wells was performed with VMD's Hydrogen Bond analysis module. The  $\Psi$  and  $\chi_3$  dihedral angles were extracted from the simulation trajectories with a Tcl script, which is available for download from the Supporting Information. A heavy atom distance of <3.5 Å and a bond angle of  $180^\circ \pm 30^\circ$  was used to define a hydrogen bond. Metadynamics free energy data and collective variable history were processed with two respective Python scripts, which can be found in the Supporting Information. Script 1 identifies the energy landscape positions for structures of interest from the potential of mean force (pmf) file. Script 2 searches the metadynamics trajectory and identifies the structures that match the RMSD values found using script 1.

**Constrained MD Simulation Details.** We used constrained MD simulations to investigate the energies of Q10 in PPII and 2.5<sub>1</sub>-helix conformations constructed with ( $\Phi$ ,  $\Psi$ ) angles of  $(-75^\circ, 150^\circ)$  and  $(-130^\circ, 177^\circ)$ , respectively. In the constrained MD simulations, harmonic constraints were placed on the  $\Phi$  and  $\Psi$  dihedral angles for the Q10 peptides. These constraints ensured that the 2.5<sub>1</sub>-helix and PPII peptides would retain their respective secondary structures throughout the simulations. A force constant of 0.25 kcal mol<sup>-1</sup> Å<sup>-1</sup> was used for each dihedral constraint, and the constraints were applied with the Collective Variables module in NAMD.

For the constrained MD simulations the solvated Q10 peptides were energy minimized for 10 000 steps using the conjugate gradient method. The peptide systems were equilibrated for 1 ns, followed by data collection runs of 200 ns. Peptide potential energies were calculated using the NAMD Energy module. Bond, angles, dihedral angles, van der Waals interaction energies, and electrostatic energies were calculated for every frame of the trajectory output files. Energies were averaged using a simple Python script. The hydrogen bonding analysis of these constrained MD simulations was also performed with VMD's Hydrogen Bond module. The hydrogen bond parameters were the same as those described in the Metadynamics section above.

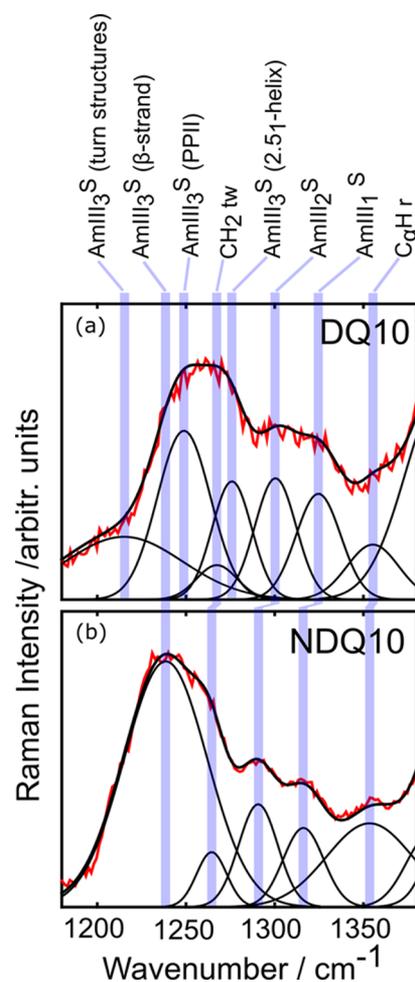
## RESULTS AND DISCUSSION

To maximize clarity, we include numbered bold headings that highlight the most important polyQ structural discoveries made here that are visually summarized in Figure 1.

**Solution-State Structures of Q10 in H<sub>2</sub>O.** Our previous study by Xiong et al.<sup>23</sup> examined the solution-state conformations of Q10 and suggested that it exists in two different

solution-state conformations that depend on the sample preparation method. Preparing Q10 using a standard “disaggregation” protocol<sup>24</sup> (forms a solution containing the peptide denoted as DQ10; see Experimental Section for details) results in an ensemble of predominately PPII-like peptide conformations. In contrast, non-disaggregated Q10 (denoted NDQ10) dissolved in water adopts a distinctly different conformation that Xiong et al.<sup>23</sup> proposed was a mainly  $\beta$ -hairpin conformation.

We remeasured the UVRR spectra of DQ10 and NDQ10. Figure 2 shows representative ~204 nm excited UVRR spectra



**Figure 2.** Representative 204 nm UVRR spectra and assignments of the AmIII<sup>S</sup> region for (a) DQ10 and (b) NDQ10 peptide solutions at 0.3 mg mL<sup>-1</sup> concentrations. Spectra from three independent replicates were globally fit using a minimum sum of Gaussian and Lorentzian bands, as described in the Supporting Information.

of the DQ10 and NDQ10 peptides dissolved in H<sub>2</sub>O. The DQ10 UVRR spectra are essentially identical to those measured previously by Xiong et al.,<sup>23</sup> who investigated DQ10 peptide solutions at pH 7. In contrast, the NDQ10 UVRR spectra (Figure S7) differ significantly from that of Xiong et al.<sup>23</sup> It appears that Xiong et al.'s<sup>23</sup> NDQ10 samples had begun to aggregate. The sensitivity of NDQ10 monomers to nucleation of aggregation may have caused these spectral differences. Thus, we have altered our assignment of the NDQ10 solution monomer conformation.

**1.0 UVRR Shows That DQ10 Populates PPII and 2.5<sub>1</sub>-Helix Conformations, While NDQ10 Populates  $\beta$ -Strand-like Conformations.** UVRR excitation at  $\sim 200$  nm occurs within the NV<sub>1</sub> electronic transitions of the secondary and primary amides of the peptide backbone and glutamine side chains.<sup>23</sup> Thus, the Figure 2  $\sim 204$  nm excitation UVRR spectra of polyQ peptides are dominated mainly by resonance enhanced bands that derive from primary and secondary amide (Am) vibrations, which we denote with the superscripts P and S, respectively.

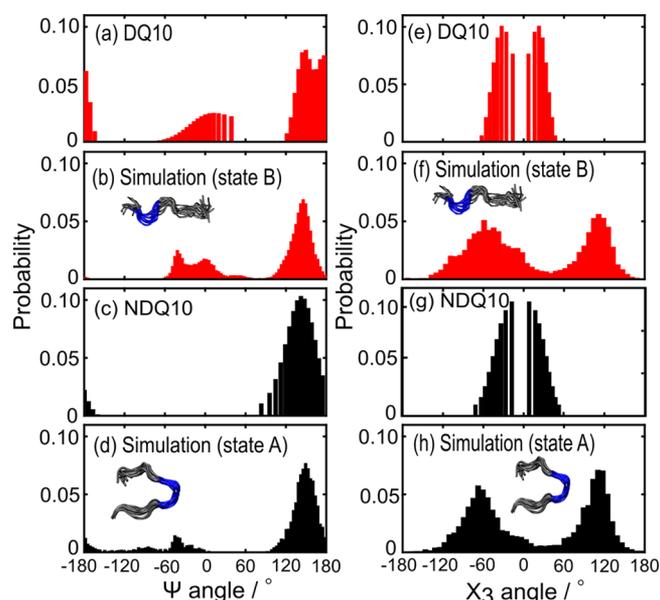
The Figure 2 spectra show the structurally sensitive extended AmIII<sup>S</sup> region between 1180 and 1380 cm<sup>-1</sup>.<sup>42</sup> The DQ10 and NDQ10 spectra were curve-fit using a new, more rigorous and self-consistent global analysis, as described in the Supporting Information (Figure S1). Both the DQ10 and NDQ10 UVRR spectra show bands at  $\sim 1355$ ,  $\sim 1320$ , and  $\sim 1295$  cm<sup>-1</sup>. The 1355 cm<sup>-1</sup> features are assigned to a mainly C $\alpha$ -H rocking mode based on our previous normal-mode analyses of glutamine.<sup>41</sup> The peaks at 1320 and 1295 cm<sup>-1</sup> are the AmIII<sup>S</sup><sub>1</sub> and AmIII<sup>S</sup><sub>2</sub> bands, respectively.<sup>40,43-45</sup> However, we note that the AmIII<sup>S</sup><sub>1</sub> and AmIII<sup>S</sup><sub>2</sub> overlap with additional weaker bands that are due to CH<sub>2</sub> twisting modes of the glutamine side chains.

The AmIII<sup>S</sup><sub>3</sub> region of DQ10 (Figure 2a) is similar to that of Xiong et al. with bands at  $\sim 1275$ ,  $\sim 1265$ ,  $\sim 1250$ , and  $\sim 1215$  cm<sup>-1</sup>. In contrast, the AmIII<sup>S</sup><sub>3</sub> region of NDQ10 contains bands at  $\sim 1240$  and  $\sim 1265$  cm<sup>-1</sup>, which significantly differ from spectra reported by Xiong et al.<sup>23</sup> The difference in the NDQ10 solution state spectra is most likely due to some aggregation of Xiong et al.'s NDQ10 sample, since their spectrum is more similar to that of our NDQ10 fibrils.<sup>46</sup> Additionally, our recent insights into the UVRR spectra of glutamine show that the  $\sim 1265$  cm<sup>-1</sup> band derives from CH<sub>2</sub> wagging of the glutamine side chain instead of an AmIII<sup>S</sup><sub>3</sub> vibration.<sup>41</sup> As discussed below, our new understanding of the UVRR spectrum of NDQ10 modifies our assignment of the NDQ10 solution conformation.

**Ramachandran  $\Psi$  Angle Distributions of DQ10 and NDQ10.** The AmIII<sup>S</sup><sub>3</sub> band is the most structurally sensitive spectroscopic marker of the peptide backbone conformation because its frequency depends on the Ramachandran  $\Psi$  dihedral angle.<sup>44,45</sup> As discussed in detail by Asher and co-workers,<sup>45</sup> this sinusoidal frequency dependence derives from the coupling of the C $\alpha$ -H bending vibration with the N-H bending component of the AmIII<sup>S</sup><sub>3</sub>. This coupling is strong for  $\beta$ -strand and PPII-like structures, where the C $\alpha$ -H and N-H groups are approximately in a cis-configuration, which downshifts the AmIII<sup>S</sup><sub>3</sub> frequency. In contrast, for  $\alpha$ -helical structures, the C $\alpha$ -H and N-H groups are in a trans configuration. This decouples the C $\alpha$ -H and N-H bending motions, which results in an upshift of the AmIII<sup>S</sup><sub>3</sub> band frequency.

We utilized the structural sensitivity of the AmIII<sup>S</sup><sub>3</sub> band to estimate the Ramachandran  $\Psi$  angle distributions for DQ10 and NDQ10 in solution. To do this, we employed the methodology of Mikhonin et al.<sup>44</sup> (see Supporting Information for details), which correlates the frequencies of the AmIII<sup>S</sup><sub>3</sub> band envelopes to the peptide bond  $\Psi$  angles. This enables us to roughly estimate the probability distribution of peptide bond  $\Psi$  angles that derive from the inhomogeneously broadened AmIII<sup>S</sup><sub>3</sub> band shape.

Figure 3a,c shows the  $\Psi$  angle distributions of DQ10 and NDQ10. The  $\Psi$  angle distribution of DQ10 is essentially



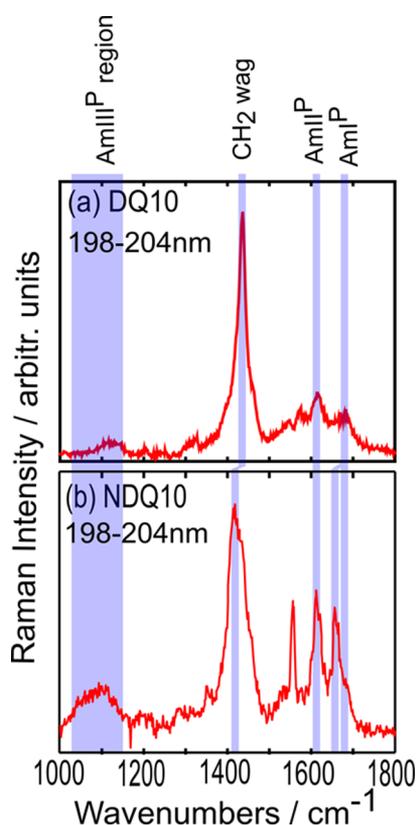
**Figure 3.** Comparison of  $\Psi$  and  $\chi_3$  angle distributions measured by UVRR and calculated by MD. Comparison between DQ10  $\Psi$  angle distributions measured by (a) UVRR and (b) MD structures found in state B. Comparison between NDQ10  $\Psi$  angle distributions measured by (c) UVRR and (d) MD structures found in state A. Comparison between DQ10  $\chi_3$  angle distributions measured by (e) UVRR and (f) MD structures found in state B. Comparison between NDQ10  $\chi_3$  angle distributions measured by (g) UVRR and (h) MD structures found in state A.

identical to that reported by Xiong et al.,<sup>23</sup> consisting of three prominent peaks located at  $\sim 10^\circ$ ,  $\sim 150^\circ$ , and  $\sim 175^\circ$ . The peak located at  $150^\circ$  is indicative of PPII-like conformations, while the peak centered at  $10^\circ$  indicates the presence of turn-like structures. The peak centered at  $175^\circ$  suggests that a significant fraction of DQ10 peptide bonds adopt 2.5<sub>1</sub>-helix-like conformations.

The NDQ10  $\Psi$  angle distribution (Figure 3c) shows a peak located at  $\sim 140^\circ$  indicative of  $\beta$ -strand conformations. In contrast to the NDQ10  $\Psi$  angle distribution of Xiong et al., we do not observe a peak at  $\psi \sim -40^\circ$  that they assigned to turn structures.

**Hydrogen Bonding Environment of DQ10 and NDQ10 Side Chains.** Our previous study of primary amide vibrations shows that the AmI<sup>P</sup> and AmII<sup>P</sup> bands report on the hydrogen bonding and dielectric environments of the glutamine side chain C=O and NH<sub>2</sub> groups, respectively.<sup>47</sup> Xiong et al. showed that  $\sim 198$  nm excitation enhances the primary amide UVRR bands significantly more than the secondary amide bands.<sup>23</sup> As a result, the primary amide bands can be selectively studied by calculating the difference spectrum between  $\sim 198$  and  $\sim 204$  nm excitation UVRR spectra.

Figure 4 shows the 198–204 nm UVRR difference spectra of DQ10 and NDQ10. The AmI<sup>P</sup> (mostly C=O stretching) and AmII<sup>P</sup> (NH<sub>2</sub> scissoring) bands of DQ10 are located at  $\sim 1680$  and  $\sim 1614$  cm<sup>-1</sup>, respectively. In NDQ10, the AmI<sup>P</sup> frequency downshifts to  $\sim 1660$  cm<sup>-1</sup>, although a prominent shoulder remains at  $\sim 1680$  cm<sup>-1</sup>. As with DQ10, the AmII<sup>P</sup> band of NDQ10 is located at  $\sim 1614$  cm<sup>-1</sup>. The most intense bands in the 198–204 nm UVRR difference spectra appear at  $\sim 1437$  cm<sup>-1</sup> for DQ10 and  $\sim 1414$  cm<sup>-1</sup> for NDQ10. We assign these bands to a complex vibration that contains CH<sub>2</sub> wagging as well as significant contributions of C–C stretching, CH<sub>2</sub> scissoring,



**Figure 4.** UVRR 198–204 nm difference spectra for (a) DQ10 reported by Xiong et al.<sup>23</sup> and (b) NDQ10 reported in this work. Panel a was adapted with permission from ref 23.

and side chain C–N stretching. The C–N stretching component of this vibration likely contributes to its UVRR enhancement.

As shown in Table 1, the AmI<sup>p</sup> and AmII<sup>p</sup> band frequencies of DQ10 are similar to those of monomeric glutamine in water.

**Table 1. Glutamine, DQ10, and NDQ10 Primary Amide UVRR Band Frequencies in Water (cm<sup>-1</sup>)<sup>a</sup>**

vibration	glutamine (cm <sup>-1</sup> )	DQ10 (cm <sup>-1</sup> )	NDQ10 (cm <sup>-1</sup> )
AmI <sup>p</sup>	1679	1680	1680 (sh), 1655
AmII <sup>p</sup>	1620	1614	1614
CH <sub>2</sub> wag	1427	1437	1432 (sh), 1414
AmIII <sup>p</sup>	1110	1099	1106

<sup>a</sup>sh: shoulder.

This indicates that the primary amide C=O and NH<sub>2</sub> groups of DQ10 are predominately hydrogen bonded to water. In contrast, the AmI<sup>p</sup> band of NDQ10 is downshifted to ~1660 cm<sup>-1</sup>, suggesting that a significant population of side chain C=O groups are involved in intrapeptide hydrogen bonding. Additionally, the shoulder at ~1680 cm<sup>-1</sup> suggests that there is also a population of NDQ10 side chain C=O groups that hydrogen bond to water molecules.

In contrast, the AmII<sup>p</sup> band frequency of NDQ10 (~1614 cm<sup>-1</sup>) indicates that its side chain NH<sub>2</sub> groups predominately hydrogen bond to water. Thus, we conclude that NDQ10 side chain C=O moieties must be hydrogen bonded to the NH groups of the peptide backbone. This conclusion is supported by a recently published report by Walsh et al.,<sup>73</sup> who showed that small, gas phase glutamine dipeptides with a  $\beta$ -strand-like

conformation preferentially form hydrogen bonds between side chain C=O groups and backbone NH groups.

**Glutamine Side Chain  $\chi_3$  Angle Distributions of DQ10 and NDQ10.** We recently discovered a side chain UVRR primary amide vibration that shows a cosinusoidal frequency dependence on the O–C–C dihedral angles of glutamine and asparagine side chains (designated as the  $\chi_3$  and  $\chi_2$  dihedral angles, respectively).<sup>41</sup> This vibration, which we call the AmIII<sup>p</sup>, is reminiscent of the AmIII<sup>s</sup> vibration because it derives from an in-phase combination of side chain C–N stretching and NH<sub>2</sub> rocking motions (instead of peptide backbone C–N stretching and N–H bending of the AmIII<sup>s</sup>). It should be noted, however, that C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>  stretching also contributes to the AmIII<sup>p</sup> vibration.

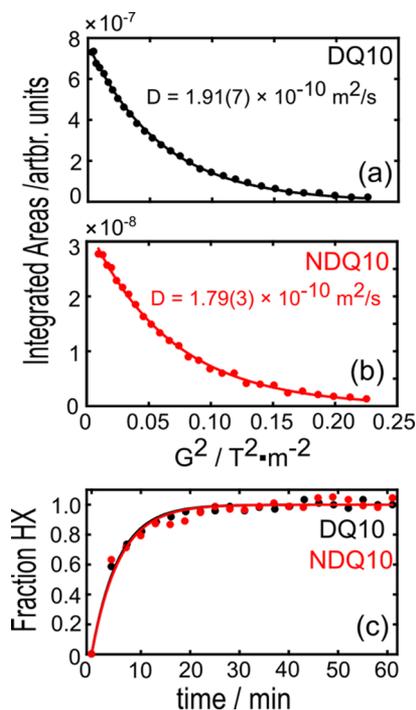
The structural sensitivity of the AmIII<sup>p</sup> derives from hyperconjugation between the side chain C=O  $\pi^*$  and C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>   $\sigma$  orbitals.<sup>41</sup> When hyperconjugation is strong (e.g., at  $\chi_3 \sim \pm 90^\circ$ ), electron density is transferred from the C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>  to the C=O bond. This elongates the C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>  bond and reduces the C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>  stretching force constant. As a result, the AmIII<sup>p</sup> frequency downshifts. In the absence of hyperconjugation (e.g., at  $\chi_3 \sim 0^\circ$ ), the C <sub>$\beta$</sub> –C <sub>$\gamma$</sub>  bond length shortens and the AmIII<sup>p</sup> frequency upshifts.

We examined the AmIII<sup>p</sup> bands of the Figure 4 UVRR difference spectra of NDQ10 and DQ10 to determine their  $\chi_3$  angle distributions (see the Supporting Information for details). As discussed previously by Punihale et al.,<sup>41</sup> the AmIII<sup>p</sup> frequency of DQ10 at ~1099 cm<sup>-1</sup> correlates to a  $\chi_3$  distribution centered at ~-30° and/or ~20° (Figure 3e). This result is similar to that of aqueous glutamine.<sup>41</sup> For NDQ10 (Figure 3g), the AmIII<sup>p</sup> band peaks at ~1106 cm<sup>-1</sup>, which correlates to  $\chi_3$  angle distributions that are centered at ~-12° and/or ~3°.

**2.0 DQ10 and NDQ10 Are Monomers in Aqueous Solution.** We used diffusion-ordered NMR spectroscopy (DOSY) to measure the diffusion coefficients of DQ10 and NDQ10. Figures 5a and 5b show representative signal decays of NMR peaks corresponding to Gln side chain N<sup>-1</sup>H resonance for DQ10 and NDQ10, respectively. We fit these decays to eq 1 and obtained diffusion coefficients of  $1.91(7) \times 10^{-10}$  and  $1.79(3) \times 10^{-10}$  m<sup>2</sup> s for DQ10 and NDQ10, respectively. Assuming a sphere, we roughly estimate that both coefficients are consistent with diffusion of monomeric peptides with Stokes radii of ~13–14 Å. The fact that DQ10 and NDQ10 are monomers in solution is also supported by our UVRR hydrogen–deuterium exchange (HX) measurements (see Supporting Information for details). As shown in Figure 5c, the HX exchange times of DQ10 and NDQ10 are essentially identical, ~5.5 min, consistent with other monomeric peptides.<sup>76</sup>

**3.0 Qualitative Evidence of a Large Activation Barrier between the DQ10 and the NDQ10 Solution Monomer Conformations.** We found that the solution conformations of DQ10 and NDQ10 do not readily interconvert. Additionally, Xiong et al.<sup>23</sup> showed that the soft, possibly crystalline solids, formed upon evaporating DQ10, retain their solution-state conformation. These observations suggest that the NDQ10 and DQ10 solution conformations lie in separate deep energy wells. The high activation barrier between NDQ10 and DQ10 effectively prevents conformational transitions and equilibration between their respective structures.

**4.0 ~2-Fold Decreased H-Bonding Solvent Transforms DQ10 from a PII-like to an  $\alpha$ -Helix-like**



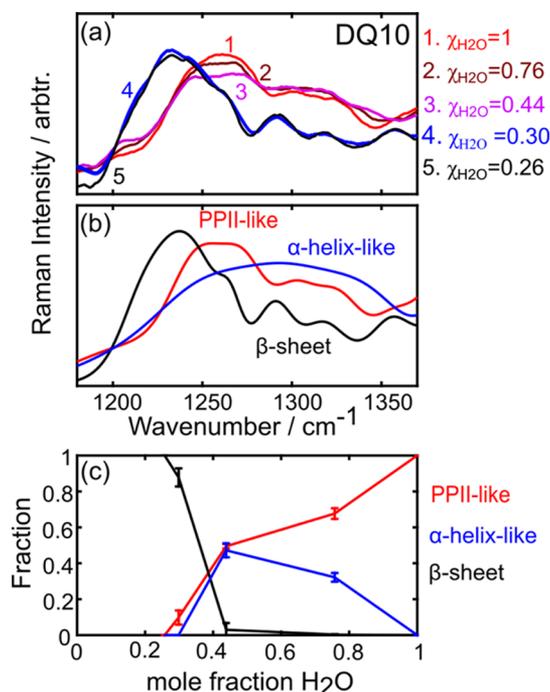
**Figure 5.** DOSY NMR and HX UVRR data indicate that DQ10 and NDQ10 are monomers in solution. Representative DOSY NMR signal decays for (a) DQ10 and (b) NDQ10. The measured diffusion coefficients for DQ10 and NDQ10 are  $1.91(7) \times 10^{-10}$  and  $1.79(3) \times 10^{-10} \text{ m}^2/\text{s}$ , respectively. (c) HX kinetics of DQ10 and NDQ10 are similar, showing time constants of  $\sim 5.5$  min, which are similar to those of monomeric peptides of similar molecular weights.

### Conformation, while a >3-Fold Decreased H-Bonding Solvent Transforms DQ10 to $\beta$ -Sheet Fibril Conformations.

We investigated the conformational transitions of DQ10 in less hydrogen bonding, lower dielectric constant acetonitrile/water solvent mixtures. The addition of acetonitrile is expected to promote inter- and intramolecular peptide hydrogen bonding by reducing the solvent's ability to hydrogen bond.<sup>48</sup> Figure 6a shows the UVRR spectra of DQ10 in a series of water/acetonitrile mixtures. Compared to DQ10 in pure water, the relative intensity of the AmIII<sub>3</sub> band decreases as the mole fraction of water decreases to  $X_{\text{H}_2\text{O}} = 0.44$ . At mole fractions below  $X_{\text{H}_2\text{O}} = 0.44$ , the AmIII<sub>3</sub> band downshifts to  $\sim 1230 \text{ cm}^{-1}$  and increases significantly in relative intensity. The dramatic spectral changes that occur at low mole fractions of water are accompanied by the formation of large, visually evident aggregates.

The spectra shown in Figure 6a were modeled as a linear combination of basis spectra to determine the relative fraction of different secondary structures. We found that the spectra could be satisfactorily modeled using a linear combination of PPII-like,  $\alpha$ -helix-like, and  $\beta$ -sheet basis spectra (shown in Figure 6b). Our modeling procedure and methods used to calculate the basis spectra are discussed in detail in the Supporting Information (Figures S1 and S2).

Figure 6c shows the relative fractions of the different secondary structure conformations of DQ10 as a function of the acetonitrile/water solvent composition, while Figure 7 shows the consequent  $\Psi$  angle distributions. The UVRR data indicate that the DQ10 peptide bonds adopt both PPII-like and  $\alpha$ -helix-like conformations (Figure 1c) at intermediate water/



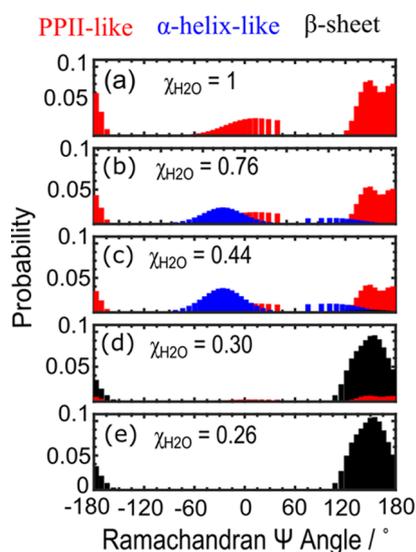
**Figure 6.** Dependence of the 204 nm UVRR spectra of DQ10 on the  $\text{H}_2\text{O}/\text{CD}_3\text{CN}$  solvent composition. (a) Dependence of the amide III band on solvent composition. The spectra were smoothed using a Savitzky-Golay filter with a second-order polynomial over a 15-point spectral window. (b) Calculated Q10  $\beta$ -sheet (black), PPII-like (red), and  $\alpha$ -helix-like (blue) UVRR basis spectra (see Figure S2 for details). (c) Relative fractions of PPII-like (red),  $\alpha$ -helix-like (blue), and  $\beta$ -sheet (black) secondary structures as a function of  $X_{\text{H}_2\text{O}}$ . The error bars were determined from three independent experiments.

acetonitrile mole fractions ( $X_{\text{H}_2\text{O}} = 0.76$  or  $X_{\text{H}_2\text{O}} = 0.42$ ), while at lower mole fractions of water ( $X_{\text{H}_2\text{O}} = 0.24$ ) DQ10 forms  $\beta$ -sheet-rich aggregates (Figure 1d). These data agree with our measured CD spectra shown in Figure S3.

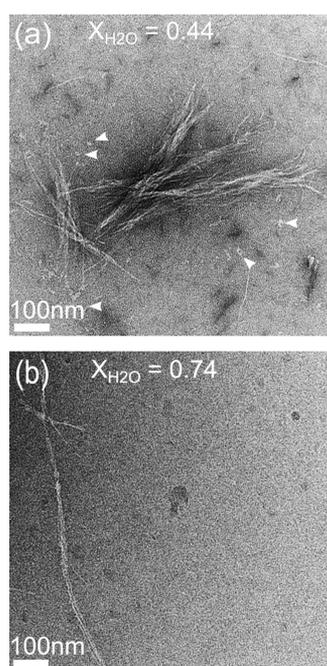
The distributions shown in Figure 7b,c that derive from peptide bonds in “ $\alpha$ -helix-like” conformations (shown in blue) possess two broad populations of  $\Psi$  angles centered at  $-25^\circ$  and  $100^\circ$ . The distributions indicate that the peptide bonds do not derive from canonical  $\alpha$ -helix conformations. Instead, they suggest that a large fraction of these structures likely derive from  $3_{10}$ -like helices (indicated by the distribution peak at  $-25^\circ$ ), interspersed by turn structures (indicated by the peak located at  $\sim 100^\circ$ ). However, the breadth of these distributions indicates that the DQ10 peptide bonds also populate  $\Psi$  angles indicative of more canonical  $\alpha$ -helix conformations as well as small populations of  $\pi$ -helices.

In addition to dominant  $\alpha$ -helix-like and PPII-like conformations, a small population of peptide bonds adopt  $\beta$ -sheet-like structures in acetonitrile/water mixtures where  $X_{\text{H}_2\text{O}} = 0.44$ . However, at lower water mole fractions ( $X_{\text{H}_2\text{O}} = 0.30$ ), the PPII-like and  $\alpha$ -helix-like structures convert almost entirely into  $\beta$ -sheet structures (Figures 6c and 7d,e). The formation of these  $\beta$ -sheet structures coincides with the appearance of visually evident aggregates.

TEM images (Figure 8) indicate that these  $\beta$ -sheet aggregates consist of oligomers and mature amyloid-like fibrils. At  $X_{\text{H}_2\text{O}} = 0.44$  (Figure 8a), we observe both fibrillar and oligomeric aggregates; however, at  $X_{\text{H}_2\text{O}} = 0.26$  (Figure 8b), we



**Figure 7.** Ramachandran  $\Psi$  angle distributions of DQ10 in different  $\text{H}_2\text{O}/\text{CD}_3\text{CN}$  mixtures. The distributions shown in red correspond to PPII-like structures, in blue to  $\alpha$ -helix-like structures, and in black to  $\beta$ -sheet structures. The  $\Psi$  angle distributions were determined from the Figure 6b basis spectra (see Supporting Information for details) and scaled according to the relative fractions of PPII-like,  $\alpha$ -helix-like, and  $\beta$ -sheet structures reported in Figure 6c.



**Figure 8.** Representative TEM images of DQ10 aggregates prepared in (a)  $X_{\text{H}_2\text{O}} = 0.44$  solution and (b)  $X_{\text{H}_2\text{O}} = 0.26$  solution. The white arrows point to nonfibrillar aggregates.

observe essentially only fibrils. Replacement of water with increasing amounts of acetonitrile destabilizes PPII-like structures because the solvent is less competitive for hydrogen bonding to the backbone peptide bonds. As a result, Q10 adopts  $\alpha$ -helix-like and turn conformations, which can satisfy hydrogen bonding through intramolecular interactions. However, at high concentrations of acetonitrile, hydrogen bonding must involve interpeptide chain hydrogen bonding, thus forming aggregates and  $\beta$ -sheet-rich fibrils.

The morphologies and UVRR spectra of the DQ10 fibrils prepared at low mole fractions of water are similar to fibrils prepared by Punihaole et al.<sup>46</sup> in aqueous NDQ10 and DQ10 monomer solutions at 60 °C and pH 7 (Figure 1e). Punihaole et al. found that the fibrils prepared from both DQ10 and NDQ10 peptide solutions in this manner were composed of extended  $\beta$ -strands that predominately assembled into antiparallel  $\beta$ -sheets, although a minor population of parallel  $\beta$ -sheets were also observed. Here, we also observe that  $\sim 95\%$  of the peptide bonds adopt antiparallel  $\beta$ -sheet conformations in fibrils prepared in acetonitrile/water mixtures.

### 5.0 Fibrils Aggregated in Low H-Bonding Solvent Redissolve in Water into the NDQ10 Monomer Solution Conformation.

In contrast to fibrils prepared from DQ10 monomers by Punihaole et al.,<sup>46</sup> those prepared from the acetonitrile solutions redissolve in water (Figure 1a,d). These redissolved aggregates show a UVRR spectrum identical to the solution-state, monomeric NDQ10 conformations (see Figure S4). This result significantly differs from fibrils formed via incubation in water, which irreversibly form fibrils.

We investigated the side chain and backbone amide hydrogen bonding in DQ10 fibrils grown in acetonitrile by examining their 197–204 nm and 204–(197–204) nm UVRR difference spectra respectively (Figures S8 and S9). We observe an  $\text{AmI}^{\text{S}}$  frequency at  $\sim 1670 \text{ cm}^{-1}$  and an  $\text{AmI}^{\text{P}}$  frequency at  $\sim 1663 \text{ cm}^{-1}$ . In contrast, DQ10 fibrils grown in water have  $\text{AmI}^{\text{P}}$  and  $\text{AmI}^{\text{S}}$  frequencies of 1664 and 1661  $\text{cm}^{-1}$ , respectively.<sup>46</sup> The  $\text{AmIII}^{\text{S}_3}$  band of fibrils prepared in acetonitrile is essentially identical to that grown in water.<sup>46</sup> This suggests that the peptide backbone structures of these fibrils are essentially the same. Thus, we conclude that the  $\text{AmI}^{\text{P}}$  and  $\text{AmI}^{\text{S}}$  frequency differences must be attributed to changes in hydrogen bonding of the side chain and backbone amide groups, respectively. Based on this conclusion, our data suggest that compared to fibrils grown in aqueous solution, fibrils prepared from DQ10 in acetonitrile have similar side chain C=O and  $\text{NH}_2$  hydrogen bonding, but much weaker peptide backbone hydrogen bonding. The weakened backbone hydrogen bonding interactions may destabilize the fibril structure, causing fibrils grown in acetonitrile to dissolve when exposed to more water.

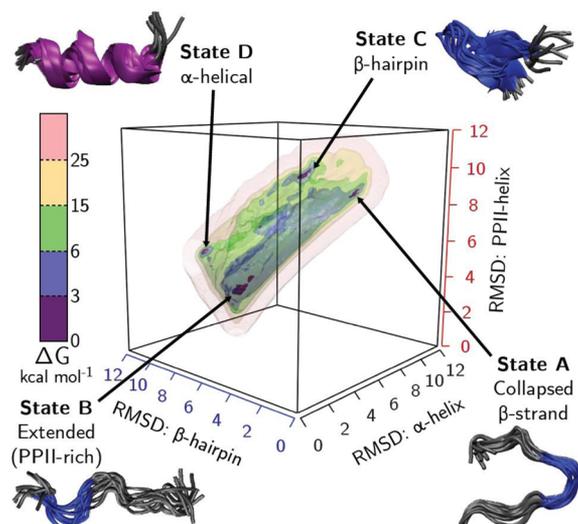
Interestingly, DQ10 fibrils grown in  $\sim 90\%$  acetonitrile and redissolved in 50% acetonitrile/50% water show CD spectral signatures characteristic of NDQ10 collapsed  $\beta$ -strands, with no indication of  $\alpha$ -helix-like conformations (Figure S5). At intermediate acetonitrile concentrations, DQ10 PPII-like structures are in equilibrium with  $\alpha$ -helix-like conformations. These  $\alpha$ -helix-like structures convert back into PPII-like conformations upon further dilution with water (Figure S5).

At high acetonitrile mole fractions ( $\chi_{\text{H}_2\text{O}} = 0.26$ ), NDQ10 shows a Raman spectrum characteristic of  $\beta$ -sheet fibrils; however, unlike DQ10, these solutions are not turbid. Upon centrifugation ( $\sim 20\,000\times g$ ), solutions of NDQ10 with high acetonitrile mole fractions do not form pellets. Thus, it appears that NDQ10 solutions with high acetonitrile mole fractions do not form high molecular weight aggregates. Copious water addition to NDQ10 in 90% acetonitrile results in a conformational transition to its original collapsed  $\beta$ -strand structure (Figure S10).

**Metadynamics Simulations of Q10.** We used metadynamics MD simulations to obtain insights into the conformational equilibrium and the high activation barrier between the

DQ10 and NDQ10 monomer forms. Metadynamics simulations are ideal for characterizing the conformational landscapes of complex biophysical systems.<sup>54,55</sup> In metadynamics, the evolution of the simulated system is biased by adding a history-dependent potential energy that discourages revisiting events already sampled and accelerates the sampling of rare events. As a consequence of the algorithm, metadynamics can also be utilized to estimate the conformational free energy landscape of a peptide or protein as a function of a suitably chosen set of reaction coordinates called collective variables.

Figure 9 shows the conformational Gibbs free energy landscape of Q10 as a function of three collective variables,

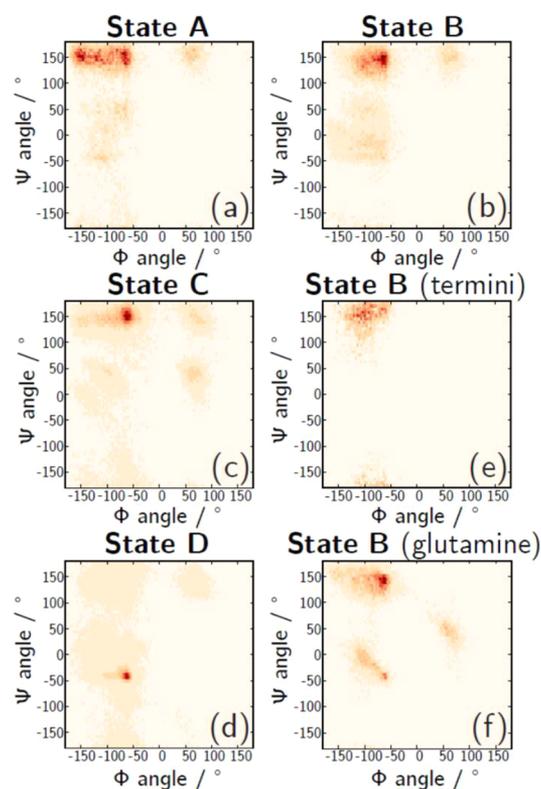


**Figure 9.** Conformational energy landscape of Q10 obtained from metadynamics. The four lowest energy states are represented by the violet isosurfaces with representative structural ensembles shown for these low-energy states.

defined in terms of the root-mean-square-deviations (RMSD) from idealized  $\beta$ -hairpin,  $\alpha$ -helix, and PPII structures. We chose these three reference structures because they are common secondary structure motifs that enable us to clearly distinguish between different possible polyQ conformations.<sup>56–58</sup> We chose a  $\beta$ -hairpin structure because it was previously implicated as an aggregation nucleus.<sup>59</sup>

We find that the conformational energy landscape of Q10 is frustrated because it contains numerous shallow, local minima. However, four deep energy minima (purple regions in Figure 9) are observed. These energy minima contain different structural states (A–D) that are defined both by their Ramachandran angles (Figure 10a–d) as well as their hydrogen bonding propensities (Figure 11).

The structures found in state A possess peptide bonds with average  $(\Phi, \Psi)$  angles of  $(-90^\circ, 140^\circ)$ , roughly consistent with  $\beta$ -strand-like conformations (Figure 10a). These conformations are best described as “collapsed”  $\beta$ -strands, rather than  $\beta$ -hairpin, because they lack the intramolecular peptide backbone hydrogen bonding that defines canonical  $\beta$ -hairpins. As shown in Figure 11d, these  $\beta$ -strand structures form, on average, more side chain–backbone hydrogen bonds than the other states. These side chain–backbone hydrogen bonds, together with the attractive electrostatic interactions between the terminal aspartic acid and lysine residues, cause the  $\beta$ -strands to adopt more compact (i.e., collapsed) conformations.

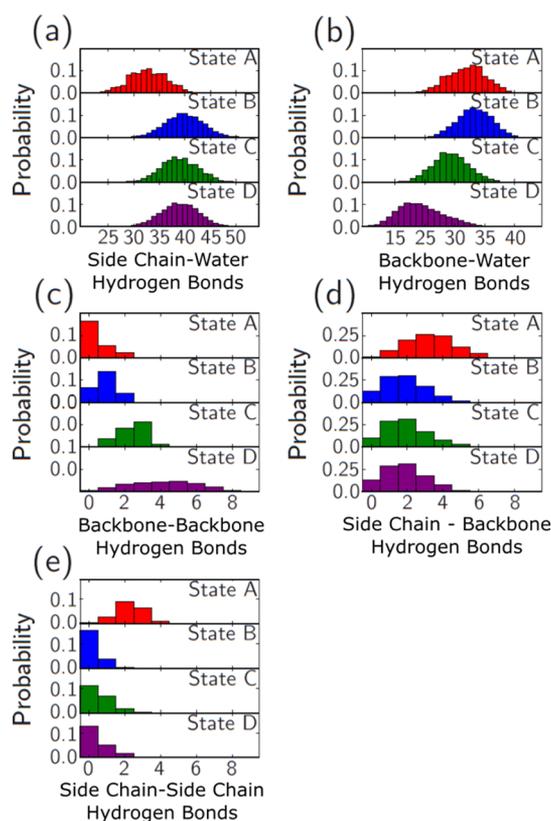


**Figure 10.** Ramachandran  $(\Phi, \Psi)$  angle distributions shown for the (a) collapsed  $\beta$ -strand structures of state A, (b) PPII-like structures of state B, (c)  $\beta$ -hairpin-like structures of state C, and (d)  $\alpha$ -helical structures of state D. Panel (e) shows the  $(\Phi, \Psi)$  angles of the terminal aspartic acid and lysine residues of state B while panel (f) shows  $(\Phi, \Psi)$  angles of the glutamine residues for state B.

The  $(\Phi, \Psi)$  angle distributions of state B are centered at about  $(-80^\circ, -150^\circ)$ , which is characteristic of PPII secondary structures (Figure 10b). However, a smaller population of Ramachandran angles, located at approximately  $(-50^\circ, -57^\circ)$ , indicate that state B also contains turn-like conformations. Taken together, these two populations of Ramachandran angles indicate that state B consists of irregular structures that are characterized by short PPII helices interspersed by turn conformations or bends. In addition, our hydrogen bonding analysis (Figure 11) indicates that compared to collapsed  $\beta$ -strand structures (state A), both the peptide backbone and side chains of PPII-like structures (state B) are preferentially hydrogen bonded to water.

The structures found in state C possess peptide bonds whose  $(\Phi, \Psi)$  angles are centered at  $(-135^\circ, 135^\circ)$ , indicative of  $\beta$ -strand structures (Figure 10c). A smaller population of peptide bonds adopts Type I' turn Ramachandran angles  $(-40^\circ, -85^\circ)$ . In contrast to the state A collapsed  $\beta$ -strand structures, the state C  $\beta$ -hairpin structures are calculated to form more hydrogen bonds between backbone peptide bonds (Figure 11c). These backbone–backbone hydrogen bonds are between  $\beta$ -strands and are characteristic of canonical  $\beta$ -hairpins defined by Milner et al.<sup>60</sup> In addition, the side chains of these  $\beta$ -hairpin structures are more solvated than those of the collapsed  $\beta$ -strand structures found in state A (Figure 11a).

The structural ensemble of state D is comprised mostly of  $\alpha$ -helical conformations. As shown in Figure 10e, the  $(\Phi, \Psi)$  angles are distributed narrowly around  $(-50^\circ, -65^\circ)$ . These structures show substantial amount of backbone–backbone



**Figure 11.** Number of interamide and amide–water hydrogen bonds formed for structures A–D in the metadynamics simulations. (a) Total number of glutamine side chain–water hydrogen bonds for each structure. (b) Total number of peptide backbone–water hydrogen bonds formed for each structure. (c) Total number of peptide backbone–backbone hydrogen bonds formed for each structure. (d) Total number of side chain–backbone hydrogen bonds for each structure. (e) Total number of side chain–side chain hydrogen bonds formed for each structure.

hydrogen bonding (Figure 11c), as expected for  $\alpha$ -helices, while the side chains are well solvated.

The collapsed  $\beta$ -strand (state A) occurs at the global minimum in the conformational energy landscape, while the extended PPII conformation (state B), the  $\beta$ -hairpin (state C), and the  $\alpha$ -helix (State D) occur at 0.5, 0.66, and 1.0 kcal mol<sup>-1</sup> higher in energy, respectively. A 3–6 kcal mol<sup>-1</sup> activation energy barrier separates states A and B. In contrast, states C and D are *both* isolated from each other and isolated from states A and B by a 6–15 kcal mol<sup>-1</sup> energy barrier.

**Consistency of MD Simulation and UVRr Results.** Our interpretation of the Ramachandran  $\Psi$  and  $\Phi$  angle distributions for the simulated states A–D (Figure 10a–d) indicate that the state A (collapsed  $\beta$ -strand) and state B (PPII-like) structures resemble the UVRr measured solution structures found for the NDQ10 and DQ10 monomers, respectively. To validate the simulation results, we compared the UVRr experimentally measured  $\Psi$  and  $\chi_3$  angles of DQ10 and NDQ10 to the corresponding distributions obtained from the MD structures populating the states A and B energy wells (Figure 3). Overall, agreement between the experimentally measured  $\Psi$  angle distributions and those obtained from the simulations is excellent. In contrast, the side chain  $\chi_3$  angle distributions obtained from the MD simulation show poor agreement with the UVRr data.

Figures 3a–d compare the  $\Psi$  angle distributions of DQ10 and NDQ10 with those of states A and B from the simulations. The corresponding experimentally measured and computationally derived  $\Psi$  angle distributions are in overall good agreement. The only major discrepancy is that the calculated distribution of state B (Figure 3b) lacks a peak at  $\sim 175^\circ$  that corresponds to 2.5<sub>1</sub>-helix conformations. This apparent discrepancy is discussed in detail below.

Metadynamics show that the  $\chi_3$  distributions for states A and B are bimodal (Figure 3f,h). Both distributions show peaks centered at  $\sim -70^\circ$  and  $\sim 110^\circ$ . These  $\chi_3$  angle distributions differ significantly from those measured experimentally by UVRr. An analysis of a side chain rotamer database by Dunbrack and Shapovalov<sup>41,61</sup> shows that glutamine residues typically possess  $\chi_3$  angles that are distributed between  $\sim -90^\circ$  and  $90^\circ$ , which is inconsistent with the distributions shown in Figure 3f,h.

We conclude that the large differences between our UVRr measurements and our MD simulated  $\chi_3$  angle distributions derive from a lack of parametrization for nonrotameric side chain dihedral angles in the CHARMM36 force field. It is known, for example, that the  $\chi_1$  and  $\chi_2$  dihedral angles of amino acid side chains have been optimized in CHARMM36 using experimental data.<sup>62</sup> However, the nonrotameric dihedral angle, such as the  $\chi_3$  angle of the glutamine side chain, has not been similarly treated in CHARMM36.

We compared the backbone and side chain amide hydrogen bonding environments observed by UVRr measurements (NDQ10 and DQ10) to those calculated from the metadynamics calculated structures of states A and B (Figure 11). Our computational results are in excellent agreement with our experimental findings. The state B structures of the metadynamics contain side chain and backbone peptide bonds predominantly hydrogen bonded to water, while the state A structures show significant interamide hydrogen bonding.

The metadynamics data validate the observation of a large activation barrier between the NDQ10 PPII (state B) and the DQ10  $\beta$ -strand (state A) conformations. We calculate that the barrier between states A ( $\beta$ -strand) and B (PPII) is 3–6  $k_B T$ . In addition, the simulated MD energy barrier region between states A and B also contains  $\geq 8$  local energy minima, with energy well depths of  $\sim 1.0$  kcal mol<sup>-1</sup> (Figure S6). These local energy minima define metastable states along the energy barrier region. Even if Q10 can overcome the large energy barrier, the kinetics of conformation equilibration will be slowed by these multiple wells to prevent PPII-like to  $\beta$ -strand structural transitions.

**Origin of Q10 2.5<sub>1</sub>-Helix Conformation Stabilization.** Our DQ10 UVRr studies surprisingly detect a significant population of the 2.5<sub>1</sub>-helix conformation previously proposed by Krimm and co-workers<sup>39,63,64</sup> and observed for PLL and PGA peptides by Asher and co-workers.<sup>43</sup> The 2.5<sub>1</sub>-helix conformation consists of an extended, left-handed,  $\beta$ -strand-like structure with  $\sim 2.5$  residues per helical turn and ( $\Phi$ ,  $\Psi$ ) angles of roughly ( $-130^\circ$ ,  $170^\circ$ ).

Asher and Krimm's work indicate that the 2.5<sub>1</sub>-helix conformations in PLL and PGA derive mainly from the electrostatic repulsions between their charged amino acid side chains forcing them to splay out.<sup>43</sup> Based on this understanding, it is surprising that a predominately polyQ-rich peptide such as Q10 (which only has four charged residues) would show a significant population of peptide bonds with 2.5<sub>1</sub>-

Table 2. Potential Energies (kcal mol<sup>-1</sup>) for Q10 Peptides in PPII-like and 2.5<sub>1</sub>-Helix Conformations

energy terms	PPII		PPII/2.5 <sub>1</sub> -helix		2.5 <sub>1</sub> -helix	
	all residues	glutamine residues	all residues	glutamine residues	all residues	glutamine residues
bonds	44(5)	33(5)	44(5)	32(5)	43(5)	32(5)
angles	134(10)	96(8)	134(10)	95(8)	134(10)	96(8)
torsions	146(5)	107(4)	160(5)	121(4)	174(6)	131(4)
van der Waals	-14(4)	-11(3)	-16(4)	-14(4)	-15(4)	-14(4)
electrostatics	-436(29)	-318(10)	-461(35)	-331(8)	-487(37)	-330(9)

helix-like  $\Psi$  angles (Figure 3a). Another disquieting fact is that our MD simulations of the state B structures (Figure 3b) appear unimodal, consisting of  $\Psi$  angles indicative of the PPII conformation. This lack of predicted 2.5<sub>1</sub>-helix conformations is troubling because it could signal misinterpretations of the UVR data.

We used classical MD simulations to more deeply examine the factors that stabilize 2.5<sub>1</sub>-helix conformations in Q10. We simulated three different Q10 peptide structures. One structure has its backbone dihedral angles constrained to the canonical 2.5<sub>1</sub>-helix conformation, the second structure has a PPII conformation, and the third structure has alternating PPII/2.5<sub>1</sub>-helix dihedral Ramachandran angles. We identified the overall lowest energy structure of each of the constrained simulations. We then calculated, for each Q10 structure, the energy contributions of bond stretches, bond angle bending, dihedral angle rotation, electrostatics, and van der Waals to the overall potential energy of the system (Table 2). Neglecting the influence of solvating waters, our simulations indicate that the potential energy of the 2.5<sub>1</sub>-helix is  $\sim 50$  kcal mol<sup>-1</sup> lower than that of the PPII structure, mainly due to intrapeptide electrostatics.

To determine the electrostatic potential energy contributions of the charged terminal residues, we compared the electrostatic energy terms of just the glutamine residues of Q10 for the 2.5<sub>1</sub>-helix and PPII structures. We find that for glutamine residues the electrostatic potential energy difference between the 2.5<sub>1</sub>-helix and PPII structures is only  $\sim 10$  kcal mol<sup>-1</sup>. Thus, the charged terminal residues contribute most heavily to the electrostatic potential energy difference between the PPII and 2.5<sub>1</sub>-helix structures.

We also compared the ( $\Phi$ ,  $\Psi$ ) angle distributions of the glutamine residues and the charged flanking residues for state B structures. As shown in Figure 10e, the terminal aspartic acid and lysine residues preferentially adopt ( $\Phi$ ,  $\Psi$ ) angles centered roughly at ( $-115^\circ$ ,  $-160^\circ$ ), closer to the ( $\Phi$ ,  $\Psi$ ) angles of the canonical 2.5<sub>1</sub>-helix conformation. In contrast, the glutamine residues (Figure 10f) adopt ( $\Phi$ ,  $\Psi$ ) angles of ( $-80^\circ$ ,  $-150^\circ$ ) and ( $-50^\circ$ ,  $-57^\circ$ ), indicative of PPII and turn-like conformations, respectively. This suggests that the 2.5<sub>1</sub>-helix-like conformations are selectively localized to the peptide bonds of the charged terminal residues, while the glutamine residues occur in a predominately PPII conformation.

Thus, the UVR DQ10 results appear consistent with the MD simulations. Assuming equal Raman cross sections, we roughly estimate from the integrated intensities of the AmIII<sub>3</sub> bands (Figure 2a) that 43% of peptide bonds adopt PPII-like  $\Psi$  angles, 22% of peptide bonds adopt 2.5<sub>1</sub>-helix  $\Psi$  angles, and 35% of the peptide bonds adopt turn structure  $\Psi$  angles. On the basis of these percentages, we calculate that a single DQ10 peptide has 2–3 peptide bonds involved in a 2.5<sub>1</sub>-helix conformation. These 2.5<sub>1</sub>-helix conformation peptide bonds most likely occur on the terminal charged residues. However,

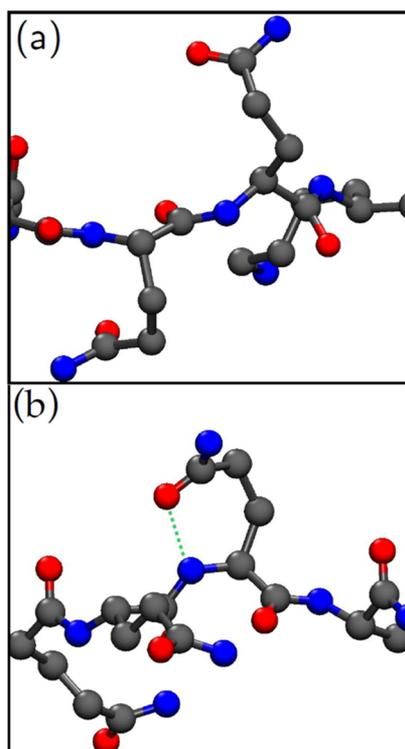
the MD simulations indicate that the state B 2.5<sub>1</sub>-helix-like ( $\Phi$ ,  $\Psi$ ) angles are centered around  $-160^\circ$ , very close to that of canonical PPII structures centered at  $-150^\circ$ . This significantly differs from the  $\Psi = -175^\circ$  value measured for DQ10 by UVR and causes the calculated  $\Psi$  angle distribution of state B (Figure 3b) to appear unimodal.

We attribute the discrepancy in the  $\Psi$  angle values for the 2.5<sub>1</sub>-helix-like peptide bonds in the simulated state B structures to an inadequacy of the CHARMM36 force field. In the case of CHARMM36, parameters for backbone dihedral angles are optimized from globular protein X-ray crystal structures and QM data for the dialanine peptide.<sup>65</sup> The 2.5<sub>1</sub>-helix is an uncommon structure that forms in the presence of adjacent charged amino acids and is unlikely to be well represented in the X-ray structures used to parametrize CHARMM36. This conclusion is supported by Liqi Feng, who showed that metadynamics of poly(L-lysine), using the CHARMM36 force field, fail to show a  $\Psi$  angle distribution at  $\sim 170^\circ$ , indicative of 2.5<sub>1</sub>-helix conformations.<sup>75</sup>

**Comparisons to Other Studies.** The solution-state structures of polyQ-rich peptides and proteins have been studied in detail.<sup>12,14,16,17,67,68</sup> Many of the important computational studies were conducted by Pappu's group.<sup>12,14,17</sup> For example, Wang et al.<sup>12</sup> characterized the conformational ensemble of *N*-acetyl-Q<sub>5</sub>-*N'*-methylamide (Q5) and *N*-acetyl-Q<sub>15</sub>-*N'*-methylamide (Q15) that are somewhat similar to our Q10 peptide.

From their MD simulations, Wang et al.<sup>12</sup> conclude that the conformational energy landscapes of Q5 and Q15 are frustrated, since they observe numerous metastable and glassy states. Their simulations find essentially no global energy minimum structures. They show that Q5 and Q15 are structurally disordered in aqueous solution, with only transient stretches of regular secondary structure elements, such as PPII-helices,  $\beta$ -strands, and  $\alpha$ -helices occurring. They also show that both Q5 and Q15 peptide bonds have a strong propensity to adopt PPII-like and  $\alpha$ -helix-like ( $\Phi$ ,  $\Psi$ ) angles.

Wang et al.'s<sup>12</sup> simulations also suggest that structurally disordered, monomeric polyQ peptides form a significant number of interamide hydrogen bonds, with side chain–backbone hydrogen bonding being the most prevalent. This is ostensibly supported by the NMR measurements by Darnell et al.<sup>72</sup> that find nuclear Overhauser effect (NOE) cross-peaks between the C <sub>$\gamma$</sub>  protons of glutamine side chains and C <sub>$\alpha$</sub>  protons of the backbone in R<sub>3</sub>GQ<sub>3</sub>GY and R<sub>3</sub>GQ<sub>3</sub>P<sub>11</sub>GY peptides that adopt predominately PPII structures. Darnell et al. interpret these signals to mean that in PPII-rich polyQ peptides the glutamine side chains adopt “folded-over” conformations (see Figure 12) that enable hydrogen bonding to the backbone amides. To our knowledge, this is the only experimental evidence of significant side chain–backbone hydrogen bonding in PPII-rich polyQ peptides.



**Figure 12.** Q10 in a PPII conformation with (a) extended, solvated side chains and (b) side chains hydrogen bonded to the peptide backbone as proposed by Darnell et al.<sup>72</sup> The structure shown in panel (a) is of the PPII state (state B) found in our metadynamics simulations and validated by our UVRr measurements of DQ10. The structure shown in panel (b) was created using a Q10 peptide with Ramachandran angles constrained to PPII values in which the side chain structure was modified to adopt side chain–backbone hydrogen bonds between the C=O and NH moieties of the side chains and backbone, respectively. We clearly do not experimentally observe significant side chain–backbone hydrogen bonding in PPII-like DQ10 peptide monomers.

The simulations of Wang et al.<sup>12</sup> lead them to hypothesize that the structural disorder of polyQ peptides stems from the many different possible combinations of intramolecular and intermolecular–amide hydrogen bonds that can form between side chain and backbone amides. This explains why a  $\beta$ -sheet-rich fibril nucleus<sup>66,69</sup> is expected to be energetically unfavorable. According to their model, the multiple possible combinations of intrapeptide hydrogen bonds promotes disorder in the peptide backbone, disrupting the formation of backbone–backbone interactions that enable secondary structures such as  $\alpha$ -helices and  $\beta$ -sheets.

Some aspects of our simulation results are consistent with those of Wang et al.<sup>12</sup> We observe that the conformational energy landscape of Q10 contains numerous shallow energy minima. Similarly, we observe that Q10 peptides can engage in a large variety of intramolecular side chain–side chain, backbone–backbone, and side chain–backbone hydrogen bonding interactions.

However, there are many aspects of our results that are inconsistent with those of Wang et al.<sup>12</sup> For example, our experimental and metadynamics simulation data reveal that Q10 can adopt stable and well-defined structural states, with deep energy wells, such as those shown in Figure 1. Interestingly, we also find that the side chains and backbone amides of the structurally “disordered” PPII-rich state of Q10

are predominately hydrogen bonded to water, which disagrees with the Wang et al.<sup>12</sup> results.

These disagreements raise two important questions. First, why does Wang et al.<sup>12</sup> not observe *any* well-defined structural states for small polyQ peptides, whereas we do? And second, why do the simulations of Wang et al.<sup>12</sup> indicate that there are significant numbers of interamide hydrogen bonds (particularly between the backbone and side chain amides) in structurally disordered Q5 and Q15 peptides, while our simulations do not find this for Q10?

To answer the first question, we note that the conformational energy landscape of polyQ peptides is frustrated, consisting of many local energy wells. Classical MD simulations generally do not efficiently sample different structural states in these situations. Thus, one possibility is that the simulated peptides in the Wang et al. study are trapped in local energy minima. Our metadynamics calculations enable us to robustly sample a greater ensemble of structures than does traditional MD simulations. Thus, we characterized the entire conformational landscape and discovered global minimum energy conformations.

We are also aware that the conformational behavior of model polyQ peptides in solution depends on both the number of glutamine repeats and the choice of non-glutamine flanking groups.<sup>77</sup> For example, the use of *N*-acetyl and methylamide flanking groups in Pappu and co-workers’ simulations are better models for polyQ tracts in proteins. In addition, fluorescence correlation spectroscopy<sup>16</sup> data by the Pappu group and dynamic light scattering data by the Murphy group<sup>19</sup> indicate that structurally disordered long polyQ peptides adopt relatively compact structures in aqueous solution. These compact structures presumably derive from the fact that interamide hydrogen bonding interactions become more prevalent and important in polyQ peptides as the glutamine repeat length increases. For example, Walters and Murphy<sup>19</sup> have shown that phosphate buffer is a good solvent for Q8 and Q12, a theta solvent for Q16, and a poor solvent for larger peptides such as Q20. Thus, the model proposed by Pappu and co-workers in the Wang et al.<sup>12</sup> study may be valid in the limit of polyQ peptides with  $\geq 20$  glutamine repeats.

The answer to the second question may lie in the differences of the force field and water model used by our study compared to Wang et al.<sup>12</sup> As discussed by Wang et al.,<sup>12</sup> the hydrogen bonding interactions observed in a simulation will vary depending on the force field utilized. Our use of the TIP3P water model and the CHARMM36 modern force field results in different hydrogen bonding interactions compared to that of the MD simulations of Wang et al.<sup>12</sup> The hydrogen bonding interactions predicted by our metadynamics are in excellent agreement with our UVRr experimental data, which show that our choice of force field appropriately models the hydrogen bonding interactions in Q10.

The congruence of both our simulation and experimental data gives us confidence in the relative accuracy and robustness of the CHARMM36 force field used in this study. Despite this, we do not dispute the validity of the Pappu et al.’s model to describe, in general, the underlying physical principles that govern the structural disorder of larger polyQ peptides or the energetic unfavorability of the coil to  $\beta$ -sheet transition. However, we do believe that the importance of side chain–side chain and side chain–backbone interactions in describing the structural disorder of PPII-rich polyQ peptides may be

overstated, at least in the context of small peptide systems, such as Q5, Q10, and Q15.

## CONCLUSIONS

We investigated the structural ensemble of the Q10 polyQ peptide in solution. From our UVRR, CD, and MD data, we constructed a molecular-level model for the structure and part of the energy landscape of DQ10 and NDQ10 monomers in aqueous solution. We show that in DQ10 the charged, terminal residues adopt  $2.5_1$ -helix conformations, while the central glutamine residues adopt short PPII segments that are interspersed with turn structures (Figure 1b). In contrast, we find that NDQ10 exists in a collapsed  $\beta$ -strand conformation (Figure 1a). There is a large activation barrier between these conformations in solution that prevents their interconversion.

Our study achieves new insights into the hydrogen bonding environments and glutamine side chain conformations of Q10. The DQ10 side chains are predominately hydrogen bonded to water, whereas the NDQ10 side chains form relatively more interamide hydrogen bonds. We were also able to determine that DQ10 possesses  $\chi_3$  angle distributions centered at  $\sim -30^\circ$  and/or  $\sim 20^\circ$ . In contrast, NDQ10 has a  $\chi_3$  angle distribution centered at  $\sim -12^\circ$  and/or  $\sim 3^\circ$ .

We also monitored the PPII to  $\beta$ -sheet structural transition of DQ10 in acetonitrile/water mixtures. Acetonitrile is a secondary structure inducing solvent because it promotes intra- and interpeptide hydrogen bonding. At intermediate acetonitrile concentrations, we observe a significant amount of  $\alpha$ -helix-like secondary structures in DQ10 (Figure 1c). At high acetonitrile concentrations, we observe a transition to  $\beta$ -sheet conformations and the formation of visible aggregate light scattering (Figure 1d). The higher acetonitrile concentrations promote intra- and interpeptide hydrogen bonding by first forming  $\alpha$ -helix-like structures and at higher acetonitrile concentrations forming  $\beta$ -sheet aggregates that include fibrils. The DQ10 fibrils and  $\beta$ -sheet aggregates formed in acetonitrile/water solutions redissolve upon addition of water. We find that the redissolved fibrils have a UVRR spectrum identical to monomeric NDQ10, indicating a collapsed  $\beta$ -strand structure. To our knowledge, this is the first report of redissolvable polyQ fibrils.

In contrast to DQ10, NDQ10 at high acetonitrile concentrations does not become turbid and does not form a pellet upon centrifugation. Thus, we conclude that NDQ10 does not form high molecular weight aggregates at high acetonitrile concentrations. However, the UVRR spectra of NDQ10 at high acetonitrile mole fractions are similar to that of fibrils and show a  $\Psi$  angle distribution characteristic of  $\beta$ -sheet. This NDQ10 collapsed  $\beta$ -strand to  $\beta$ -sheet transition upon the addition of acetonitrile is reversible.

We also used metadynamics to investigate the energy barriers between the collapsed  $\beta$ -strand and the DQ10 PPII/ $2.5_1$  helix conformation. We calculate a 3–6  $k_B T$  activation barrier with additional smaller minima along the reaction coordinate between these conformations (Figure 1a,b). This is consistent with our experimental observation that NDQ10 and DQ10 conformations do not readily interconvert.

Another important result of this study is that we rigorously benchmarked the CHARMM36 force field against experimental data and determined both the strengths and limitations of this force field's accuracy. We will continue to investigate and improve the utility of CHARMM36 in describing  $2.5_1$ -helix

structures as well as the nonrotameric  $\chi_3$  dihedral angles calculated for glutamine side chains.

Until now investigations of the aggregation mechanisms of polyQ peptides and proteins have been hindered by the inability to study intermediate polyQ species along the aggregation and fibril formation reaction coordinate. The fact that we can poise our Q10 peptide in different conformations and reversibly transition these conformations along reaction coordinates that form  $\beta$ -sheet aggregates, such as fibrils, makes us hopeful that this simple Q10 system will enable elucidation of fibril aggregation mechanisms. We will be examining polyQ structural changes in more physiologically relevant polyQ peptides in future studies.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.7b04060.

UVRR spectral processing and fitting, determination of  $\Psi$  and  $\chi_3$  angle distributions, hydrogen–deuterium exchange experiments, and CD measurements (PDF)

Metadynamics  $\beta$ -hairpin collective variable reference structure (PDB)

Metadynamics  $\alpha$ -helix collective variable reference structure (PDB)

Metadynamics PPII collective variable reference structure (PDB)

Python scripts used to process Metadynamics free energy data and collective variable history (ZIP)

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [madura@duq.edu](mailto:madura@duq.edu) (J.D.M.).

\*E-mail: [asher@pitt.edu](mailto:asher@pitt.edu) (S.A.A.).

### ORCID

Ryan S. Jakubek: 0000-0001-7880-9422

### Present Addresses

D.P.: Department of Chemistry, University of Minnesota, Minneapolis, MN 55455.

L.E.M.: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

### Author Contributions

D.P. and R.S.J. contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

Funding for this work was provided by the University of Pittsburgh (DP, RSJ, SAA), Defense Threat Reduction Agency HDTRA1-09-14-FRCWMD (RSJ, SAA), and partially supported by NIH R01 DA027806 (JDM and RJW). The MD simulation computer time was supported by XSEDE MCB060069, and computer equipment was purchased from NSF funds (CHE-1126465 and P116Z080180).

## REFERENCES

- Orr, H. T.; Zoghbi, H. Y. Trinucleotide Repeat Disorders. *Annu. Rev. Neurosci.* **2007**, *30*, 575–621.
- Murphy, R. M.; Walters, R. H.; Tobelmann, M. D.; Bernacki, J. P. When More Is Not Better: Expanded Polyglutamine Domains in Neurodegenerative Disease. In *Non-fibrillar Amyloidogenic Protein*

*Assemblies - Common Cytotoxins Underlying Degenerative Diseases*; Rahimi, F., Bitan, G., Eds.; Springer: Dordrecht, Netherlands, 2012; pp 337–375.

(3) Myers, R. H. Huntington's Disease Genetics. *NeuroRx* **2004**, *1*, 255–262.

(4) Bhattacharyya, A.; Thakur, A. K.; Chellgren, V. M.; Thiagarajan, G.; Williams, A. D.; Chellgren, B. W.; Creamer, T. P.; Wetzel, R. Oligoproline Effects on Polyglutamine Conformation and Aggregation. *J. Mol. Biol.* **2006**, *355*, 524–535.

(5) Thakur, A. K.; Jayaraman, M.; Mishra, R.; Thakur, M.; Chellgren, V. M.; Byeon, I.-J.; Anjum, D. H.; Kodali, R.; Creamer, T. P.; Conway, J. F.; et al. Polyglutamine Disruption of the Huntingtin Exon 1 N Terminus Triggers a Complex Aggregation Mechanism. *Nat. Struct. Mol. Biol.* **2009**, *16*, 380–389.

(6) Chen, S.; Berthelie, V.; Hamilton, J. B.; O'Nuallain, B.; Wetzel, R. Amyloid-like Features of Polyglutamine Aggregates and Their Assembly Kinetics. *Biochemistry* **2002**, *41*, 7391–7399.

(7) Wetzel, R. Physical Chemistry of Polyglutamine: Intriguing Tales of a Monotonous Sequence. *J. Mol. Biol.* **2012**, *421*, 466–490.

(8) Altschuler, E. L.; Hud, N. V.; Mazrimas, J. A.; Rupp, B. Random Coil Conformation for Extended Polyglutamine Stretches in Aqueous Soluble Monomeric Peptides. *J. Pept. Res.* **1997**, *50*, 73–75.

(9) Chellgren, B. W.; Miller, A.-F.; Creamer, T. P. Evidence for Polyproline II Helical Structure in Short Polyglutamine Tracts. *J. Mol. Biol.* **2006**, *361*, 362–371.

(10) Klein, F. A. C.; Pastore, A.; Masino, L.; Zeder-Lutz, G.; Nierengarten, H.; OuladAbdelghani, M.; Altschuh, D.; Mandel, J.-L.; Trotter, Y. Pathogenic and Nonpathogenic Polyglutamine Tracts Have Similar Structural Properties: Towards a Length-dependent Toxicity Gradient. *J. Mol. Biol.* **2007**, *371*, 235–244.

(11) Nagai, Y.; Inui, T.; Popiel, H. A.; Fujikake, N.; Hasegawa, K.; Urade, Y.; Goto, Y.; Naiki, H.; Toda, T. A Toxic Monomeric Conformer of the Polyglutamine Protein. *Nat. Struct. Mol. Biol.* **2007**, *14*, 332–340.

(12) Wang, X.; Vitalis, A.; Wyczalkowski, M. A.; Pappu, R. V. Characterizing the Conformational Ensemble of Monomeric Polyglutamine. *Proteins: Struct., Funct., Genet.* **2006**, *63*, 297–311.

(13) Vitalis, A.; Wang, X.; Pappu, R. V. Atomistic Simulations of the Effects of Polyglutamine Chain Length and Solvent Quality on Conformational Equilibria and Spontaneous Homodimerization. *J. Mol. Biol.* **2008**, *384*, 279–297.

(14) Vitalis, A.; Lyle, N.; Pappu, R. V. Thermodynamics of  $\beta$ -Sheet Formation in Polyglutamine. *Biophys. J.* **2009**, *97*, 303–311.

(15) Nakano, M.; Watanabe, H.; Rothstein, S. M.; Tanaka, S. Comparative Characterization of Short Monomeric Polyglutamine Peptides by Replica Exchange Molecular Dynamics Simulation. *J. Phys. Chem. B* **2010**, *114*, 7056–7061.

(16) Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence Correlation Spectroscopy Shows that Monomeric Polyglutamine Molecules form Collapsed Structures in Aqueous Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 16764–16769.

(17) Vitalis, A.; Wang, X.; Pappu, R. V. Quantitative Characterization of Intrinsic Disorder in Polyglutamine: Insights from Analysis Based on Polymer Theories. *Biophys. J.* **2007**, *93*, 1923–1937.

(18) Singh, V. R.; Lapidus, L. J. The Intrinsic Stiffness of Polyglutamine Peptides. *J. Phys. Chem. B* **2008**, *112*, 13172–13176.

(19) Walters, R. H.; Murphy, R. M. Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation. *J. Mol. Biol.* **2009**, *393*, 978–992.

(20) Vitalis, A.; Pappu, R. V. Assessing the Contribution of Heterogeneous Distributions of Oligomers to Aggregation Mechanisms of Polyglutamine Peptides. *Biophys. Chem.* **2011**, *159*, 14–23.

(21) Ladurner, A. G.; Itzhaki, L. S.; Daggett, V.; Fersht, A. R. Synergy between Simulation and Experiment in Describing the Energy Landscape of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 8473–8478.

(22) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. Characterization of the Freeenergy Landscapes of Proteins by NMR-

guided Metadynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6817–6822.

(23) Xiong, K.; Punihale, D.; Asher, S. A. UV Resonance Raman Spectroscopy Monitors Polyglutamine Backbone and Side Chain Hydrogen Bonding and Fibrillization. *Biochemistry* **2012**, *51*, 5822–5830.

(24) Chen, S.; Wetzel, R. Solubilization and Disaggregation of Polyglutamine Peptides. *Protein Sci.* **2001**, *10*, 887–891.

(25) Bykov, S.; Lednev, I.; Ianoul, A.; Mikhonin, A.; Munro, C.; Asher, S. A. Steady-State and Transient Ultraviolet Resonance Raman Spectrometer for the 193–270 nm Spectral Region. *Appl. Spectrosc.* **2005**, *59*, 1541–1552.

(26) Piotta, M.; Saudek, V.; Sklenár, V. Gradient-tailored Excitation for Single-quantum NMR Spectroscopy of Aqueous Solutions. *J. Biomol. NMR* **1992**, *2*, 661–665.

(27) Stejskal, E. O.; Tanner, J. E. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J. Chem. Phys.* **1965**, *42*, 288–292.

(28) Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2016.

(29) Phillips, J.; Braun, R.; et al. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(30) Huang, J.; Mackerell, A. D. CHARMM36 All-atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.

(31) Mackerell, A. D. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604.

(32) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(33) Ascitutto, E. K.; Mikhonin, A. V.; Asher, S. A.; Madura, J. D. Computational and Experimental Determination of the Alpha-helix Unfolding Reaction Coordinate. *Biochemistry* **2008**, *47*, 2046–2050.

(34) Ascitutto, E. K.; General, I. J.; Xiong, K.; Asher, S. A.; Madura, J. D. Sodium Perchlorate Effects on the Helical Stability of a Mainly Alanine Peptide. *Biophys. J.* **2010**, *98*, 186–196.

(35) Xiong, K.; Ascitutto, E. K.; Madura, J. D.; Asher, S. A. Salt Dependence of an  $\alpha$ -Helical Peptide Folding Energy Landscapes. *Biochemistry* **2009**, *48*, 10818–10826.

(36) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered Metadynamics: A Smoothly Converging and Tunable Free-energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.

(37) Laio, A.; Parrinello, M. Escaping Free-energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562–12566.

(38) O'Nuallain, B.; Thakur, A. K.; Williams, A. D.; Bhattacharyya, A. M.; Chen, S.; Thiagarajan, G.; Wetzel, R. Kinetics and Thermodynamics of Amyloid Assembly using a High-performance Liquid Chromatography-based Sedimentation Assay. *Methods Enzymol.* **2006**, *413*, 34–74.

(39) Tiffany, M.; Krimm, S. New Chain Conformations of Poly(glutamic acid) and Polylysine. *Biopolymers* **1968**, *6*, 1379–1382.

(40) Mikhonin, A. V.; Ahmed, Z.; Ianoul, A.; Asher, S. A. Assignments and Conformational Dependencies of the Amide III Peptide Backbone UV Resonance Raman Bands. *J. Phys. Chem. B* **2004**, *108*, 19020–19028.

(41) Punihale, D.; Hong, Z.; Jakubek, R. S.; Dahlburg, E. M.; Geib, S.; Asher, S. A. Glutamine and Asparagine Side Chain Hyperconjugation-Induced Structurally Sensitive Vibrations. *J. Phys. Chem. B* **2015**, *119*, 13039–13051.

(42) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. *Chem. Rev.* **2012**, *112*, 2604–2628.

(43) Mikhonin, A. V.; Myshakina, N. S.; Bykov, S. V.; Asher, S. A. UV Resonance Raman Determination of Polyproline II, Extended 2.5 1-helix, and  $\beta$ -sheet  $\Psi$  Angle Energy Landscape in Poly-L-Lysine and Poly-L-Glutamic Acid. *J. Am. Chem. Soc.* **2005**, *127*, 7712–7720.

- (44) Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. Peptide Secondary Structure Folding Reaction Coordinate: Correlation Between UV Raman Amide III Frequency,  $\psi$  Ramachandran Angle, and Hydrogen Bonding. *J. Phys. Chem. B* **2006**, *110*, 1928–1943.
- (45) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. Dihedral  $\psi$  Angle Dependence of the Amide III Vibration: a Uniquely Sensitive UV Resonance Raman Secondary Structural Probe. *J. Am. Chem. Soc.* **2001**, *123*, 11775–11781.
- (46) Punihaole, D.; Workman, R. J.; Hong, Z.; Madura, J. D.; Asher, S. A. Polyglutamine Fibrils: New Insights into Antiparallel  $\beta$ -Sheet Conformational Preference and Side Chain Structure. *J. Phys. Chem. B* **2016**, *120*, 3012–3026.
- (47) Myshakina, N. S.; Ahmed, Z.; Asher, S. A. Dependence of Amide Vibrations on Hydrogen Bonding. *J. Phys. Chem. B* **2008**, *112*, 11873–11877.
- (48) Gekko, K.; Ohmae, E.; Kameyama, K.; Takagi, T. Acetonitrile-protein Interactions: Amino Acid Solubility and Preferential Solvation. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1998**, *1387*, 195–205.
- (49) Sugita, Y.; Okamoto, Y. Replica Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (50) Darve, E.; Rodríguez-Gómez, D.; Pohorille, A. Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. *J. Chem. Phys.* **2008**, *128*, 144120.
- (51) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: a Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- (52) Grubmüller, H. Predicting Slow Structural Transitions in Macromolecular Systems: Conformational Flooding. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1995**, *52*, 2893–2906.
- (53) Pande, V. S.; Beauchamp, K.; Bowman, G. R. Everything you Wanted to Know About Markov State Models but were Afraid to Ask. *Methods* **2010**, *52*, 99–105.
- (54) Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein, M. Metadynamics as a Tool for Exploring Free Energy Landscapes of Chemical Reactions. *Acc. Chem. Res.* **2006**, *39*, 73–81.
- (55) Vymětal, J.; Vondrášek, J. Metadynamics as a Tool for Mapping the Conformational and Free-energy Space of Peptides - The Alanine Dipeptide Case Study. *J. Phys. Chem. B* **2010**, *114*, 5632–5642.
- (56) Laio, A.; Gervasio, F. L. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71*, 126601.
- (57) Hénin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. Exploring Multidimensional Free Energy Landscapes Using Time-dependent Biases on Collective Variables. *J. Chem. Theory Comput.* **2010**, *6*, 35–47.
- (58) Gaborek, T. J.; Chipot, C.; Madura, J. D. Conformational Free-energy Landscapes for a Peptide in Saline Environments. *Biophys. J.* **2012**, *103*, 2513–2520.
- (59) Kar, K.; Hoop, C. L.; Drombosky, K. W.; Baker, M. A.; Kodali, R.; Arduini, I.; van der Wel, P. C. A.; Horne, W. S.; Wetzel, R.  $\beta$ -Hairpin-Mediated Nucleation of Polyglutamine Amyloid Formation. *J. Mol. Biol.* **2013**, *425*, 1183–1197.
- (60) Milner-White, E. J.; Poet, R. Four Classes of Beta-hairpins in Proteins. *Biochem. J.* **1986**, *240*, 289–292.
- (61) Shapovalov, M. V.; Dunbrack, R. L. A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure* **2011**, *19*, 844–858.
- (62) Best, R. B.; Zhu, X.; Shim, J.; et al. Optimization of the Additive CHARMM All-atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (63) Tiffany, M. L.; Krimm, S. Effect of Temperature on the Circular Dichroism Spectra of Polypeptides in the Extended State. *Biopolymers* **1972**, *11*, 2309–2316.
- (64) Krimm, S.; Mark, J. E. Conformations of Polypeptides with Ionized Side Chains of Equal Length. *Proc. Natl. Acad. Sci. U. S. A.* **1968**, *60*, 1122–1129.
- (65) Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; MacKerell, A. D. Importance of the CMAP Correction to the CHARMM22 Protein Force Field: Dynamics of Hen Lysozyme. *Biophys. J.* **2006**, *90*, L36–L38.
- (66) Bhattacharyya, A. M.; Thakur, A. K.; Wetzel, R. Polyglutamine Aggregation Nucleation: Thermodynamics of a Highly Unfavorable Protein Folding Reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15400–15405.
- (67) Moradi, M.; Babin, V.; Roland, C.; Sagui, C. Are Long-range Structural Correlations behind the Aggregation Phenomena of Polyglutamine Diseases? *PLoS Comput. Biol.* **2012**, *8*, e1002501.
- (68) Miettinen, M. S.; Knecht, V.; Monticelli, L.; Ignatova, Z. Assessing Polyglutamine Conformation in the Nucleating Event by Molecular Dynamics Simulations. *J. Phys. Chem. B* **2012**, *116*, 10259–10265.
- (69) Kar, K.; Jayaraman, M.; Sahoo, B.; Kodali, R.; Wetzel, R. Critical Nucleus Size for Disease-Related Polyglutamine Aggregation is Repeat-Length Dependent. *Nat. Struct. Mol. Biol.* **2011**, *18*, 328–336.
- (70) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33–38.
- (71) Martyna, G. J.; Tobias, T. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189.
- (72) Darnell, G. D.; Derryberry, J.; Kurutz, J. W.; Meredith, S. C. Mechanism of Cis-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect. *Biophys. J.* **2009**, *97*, 2295–2305.
- (73) Walsh, P. S.; Blodgett, K. N.; McBurney, C.; Gellman, S. H.; Zwier, T. S. Inherent Conformational Preferences of Ac-Gln-Gln-NH<sub>2</sub>: Sidechain Hydrogen Bonding Supports a  $\beta$  Turn in the Gas Phase. *Angew. Chem., Int. Ed.* **2016**, *55*, 14618–14622.
- (74) Chen, M.; Tsai, M.; Zheng, W.; Wolynes, P. G. Exploring the Aggregation Free Energy Landscape of the Amyloid- $\beta$  Protein (1–40). *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (42), 11835–11840.
- (75) Feng, L. Molecular Dynamics Predicts the Solution Conformations of Poly-L-lysine in Salt Solutions. 2014 Master's Thesis, University of Pittsburgh.
- (76) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins: Struct., Funct., Genet.* **1993**, *17* (1), 75–86.
- (77) Heck, B. S.; Doll, F.; Hauser, K. Length-dependent Conformational Transitions of Polyglutamine Repeats as Molecular Origin of Fibril Initiation. *Biophys. Chem.* **2014**, *185*, 47–57.