PNAS
www.pnas.org

Supplementary Information for

Probing initial transient oligomerization events facilitating Huntingtin fibril nucleation at atomic resolution by relaxation-based NMR

Samuel A. Kotler[a], Vitali Tugarinov[a], Thomas Schmidt[a], Alberto Ceccon[a], David S. Libich[a,1], Rodolfo Ghirlando[b], Charles D. Schwieters[c], and G. Marius Clore[a,2]

[a]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520. [b]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0530. [c]Imaging Sciences Laboratory Center for Information Technology, National Institutes of Health, Bethesda, MD 20892-5624

[1]Present address: Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio.

[2]Corresponding author: G. Marius Clore. Email: mariusc@mail.nih.gov

This PDF file includes:

- Supplementary text: (a) SI Materials and Methods. (b) SI Theory and Data Fitting of Relaxation Dispersion and Exchange-Induced Chemical Shift Data. (c) SI PRE-Based Structure Modeling and Xplor-NIH Calculations
- Figs. S1 to S22
- References for SI reference citations

SI Materials and Methods

Expression and purification of $htt^{NT}Q_n$ peptides. Huntingtin N-terminal fragments $htt^{NT}Q_n$ (n = 0, 7, 10, 14) (Fig. 1, main text) were expressed in *E. coli* as a fusion protein with the immunoglobulin-binding domain of streptococcal protein G (GB1), following a similar protocol to that described previously (Ceccon et al., 2018). Codon optimized versions of $htt^{NT}Q_n$ constructs were synthesized by Genscript Inc. (Piscataway, NJ), and sub-cloned into a pET-21d(+) plasmid using NcoI and BamHI cloning sites. The plasmid had a Factor Xa cleavage site to avoid having the inclusion of any extra residues prior to the htt^{NT} sequence. Additionally, the $htt^{NT}Q_n$ peptides were designed with sequences lacking an N-terminal methionine, as the latter is likely enzymatically removed in the cell (1). Thus, the first amino acid is alanine and the htt^{NT} segment consists of 16 residues (ATLEKLMKAFESLKSF). The resulting constructs were expressed in BL21 Star (DE3) *E. coli* cells (Life Technologies). Two additional constructs of $htt^{NT}Q_7$ were made for nitroxide spin labeling with single point mutations, S12C or S15C, using the QuikChange site-directed mutagenesis kit (Stratagene).

E. coli cells containing the expression plasmid were grown at 37 °C in M9 minimal medium supplemented with the relevant ^{15}N and/or ^{13}C isotope source to OD_{600} ~ 0.6 and subsequently induced with 0.5 mM isopropyl β-D-1 thiogalactopyranoside (IPTG). For uniform ^{15}N and ^{13}C labeling, $^{15}NH_4Cl$ and $^{13}C_6$-D-glucose were used as the sole nitrogen and carbon sources, respectively. For fractional $^{13}C\alpha$ enrichment, [2-^{13}C]-D-glucose was employed as the sole carbon source (2).The protein was expressed at 20 °C for an additional 18 h, after which cells were harvested by centrifugation (4000 g x 20 min), resuspended in lysis buffer (50 mM Tris-HCl, pH 8, 100 mM NaCl, and EDTA-free protease inhibitor mixture (Roche Life Science), and lysed by heating to 80 °C for 10 min. The cell debris was removed by centrifugation (40,000 g x 25 min) and the supernatant dialyzed against 2 x 4L lysis buffer over a 24 h period. Following dialysis, 2 mM $CaCl_2$ was added to the lysis buffer and the fusion protein (~4 mg/ml) was treated for 24 h at 4 °C with Factor Xa (0.0065 mg/mL) to cleave off the GB1 fusion domain. Following the cleavage reaction, the $htt^{NT}Q_n$ peptides were purified by reverse phase high performance liquid chromatography (HPLC) using a preparative scale C4 column (Vydac). The peptide was eluted with a gradient of 20-45% (v/v) solvent B (95% acetonitrile, 5% H_2O, and 0.1% TFA) in solvent A (5% acetonitrile, 95% H_2O, and 0.1% TFA) over 30 min. The cleavage reaction, peptide identity and isotopic incorporation were confirmed by electrospray ionization mass spectrometry (ESI MS). In the case of $^{13}C\alpha$ labeled samples, fractional enrichment at the Cα positions was further assessed by comparing cross-peak intensities in ^{1}H-^{13}C constant-time heteronuclear single quantum coherence (CT-HSQC) spectra of uniformly ^{13}C-labeled and $^{13}C\alpha$-labeled samples.

Following peptide purification, a previously described disaggregation protocol (Chen and Wetzel, Protein Sci, 2001) was used to ensure that samples were devoid of preformed aggregates as these can accelerate the aggregation of monomeric peptides. After HPLC purification, eluted fractions containing pure $htt^{NT}Q_n$ peptides were lyophilized and resuspended in a 1:1 (v/v) mixture of trifluoroacetic acid (TFA) and hexafluoroisopropanol (HFIP) at a peptide concentration of 1

mg/mL. The TFA:HFIP solvent mixture was removed with N_2 gas and the resulting peptide film was further lyophilized for 4 h to ensure complete removal of both solvents. The peptide film was then dissolved in 0.1 mM TFA, pH 3, aliquoted, and freeze-dried for 24 h. Peptide aliquots were stored at -30 °C and each aliquot used only once.

NMR sample preparation. All NMR samples were prepared by dissolving a single aliquot of $htt^{NT}Q_n$ peptide in 13.8 mM monobasic sodium phosphate, pH 4.6, 50 mM NaCl, 10% D_2O (v/v); the peptide concentration was adjusted to 1.5 mM for htt^{NT} and $htt^{NT}Q_7$ or 0.5 mM for $htt^{NT}Q_{10}$ and $htt^{NT}Q_{14}$. Since the peptide sequence contains no tyrosine or tryptophan residues, the peptide concentration was measured using absorbance at 205 nm (3). The $htt^{NT}Q_n$ peptides were dissolved initially at pH 4.6 as this improved peptide solubility. The pH of the buffer was then adjusted to 6.5 by adding dibasic sodium phosphate such that the final sodium phosphate concentration was 20 mM. From the 1.5 mM stock peptide solution, carefully matched samples were diluted to appropriate final concentrations used in NMR measurements. Unless otherwise noted, all samples were prepared in the same manner such that the final NMR buffer was 20 mM sodium phosphate, pH 6.5, 50 mM NaCl, and 10% D_2O (v/v). Methionine oxidized $htt^{NT}Q_7$ (Met^7O-$htt^{NT}Q_7$) was prepared by adding 0.5% (v/v) H_2O_2 to the NMR buffer. Complete methionine oxidation to a sulfoxide was verified by ESI MS resulting in a +16 Da increase in molecular mass, as well as by NMR.

Nitroxide spin labeling of N-terminal Huntingtin fragments and GB1. Following HPLC purification, GB1-$htt^{NT}Q_7$ (S12C), GB1-$htt^{NT}Q_7$ (S15C) or $htt^{NT}Q_7$ (S15C) were aliquoted and lyophilized for at least 24 h to remove all HPLC solvents. GB1- $htt^{NT}Q_7$ (S12C) and GB1-$htt^{NT}Q_7$ (S15C) are samples in which the GB1 domain located N-terminal to the $htt^{NT}Q_7$ cysteine mutant were not removed by factor Xa cleavage (note residue numbers are referenced throughout relative to the htt^{NT} sequence). GB1-$htt^{NT}Q_7$ spin-labeled constructs were employed for PRE experiments as the spin-labeled S12C and S15C mutants of $htt^{NT}Q_7$ displayed enhanced aggregation properties resulting in severe limitations on yields necessary for NMR measurements. Aliquots of purified material were re-suspended in a 250 μL solution of 2M guanidine hydrochloride (GdnHCl), pH 2.5, and kept on ice to ensure the peptide remained monomeric. The peptide solution was then re-equilibrated to a final volume of 500 μL to yield a solution comprising 50 mM Tris, pH 8, 50 mM NaCl, 1M GdnHCl, 0.5 mM peptide and 2.5 mM (1-oxyl-2,2,2,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL, Toronto Research Chemicals). The MTSL spin-labeling reaction was allowed to proceed in the dark overnight. Unreacted spin-label was removed by size-exclusion chromatography using a Superdex Peptide 10/300 GL column (GE Healthcare Life Sciences) with an elution buffer of 20 mM sodium phosphate, pH 6.5, and 50 mM NaCl. Spin-labeling was tested for completion by ESI MS before and after the samples were passed though the Superdex Peptide column. Following buffer exchange, MTSL-labeled samples at natural isotopic abundance were mixed with ^{15}N-labeled $htt^{NT}Q_7$. Samples for PRE experiments comprised 600 μM total peptide concentration in a molar ratio of 1:40 MTSL-labeled (S12C-R1 or S15C-R1) GB1-

htt^{NT}Q$_7$ to ^{15}N-labeled htt^{NT}Q$_7$. The diamagnetic control samples comprised 600 μM ^{15}N-labeled htt^{NT}Q$_7$.

Control experiments to exclude solvent PREs and PREs arising from non-specific binding of the nitroxide spin label were also carried out on the following two samples: (i) 150 μM ^{15}N-labeled htt^{NT}Q$_7$ and 150 μM TEMPO (Sigma-Aldrich), and (ii) 150 μM ^{15}N-labeled htt^{NT}Q$_7$ and 150 μM MTSL-labeled N-Cys GB1 (GB1 domain with an N-terminal cysteine). N-Cys GB1 was a gift from John M. Louis, (Laboratory of Chemical Physics, NIDDK) and was spin labeled using the same protocol described above. Unreacted spin-label in the N-Cys GB1 preparation was removed by size exclusion chromatography using a Superdex Peptide column and exchanged directly into NMR buffer. Spin-labeling of N-Cys GB1 was verified for completion by ESI MS.

NMR spectroscopy. All NMR experiments were recorded at 5 °C or 10 °C, using 900, 800 and 600 MHz Bruker Avance-III spectrometers, equipped with TCI z-axis gradient cryogenic probes, unless otherwise noted. Temperature differences between spectrometers were corrected by matching the chemical shift difference between residual water and the methyl resonance of 4,4-dimethyl-4-silapentance-1-sulfonic acid (DSS) (0 ppm) in a sample comprising 20 mM sodium phosphate, pH 6.5, 50 mM NaCl, 0.5% DSS, and 99.9% D_2O (v/v). Backbone ^{1}H, ^{13}C and ^{15}N resonance assignments for ^{13}C/^{15}N-labeled htt^{NT}Q$_n$ peptides and GB1-htt^{NT}Q$_7$ were obtained using 3D triple resonance experiments (HNCO, HNCACB, and CBCA(CO)NH (4)). All NMR spectra were processed using NMRPipe(5).

^{15}N transverse relaxation rates in the rotating frame. ^{15}N-$R_{1\rho}$ and ^{15}N-R_1 measurements at 600 and 900 MHz were recorded in an interleaved manner using heat-compensated ^{1}H-^{15}N HSQC-based pulse schemes (6) on samples of ^{15}N-labeled htt^{NT} and htt^{NT}Q$_7$ at various peptide concentrations (see Fig. 2, main text). The effective spin-lock field for the ^{15}N-$R_{1\rho}$ experiments was 1.0 kHz. ^{15}N-R_2 (in-phase) values in the rotating frame were determined from $R_2 = (R_{1\rho} - R_1\cos^2\theta)/\sin^2\theta$, where θ is the angle between the effective spin-lock field and the external magnetic (B_0) field. ^{15}N-$\Delta R_2^{1\,\mathrm{kHz}}$ values were obtained by taking the difference between ^{15}N-$R_2^{1\,\mathrm{kHz}}$ values at a particular concentration and in the absence of exchange. The latter were determined by fitting the concentration dependence of $R_2^{1\,\mathrm{kHz}}$ (see Fig. S5) to a second order polynomial to extrapolate the value at zero concentration.

^{1}H$_N$ PRE measurements. ^{1}H$_N$-R_2 values were obtained using a 2D ^{1}H-^{15}N HSQC-based pulse sequence at 600 MHz using a two-time point measurement (separated by 28 ms) as described previously (7). ^{1}H$_N$-Γ_2 values, given by the difference in ^{1}H$_N$-R_2 values between paramagnetic and diamagnetic samples, and associated errors were calculated using Eqs. (5) and (6) of ref. (7).

CPMG relaxation dispersion measurements. ^{15}N CPMG relaxation dispersions were recorded at 600 and 900 MHz using a pulse scheme with amide proton decoupling to measure the rates of in-phase ^{15}N coherences (8) with CPMG fields, ν_{CPMG} (= $1/2\tau_{CP}$, where τ_{CP} is the time

between 180° ^{15}N CPMG pulses) of 0, 10, 20, 40, 70, 90, 110, 130, 150, 180, 210, 240, 280, 320, 420, 500, 580, 660, 740, 820, 900, and 1000 Hz. The ^{15}N-CPMG evolution period was set to 100 ms. $^{1}H_N$ constant wave (CW) decoupling was applied at a radiofrequency (RF) field strength of 11 kHz. The experiment recorded with the relaxation period omitted served as a reference for the calculation of $R_{2,eff}$ rates as a function of ν_{CPMG} as described previously (8). Uncertainties in $R_{2,eff}$ values were obtained from duplicate measurements at two different ν_{CPMG} values. 4, 4, 8 and 32 scans per t_1 increment were acquired for the 1, 0.75, 0.4 and 0.05 mM ^{15}N-labeled htt^{NT}Q$_7$ samples, respectively.

^{13}Cα relaxation dispersion experiments were acquired using a ^{13}Cα-CPMG pulse scheme with Hα decoupling (9) at 600 MHz and 900 MHz with ν_{CPMG} values of 0, 25, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 750, 925, 1125, 1375, 1625, and 2000 Hz. ^{13}Cα CPMG evolution period was set to 40 ms. ^{1}Hα CW decoupling was applied with an RF field strength of 12 kHz. The experiment recorded with the relaxation period omitted served as a reference for the calculation of $R_{2,eff}$ rates as a function of ν_{CPMG} (9). 8, 8, 16 and 32 scans per t_1 increment were acquired for the 1, 0.75, 0.4 and 0.05 mM ^{13}Cα/^{15}N-labeled htt^{NT}Q$_7$ samples, respectively. Relaxation dispersion on ^{13}Cα/^{15}N-labeled htt^{NT} was only acquired on a 0.9 mM sample as no dispersions were observed for this peptide.

On-resonance ^{15}N-$R_{1\rho}$ relaxation dispersion. On-resonance ^{15}N-$R_{1\rho}$ relaxation dispersions experiments were performed using the same pulse scheme as that used for ^{15}N-$R_{1\rho}$ measurements (6) with the carrier placed at 121.5 ppm and spin-lock field strengths of 1000, 1250, 1500, 1750, 2000, 2250, and 2500 Hz. For each spin-lock field, effective $R_{1\rho}$ rates were measured using a two-time-point measurement, $R_{1\rho} = (1/\Delta T)\ln(I/I_n)$ where, ΔT is the difference in the two relaxation delays. The maximum relaxation delay was adjusted such that the NMR intensity was attenuated by 50-60% of the intensity in the minimum delay and errors in $R_{1\rho}$ measurements were propagated from the two-point time calculation as in (7).

^{13}C and ^{15}N exchange-induced chemical shifts. Changes in ^{13}Cα and ^{15}N chemical shifts of uniformly ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ as a function of concentration between 10 μM and 1.2 mM (13 data points) were obtained from ^{1}H-^{13}C CT-HSQC and ^{1}H-^{15}N HSQC spectra, respectively, recorded at 800 MHz. A concentrated 1.2 mM ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ sample was prepared (as described above) and serially diluted with NMR buffer (20 mM sodium phosphate, pH 6.5, 50 mM NaCl, and 10% v/v D_2O). The chemical shifts as a function of concentration were fit to a second-order polynomial and the fitted *y*-intercept (corresponding to zero peptide concentration) was used as the reference chemical shift (δ_{ref}) from which the ^{13}Cα and ^{15}N exchanged-induced chemical shifts, $\delta_{ex}(i)$, were calculated as $\delta_{obs}(i) - \delta_{ref}$, where $\delta_{obs}(i)$ is the observed chemical shift at a concentration i of peptide. In the case of ^{13}C/^{15}N-labeled htt^{NT}, exchange-induced chemical shifts were only calculated between concentrations of 800 μM and 20 μM because no measurable shifts were observed between these two samples (see Fig. 2).

$^{13}C\alpha$ shifts were measured from high-resolution ^{1}H-^{13}C CT-HSQC spectra (10) acquired with 420* x 1536* complex data points in the indirect (^{13}C) and direct (^{1}H) dimensions, corresponding to acquisitions times of 42.0 and 121.7 ms, respectively. To achieve the necessary high resolution to permit accurate and precise determination of small $^{13}C\alpha$ chemical shift differences, the total constant time period was set to 56 ms ~$2/J_{CC}$. ^{15}N shifts were measured from high-resolution ^{1}H-^{15}N HSQC spectra acquired with 128* x 1536* complex data points in the indirect (^{15}N) and direct (^{1}H) dimensions, corresponding to acquisitions times of 107.0 and 121.7 ms, respectively. The heteronuclear chemical shifts in each spectrum were extracted by peak picking and subsequent multidimensional line shape fitting using the *nlinLS* routine in the NMRPipe software package (5). Errors in peak positions due to random noise were determined using NMRPipe as described previously (11). In addition, the $^{13}C\alpha$ shifts of Thr2 and the three C-terminal glutamines (residues 21-23), and the ^{15}N shifts of the side chain NH_2 groups of the three C-terminal glutamines served as internal references as they showed no evidence of any changes in peak position as a function of concentration, thereby ensuring continuity of sample conditions over the entire concentration range. The standard deviation in peak positions for these reference peaks was 0.16 and 0.10 Hz for ^{13}C and ^{15}N, respectively.

Atomic force microscopy (AFM). For AFM imaging, a 10 μL aliquot of an $htt^{NT}Q_7$ solution was taken from a 0.6 mM sample after completion of NMR experiments (~3 weeks after initially solubilizing the peptide in NMR buffer) and deposited on a freshly cleaved, dry mica surface. The solution was adsorbed for 10 min before blotting with a Kimwipe. The mica surface was then washed twice with 10 μL deionized water to remove buffer salts and dried under a gentle N_2 gas flow. AFM images were recorded in tapping mode using a Veeco MultiMode instrument and Nancoscopy IV controller, equipped with a MikroMasch (Wastonville, CA) tip. The tip oscillation was ~200 kHz with a drive amplitude of 100-150 mV and a detector set point of ~0.6 V. Images typically contained 256 samples per line in a 5 μm x 5 μm area with the tip velocity set at 1 μm/s. Height, feedback error signal, and phase images were recorded simultaneously.

Pulsed Q-band EPR spectroscopy. Pulsed EPR data were collected at Q-band (33.98 GHz) at a temperature of 50 K on a Bruker E-580 spectrometer equipped with a 150 W traveling-wave tube amplifier, a model ER5107D2 resonator, and a cryo-free cooling unit. Samples (in 30% v/v glycerol) were placed in 1 mm internal diameter quartz tubes (Wilmad WG-221T-RB) and flash frozen in liquid nitrogen.

Q-band DEER data were recorded with a four-pulse scheme (12). The observe and pump pulses were separated by ~85 MHz. The lengths of the observe π and π/2 pulses were 24 and 12 ns, respectively; the length of the ELDOR π pump pulse was 10 ns. The pump frequency was centered at the maximum of the Q-band nitroxide spectrum located at +40 MHz from the center of the resonator frequency. The τ_1 value (350 ns) of the first echo period ($2\tau_1$) was incremented ten times in 2 ns increments to average out ^{1}H modulation; the position of the pump pulse was incremented in steps (Δt) of 2 ns. (Note that the data points in the DEER echo curves in Fig. 10A and Figs. S19

and S20 are plotted at Δt intervals of 8 ns). The bandwidth of the overcoupled resonator was ~120 MHz. Data collection during the second echo period ($2\tau_2$) was not carried out over the full τ_2 range because of a persistent "2+1" echo perturbation of the DEER echo curve at a time of about τ_1 from the final observe π pulse. The pulse gate time used for echo integration was 32–38 ns. Because of the short phase memory time T_m of the samples arising from the use of a protonated environment, the maximum dipolar evolution time t_{max} was set to 2.5 μs.

Q-band DEER echo curves were analyzed with the program DeerAnalysis 2016 (13) to derive $P(r)$ distributions using Tikhonov regularization. Validated Tikhonov regularization was carried out varying the modulation depth (11 steps), background density (11 steps) and background start (11 steps) for a total of 1331 permutations. The Tikhonov regularization parameter α was automatically determined by DeerAnalysis for each iteration, and the value of α ranged from 10 to 100. The position of a given $P(r)$ maximum and width was determined by evaluating the mean and width of the distribution.

Inversion-modulation DEER (IM-DEER) experiments were used to determine the oligomerization state of GB1-htt^{NT}Q$_7$ (S15C-R1) by analyzing the normalized modulation depth (Δ/Δ_{max}) as a function of ELDOR pulse flip angle (ranging from ~20 to 180°) as described previously (14). The length of the 180° ELDOR pulse as a function of attenuation and the maximum inversion efficiency (λ_{max}) were determined using a spin-echo nutation experiment as described (14). The attenuation settings for the ELDOR pulse were restricted to experimental λ_{max} values $\geq$0.7. The dependence of $\Delta(\theta)/\Delta_{max}$ on the ELDOR pulse flip angle θ (where Δ_{max} is the modulation depth at θ = 180°) and the number N of nitroxide spin labels was described in (14).

X-band constant wave (CW) EPR. GB1-htt^{NT}Q$_7$ (S12C-R1 or S15C-R1) samples were loaded into quartz capillaries (0.6-mm inner diameter × 0.84-mm outer diameter), and room temperature X-band CW EPR spectra were recorded on a Bruker EMX spectrometer with a scan width of 150 Gauss using a high sensitivity (HS) resonator cavity and an incident microwave power of 20 mW. To facilitate direct comparison all CW EPR spectra were normalized to the same amount of spins by double integration. Line broadening arising from strong dipolar coupling between closely separated nitroxide labels (Fig. 11A) and the resulting distance distributions (Fig. 11B) were determined from the concentrated (dimer for S15C-R1, see Fig. 10C and S15; dimer and tetramer for S12C-R1, see Fig. S14) samples with the dilute (10 μM) monomeric samples employed as a reference using the program 'ShortDistances' by Christian Altenbach (available for free download at http://www.chemistry.ucla.edu/directory/hubbell-wayne-l) (15). A single Gaussian distance distribution was employed in the fitting procedure with a Tikhonov factor α of 4 x 10^{-4} for the GB1-htt^{NT}Q$_7$ (S12C-R1) data. ShortDistances also estimates the fraction of monomer in the concentrated (3 mM) GB1-htt^{NT}Q$_7$ (S12C-R1) sample to be 11%. Dipolar broadening in the case of GB1-htt^{NT}Q$_7$ (S15C-R1) is very small and hence one can only ascertain that the distance between nitroxide labels lies in the 20-30 Å range.

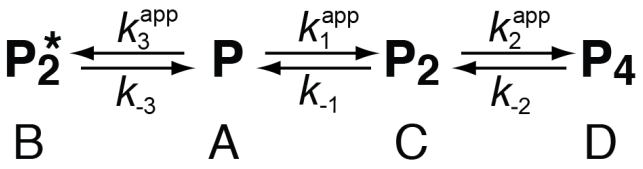
Sedimentation velocity analytical ultracentrifugation. Sedimentation velocity experiments were carried out at 60,000 rpm and 10 °C for $htt^{NT}Q_7$ on a Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge and An60-Ti rotor following standard protocols (16). 620 and 310 μM samples of ^{15}N-labeled $Htt^{NT}Q_7$ in 50 mM NaCl, 20 mM sodium phosphate (pH 6.5) and 10% (v/v) D_2O were loaded in 3-mm two-channel centerpiece cells, and sedimentation data were collected using the absorbance (230 nm) and interference (655 nm) optical detection systems. GB1-$htt^{NT}Q_7$ (S15C-R1) was studied at concentrations ranging from 0.09 to 1.1 mM at 50,000 rpm and 20 °C using a An50-Ti rotor. The buffer comprised 50 mM sodium phosphate (pH 6.5) and 100 mM NaCl; a 12 mm cell was used for the lowest concentration, and 3 mm cells for the other concentrations.

Time corrected (17) data were analyzed in SEDFIT 15.01c (18) in terms of a continuous c(*s*) distribution of sedimenting species with a maximum entropy regularization confidence interval of 0.68. Excellent fits are noted with root mean square deviations of 0.0052 – 0.0062 absorbance units, and 0.0032 – 0.0040 fringes for ^{15}N labeled $Htt^{NT}Q_7$. In the case of GB1-$htt^{NT}Q_7$ (S15C-R1) values ranged from 0.0045 to 0.0095 absorbance units, and 0.0059 to 0.013 fringes.. The solution density, solution viscosity, and protein partial specific volume were calculated based on their composition in SEDNTERP (19)and corrected for their isotopic composition. Sedimentation coefficients were corrected to standard conditions in water at 20 °C, $s_{20,w}$. Monomer and dimer populations obtained from integration of the GB1-$htt^{NT}Q_7$ (S15C-R1) c*(s)* distributions were used to construct partial concentration isotherms. Data were analyzed in SEDPHAT 14.0 (20)in terms of a monomer-dimer self-association to obtain the equilibrium constants; errors were determined using the method of F-statistics with a confidence level of 68%.

Sedimentation equilibrium analytical ultracentrifugation. Sedimentation equilibrium experiments were carried out to determine the oligomeric state of GB1-$htt^{NT}Q_7$ and GB1-$htt^{NT}Q_7$ (S12C-R1) in solution. At equilibrium the method reports on all species present in solution, independent of the exchange rates (k_{off}) between species. Samples were loaded at various concentrations, ranging from 17 to 560 μM for GB1-$htt^{NT}Q_7$ and from 16 to 300 μM for GB1-$htt^{NT}Q_7$ (S12C-R1), into 3 mm 2-channel epon centerpiece cells (40 μL) and 12 mm 2-channel epon centerpiece cells (160 μL) depending on the concentration. The buffer was 50 mM sodium phosphate (pH 6.5), and either 150 mM NaCl for GB1-$htt^{NT}Q_7$ or 100 mM NaCl for GB1-$htt^{NT}Q_7$ (S15C-R1). Sedimentation equilibrium experiments were conducted at 20°C and 15,000, 25,000 and 40,000 rpm on a Beckman Optima XL-A analytical ultracentrifuge and An50-Ti rotor following standard protocols (16), with absorbance data collected at 250 and 280 nm. Data were analyzed globally in terms of a monomer-dimer or monomer-dimer-tetramer self-association with implicit mass conservation in SEDPHAT 14.0 essentially as described (21). The buffer solution density and protein partial specific volume were calculated based on the composition using SEDNTERP (19).

SI Theory and Data Fitting of Relaxation Dispersion and Exchange-Induced Chemical Shift Data

The 4-state kinetic scheme used to analyze the $htt^{NT}Q_7$ experimental NMR data is (*cf.* Fig. 5 of the main text):



Scheme S1

where **P** and $\mathbf{P}_n$ are the $htt^{NT}Q_7$ monomer and oligomer species of order n, respectively, k_i^{app} are 'apparent', pseudo-first-order association rate constants for dimerization/tetramerization, and k_{-i} are dissociation rate constants. $\mathbf{P_2^*}$ denotes an off-pathway, 'abortive' dimer that by virtue of its structural properties cannot oligomerize further (form tetramers). For simplicity of notation, in what follows we refer to each of the states by the corresponding letter-code as shown in Scheme S1. As described in detail in the main text, the inter-conversion between the monomeric species A and the abortive, off-pathway dimer B occurs on a relatively slow timescale (~2,000 s^{-1}) making it amenable to analysis by CPMG relaxation dispersion, whereas the inter-conversion between A and the dimeric/tetrameric species C/D proceeds much faster (> ~20,000 s^{-1}) and can be characterized by exchange-induced chemical shifts and $R_{1\rho}$ relaxation dispersion. A combination of these NMR techniques in conjunction with simultaneous analysis of all data is therefore necessary for a detailed characterization of the kinetics of oligomerization in Scheme S1.

Analysis of ^{15}N- and ^{13}Cα-CPMG profiles and exchange-induced chemical shifts. The evolution of magnetization during the constant-time CPMG period is given by: $M(t) = (AA^*A^*A)^n M(0)$, where $M = [M^{\mathrm{A}}; M^{\mathrm{B}}; M^{\mathrm{C}}; M^{\mathrm{D}}]^T$; M^N denotes the transverse magnetization of state N; T denotes transposition; $A = exp(-\tilde{R}\tau_{CP}/2)$; A^* is the complex conjugate of A; n is the number of CPMG cycles employed; τ_{CP} is the time interval between two consecutive 180° pulses in the CPMG pulse train; and the effective relaxation matrix $\tilde{R} = \tilde{R}^{cs} + \tilde{R}^{rel} + \tilde{R}^{ex}$, where

$$\tilde{R}^{cs} = i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \Delta\varpi^{\mathrm{B}} & 0 & 0 \\ 0 & 0 & \Delta\varpi^{\mathrm{C}} & 0 \\ 0 & 0 & 0 & \Delta\varpi^{\mathrm{D}} \end{bmatrix} \qquad \text{(S1.1)}$$

$$\tilde{R}^{rel} = \begin{bmatrix} R_2^{\mathrm{A}} & 0 & 0 & 0 \\ 0 & R_2^{\mathrm{B}} & 0 & 0 \\ 0 & 0 & R_2^{\mathrm{C}} & 0 \\ 0 & 0 & 0 & R_2^{\mathrm{D}} \end{bmatrix} \tag{S1.2}$$

$$\tilde{R}^{ex} = \begin{bmatrix} k_1^{\mathrm{app}} + k_3^{\mathrm{app}} & -k_{-3} & -k_{-1} & 0 \\ -k_3^{\mathrm{app}} & k_{-3} & 0 & 0 \\ -k_1^{\mathrm{app}} & 0 & k_{-1} + k_2^{app} & -k_{-2} \\ 0 & 0 & -k_2^{\mathrm{app}} & k_{-2} \end{bmatrix} \tag{S1.3}$$

where $\Delta\varpi^{\mathrm{N}}$ is the difference between chemicals shifts of state *N* and the (observable) state A (rad/s), R_2^{A} is the transverse relaxation rate of state A in the absence of exchange, and k_n^{app} and k_{-n} in the exchange matrix $\tilde{R}^{ex}$ (Eq S1.3) are pseudo first-order apparent association and dissociation rate constants, respectively (see scheme S1).

All the states were assumed to be present at the start of the CPMG train, $M(0) = (\alpha p_{\mathrm{A}}, \alpha p_{\mathrm{B}}, \alpha p_{\mathrm{C}}, \alpha p_{\mathrm{D}})$, where p_{i} is the equilibrium population of state *i*, and α is a residue-specific scaling factor that takes into account that ^{13}C enrichment at Cα positions is fractional and varies from ~15% to ~45% depending on the type of residue (2), while all the peptide molecules (^{13}Cα-labeled monomers + unlabeled monomers) participate in exchange involving formation of oligomers, *i.e.* oligomers can be formed from either ^{13}Cα-labeled or unlabeled constituents as well as any combination thereof ($\alpha = 1$ for ^{15}N-CPMG dispersions). The population of the main (observable) state A at each concentration was calculated from the rate constants using the relationship,

$$p_{\mathrm{A}} = \frac{1}{1 + \dfrac{k_3^{\mathrm{app}}}{k_{-3}} + \left(\dfrac{k_1^{\mathrm{app}}}{k_{-1}}\right)\left(1 + \dfrac{k_2^{\mathrm{app}}}{k_{-2}}\right)} \tag{S2}$$

The remaining populations are given by: $p_{\mathrm{B}} = \left(k_3^{\mathrm{app}} / k_{-3}\right) p_{\mathrm{A}}$, $p_{\mathrm{C}} = \left(k_1^{\mathrm{app}} / k_{-1}\right) p_{\mathrm{A}}$, and $p_{\mathrm{D}} = 1 - p_{\mathrm{A}} - p_{\mathrm{B}} - p_{\mathrm{C}}$.

The phenomenological ('apparent'), pseudo-first-order forward rate constants k_i^{app} are related to the corresponding 'true' kinetic rate constants k_{i} through,

$$k_1^{\text{app}} = 2k_1\left[\text{A}\right], \qquad k_3^{\text{app}} = 2k_3\left[\text{A}\right] \tag{S3.1}$$

for dimerization, and

$$k_2^{\text{app}} = 2k_2\left(\frac{k_1}{k_{-1}}\right)\left[\text{A}\right]^2 \tag{S3.2}$$

for tetramerization, where [A] is the absolute (molar) concentration of the main state A (see (22) and(23) for derivation of the dependence of oligomerization rate constants on protein concentration pertinent to NMR experiments). Since the population of the oligomeric states is small under the experimental conditions employed, [A] can be approximated by the total peptide concentration. Note that the 'true' kinetic rate constants k_i and k_{-i} were used as free variable parameters in the fits (see below), with the apparent forward rate constants entering the exchange matrix $\tilde{R}^{ex}$ (Eq S1.3) recast according to the relationships in Eqs S3.1-2.

Exchange-induced chemical shifts, δ_{ex}, were calculated from the imaginary part of the smallest (by absolute magnitude) eigenvalue of the matrix $\tilde{R} = \tilde{R}^{cs} + \tilde{R}^{rel} + \tilde{R}^{ex}$ (Eq S1): $\delta_{\text{ex}}(\text{Hz}) = Im(\min[eig\{\tilde{R}\}])/2\pi$, where $eig\{\tilde{R}\}$ is a vector of complex eigenvalues of $\tilde{R}$. An offset φ was added to δ_{ex} to account for any small errors in the extrapolated zero concentration reference value, δ_{ref}, obtained from fitting the concentration dependence of the shifts to a second order polynomial. The residue-specific values of φ were included as free variable parameters in the fits (see below) and had very small optimized values (generally less than 0.1 to 0.2 Hz).

Analysis of ^{15}N $R_{1\rho}$ relaxation dispersion profiles. ^{15}N $R_{1\rho}$ relaxation dispersion profiles were analyzed by solving the set of homogeneous Bloch-McConnell equations (24) for the exchanging system shown in Scheme 1 via the propagation of a 12x12 dimensional Liouvillian $\tilde{L}$ given by,

$$\tilde{L}_{12} = \begin{bmatrix} \tilde{R}^{\text{A}} & 0 & 0 & 0 \\ 0 & \tilde{R}^{\text{B}} & 0 & 0 \\ 0 & 0 & \tilde{R}^{\text{C}} & 0 \\ 0 & 0 & 0 & \tilde{R}^{\text{D}} \end{bmatrix} + \tilde{R}^{ex} \otimes \tilde{I}_3 \tag{S4}$$

where $\tilde{R}_{ex}$ is the exchange matrix (Eq. S1.3), $\tilde{I}_3$ is the 3-dimensional identity matrix, '$\otimes$' denotes the Kronecker product, and each of the 3x3 matrix blocks $\tilde{R}^N$ of the block-diagonal matrix on the right-hand side of Eq. S4 (where N is a particular state, $N \in \{\text{A; B; C; D}\}$), has the form,

$$\widetilde{R}^{N} = \begin{bmatrix} R_2^N & \Omega^N & 0 \\ -\Omega^N & R_2^N & \omega_1 \\ 0 & -\omega_1 & R_1^N \end{bmatrix} \quad \text{(S5)}$$

where R_1^N and R_2^N are intrinsic longitudinal and transverse spin relaxation rates, respectively, of state N in the absence of exchange; Ω^N are offsets (in rad/sec) of the strong irradiation field from the resonance frequency of state N, and ω_1 is the strength of the continuous-wave irradiation ('spin-lock') field applied along the x-axis. The additional set of concentration-dependent in-phase transverse spin relaxation rates R_2 obtained at a single spin-lock field strength of 1.0 kHz (see 'Materials and Methods'), was analyzed in the same manner (via propagation of the Liouvillian in Eqs. S4-S5).

Simultaneous analysis of ^{15}N- and ^{13}Cα-CPMG relaxation dispersions, ^{15}N- and ^{13}Cα-exchange-induced chemical shifts, and ^{15}N-$R_{1\rho}$ relaxation dispersion data. The following simplifying assumptions were made with respect to the kinetic model shown in Scheme 1 in all calculations: the values of the ^{15}N and ^{13}Cα chemical shift changes for the 'productive' dimer, $\Delta\varpi^{C}$, were assumed to be equal to the corresponding values for the tetramer , $\Delta\varpi^{D}$ (Eq. S1.1), and as a consequence Ω^{C} equal to Ω^{D} in Eq. S5. The transverse spin relaxation rates of the dimeric species R_2^B and R_2^C were assumed to be equal to $2R_2^A$, while those of the tetramer (R_2^D) equal to $4R_2^A$. The longitudinal spin relaxation rates of the major state A (R_1^A ; Eq. S5) were measured in a separate experiment (see ' SI Materials and Methods'). The R_1 values for all minor states (B, C and D) were assumed to be equal to R_1^A .

All of the experimental data for htt^{NT}Q$_7$ obtained at a number of different concentrations (see 'SI Materials and Methods') were fit simultaneously by minimizing the following sum of squared differences between the observed ('obs') and calculated ('calc') values of experimental observables using an in-house Matlab program (MathWorks Inc., MA):

$$F = \alpha_1 \sum_i \sum_k \sum_l \sum_{j=1}^{2} \left(\frac{R_{2,\mathrm{eff}}^{\mathrm{obs},i,k,l,j} - R_{2,\mathrm{eff}}^{\mathrm{calc},i,k,l,j}}{\sigma_{R_{2,\mathrm{eff}}}^{\mathrm{obs},i,k,l,j}} \right)^2 + \alpha_2 \sum_i \sum_l \left(\frac{\delta_{\mathrm{ex}}^{\mathrm{obs},i,l} - \delta_{\mathrm{ex}}^{\mathrm{calc},i,l}}{\sigma_{\delta_{\mathrm{ex}}}^{\mathrm{obs},i,l}} \right)^2$$
$$+\alpha_3 \sum_i \sum_l \left(\frac{R_{2,1\mathrm{kHz}}^{\mathrm{obs},i,l} - R_{2,1\mathrm{kHz}}^{\mathrm{calc},i,l}}{\sigma_{R_{2,1\mathrm{kHz}}}^{\mathrm{obs},i,l}} \right)^2 + \alpha_4 \sum_i \sum_m \sum_{j=1}^{2} \left(\frac{R_{2,R_{1\rho}}^{\mathrm{obs},i,m,j} - R_{2,R_{1\rho}}^{\mathrm{calc},i,m,j}}{\sigma_{R_{2,R_{1\rho}}}^{\mathrm{obs},i,m,j}} \right)^2 \quad \text{(S6)}$$

where $R_{2,\mathrm{eff}}$ is the effective transverse relaxation rate in ^{15}N- and ^{13}Cα-CPMG relaxation dispersion experiments, δ_{ex} is the exchange-induced chemical shift, $R_{2,1\mathrm{kHz}}$ is the transverse relaxation rate

measured at a single spin-lock field strength of 1 kHz, and $R_{2,R1\rho}$ is the effective transverse relaxation rate measured in on-resonance ^{15}N-$R_{1\rho}$ relaxation dispersion experiments. The subscripts *i*, *j*, *k*, *l*, and *m* refer to the residue number, ^{1}H spectrometer frequency (600 and 900 MHz for CPMG, and 600 and 800 MHz for $R_{1\rho}$ dispersions), CPMG field strength, htt^{NT}Q$_7$ concentration, and on-resonance spin-lock RF field strength, respectively. The empirically determined factors α_1, α_2, α_3, and α_4, used to appropriately weight the different data types have numerical values of $\alpha_1 = \alpha_4 = 5$ and $\alpha_2 = \alpha_3 = 1$.

The set of global variable parameters in minimization of the target function in Eq. S6 thus comprised: $\{k_1, k_{-1}, k_2, k_{-2}, k_3, k_{-3}\}$ (see Eq. S3.1-3.2 for the relationship between the association rate constants k_1, k_2 and k_3 and the apparent pseudo-first order rate constant k_1^{app}, k_2^{app} and k_3^{app}, respectively). The space of local (residue-specific) variable parameters typically included: $\{[R_{2,CPMG}^{N,900}; R_{2,CPMG}^{N,600}; R_{2,CPMG}^{CA,900}; R_{2,CPMG}^{CA,600}]_{conc1}, [R_{2,\rho}^{N,800}; R_{2,\rho}^{N,600}]_{conc2}, \Delta\varpi^B, \Delta\varpi^C, \varphi^N, \varphi^{CA}\}$, where the vectors of parameters enclosed in square brackets are distinct (separate) for each concentration of htt^{NT}Q$_7$ (denoted by the subscripts 'conc$_n$', $n \in \{1,2\}$; see Table 1 in the main text for the values of htt^{NT}Q$_7$ concentrations used for CPMG and $R_{1\rho}$ dispersion measurements); the superscripts denote the type of nucleus ('N' = ^{15}N; 'CA' = ^{13}Cα) and ^{1}H spectrometer frequency; the subscripts denote the type of experimental measurement ('CPMG' or 'ρ' - $R_{1\rho}$), and the remaining notation as above. For simultaneous fitting of CPMG and $R_{1\rho}$ relaxation dispersion data, the values of Ω^B and $\Omega^C = \Omega^D$ (Eq. S5) were re-cast in terms of $\Delta\varpi$ values (as the difference between the offset from the carrier of the minor (B, C/D) and the major state (A), where the latter is readily available from the spectra. The uncertainties in the values of the optimized parameters, corresponding to confidence intervals of ± 1 standard deviation were determined from the Jacobian matrix of the non-linear fit. Convergence of the solution was confirmed by varying initial values for all parameters and obtaining the same solution within reported uncertainties.

Global fitting of ^{13}Cα data for 10 htt^{NT}Q$_7$ residues and ^{15}N data for 7 htt^{NT}Q$_7$ residues (see Table 1 main text) was performed using the expression for the target function in Eq. S6. This analysis was futher extended *a posteriori* to include exchange-induced chemical shifts for ^{1}H$_N$, ^{1}Hα, and ^{13}Cβ nuclei of htt^{NT}Q$_7$, as well as for an additional set of ^{15}N and ^{13}Cα chemical shifts of the residues that do not show significant relaxation dispersion. Although the values of $\Delta\varpi$ for all the states in Scheme S1 cannot be determined with certainty due to the absence of significant relaxation dispersion for this subset of htt^{NT}Q$_7$ nuclei, approximate values of $\Delta\varpi^C = \Delta\varpi^D$ for these sites could nevertheless be extracted from the global fitting and are reported in Figure 6A (main text) and Fig. S9.

It is also worth noting that the optimized ^{13}Cα and ^{15}N R_2^A values for the major monomeric state are concentration independent within error, as expected since the highest concentration (1 mM) used in the relaxation dispersion measurements only corresponds to 2.7 mg/ml of protein and, hence, would not be expected to affect the viscosity of the solution in any significant way. This is also borne out by the data in Figure 2B which show no difference in ^{15}N-R_2^{1KHz} values between 10

μM $htt^{NT}Q_7$ (monomeric, green circles) and 1 mM (Met^7O)-$htt^{NT}Q_7$ (also monomeric, blue circles) samples, and between 10 μM (grey circles) and 1 mM (black circles) htt^{NT} (monomeric) samples.

Effects of the faster tetramerization process on CPMG relaxation dispersion and exchange-induced shifts. The effects of the (faster) tetramerization process (A ↔ C ↔ D; scheme S1) on $^{13}C\alpha$ CPMG relaxation dispersion profiles at a single total protein concentration $[P]_T$ and on $^{13}C\alpha$ exchange-induced chemical shifts δ_{ex} as a function of $[P]_T$ (at a spectrometer frequency of 900 MHz) are illustrated in Figure S10. The black curves in Figures S10A ($^{13}C\alpha$-CPMG) and S10B ($^{13}C\alpha$-δ_{ex}) are simulated for a two-state dimerizing system (B ↔ A process only) using the parameters of exchange that are approximately the same as reported in the main text with $\Delta\omega_B$ = 2.0 ppm and R_2^A = 9 s^{-1}), while the profiles shown in red are calculated for the full four-state exchanging system (B ↔ A ↔ C ↔ D; Scheme S1) using the parameters of exchange very similar to those reported in the current work with $\Delta\omega_C = \Delta\omega_D$ = 3.0 ppm. The main effect of inclusion of a faster tetramerization process (A ↔ C ↔ D) on the CPMG profiles is the 'elevation' of the profile by ~10 s^{-1} (*cf.* the red and black curves in Figure S10A). It is therefore not surprising that the CPMG relaxation dispersion profiles of $htt^{NT}Q_7$ at 0.4, 0.75 and 1 mM can be reproduced by a two-state model, with the resulting values of R_2^A significantly overestimated at higher peptide concentrations (up to a factor of ~2 at $[P]_T$ = 1.0 mM for many residues). These spurious transverse relaxation rates are the direct consequence of the omission of a much faster process from the model of exchange. The effects of this fast process which results in the formation of a tetramer on δ_{ex} are much more profound (*cf.* the black and red in Figure S10B) implying that the chemical shift changes cannot be explained by a simple two-state dimerization process (B ↔ A). Further, the curvature appearing in the δ_{ex} plot at higher peptide concentrations (red curve in Figure S10B) is characteristic of the formation of higher order oligomers.

The curves shown in magenta and green in Figures S10A and B correspond to the same simulated kinetic scheme as shown in red, but with the tetramerization process (A ↔ C ↔ D) in Scheme S1 accelerated and slowed-down ~10-fold, respectively. When the tetramerization process becomes very fast compared to the rate of 'non-productive' dimerization (B ↔ A), the CPMG profiles become almost indistinguishable from those of the two-state dimerization process (*cf.* the magenta and black curves in Figure S10A), whereas the profiles of δ_{ex} *vs.* $[P]_T$ in the same limit are virtually identical to those obtained experimentally in the present work (*cf.* the magenta and red curves in Figure S10B). At the same time, tetramerization on a time-scale similar to that of 'non-productive' dimerization (*i.e.* much slower), would not reproduce either the CPMG relaxation dispersion or δ_{ex} experimental data (*cf.* the green and red curves in Figures S10A and S10B).

SI PRE-Based Structure Modeling and Xplor-NIH Calculations

The φ and ψ values of the subunit helix were set to ideal values of -57.8° and -47°, respectively. For each residue mutated to cysteine and subsequently linked to a MTSL (R1) tag, sidechain and tag coordinates were represented by three copies. Since the multiple copies are used to represent the conformational space sampled by the R1 tag, atomic overlap between R1 copies in a given subunit is allowed (25). Further, atomic overlap is also allowed between R1 copies of different subunits as the labeling scheme employed ensures that effectively either only one subunit of the tetramer is paramagnetically labeled or two subunits within the dimer. Structures were calculated with Xplor-NIH's strict symmetry facility (26) such that explicit degrees of freedom were allowed for only one subunit, the protomer, and those of its dimer partner were computed by making a rotation of 180° about the *z*-axis. The full tetramer was then formed by rotating these dimer coordinates 180° about the axis along the line described by $z = 0, x = y$. The allowed degrees of freedom of the protomer were the sidechain and terminal backbone torsion angles. Backbone torsion angles of residues 3-17 were held rigid throughout.

The structure calculation consisted of an initial subunit docking phase (described below), followed by high temperature dynamics and molecular dynamics simulated annealing from 3000 K to 25 K in 12.5 K increments. Energy scales on the energy terms were geometrically ramped from smaller to larger values during simulated annealing to allow greater sampling at earlier stages. The small energy scales were also used during the high-temperature molecular dynamics and docking phases, except where specified below. The resulting lowest 10 energy structures were then subject to a second round of simulated annealing refinement using the same protocol.

The PREPot (25) term was used to represent the PRE data by the Solomon-Blombergen equation, explicitly including the effect of tag conformational inhomogeneity. Because of the small ratio of tagged subunits to untagged (1 to 40), one can safely make the assumption that each tetramer from which a PRE could be measured contains only a single tagged copy, and that the PRE signal is the sum of contributions from this labeled subunit to the three others.

The experimental NMR data are indicative of a dimer/tetramer mixture, but the structural information of the restraints is limited, so we chose to make the assumption that the dimer species has exactly the same structure as half of the tetramer. Under this assumption, the dimer/tetramer ratio can be altered by multiplying the associated $1/r^6$ distance contributions to the computed PRE value by appropriate factors. The overall PRE correlation time τ_C is given by $\tau_R\tau_s/(\tau_R + \tau_s)$, where τ_s and τ_R are the electron and rotational correlation times, respectively. For a nitroxide spin label, $\tau_s > 100$ ns, and hence τ_C is dominated by τ_R which we assume to be 50% larger for the tetramer relative to the dimer. (Note the tetramer is not double the molecular weight of the dimer since, under the experimental conditions of the PRE measurements, the tetramer contains either only one GB1 fusion subunit or at best two GB1 fusion subunits within one of the dimers making up the tetramer.) Each measured PRE is the sum of 3 components corresponding to one set of distances between tag copies and dimer partner and two sets of distances between tag copies and the two subunits corresponding to the second dimer. We can simulate the effect of the presence of a dimer/tetramer mixture by scaling the distances to the second and third contributions by a factor λ.

In this case $\lambda = w_\lambda/(2 - \omega_\lambda)$, where ω_λ is the fractional population of tetramer. The PREPot term was modified to allow this sort of distance contribution scaling. Calculations were performed with 11 values of ω_λ corresponding to populations of the tetramer sampled uniformly between 0 and 100%.

As there is an unknown prefactor in the PRE calculation due to the unknown absolute concentration of singly labeled tetramers giving rise to the observed PRE, a correlation target function was added to the PREPot term, where the target energy is:

$$E_{\mathrm{PRE}} = w_{\mathrm{PRE}}(1 - C_{\mathrm{PRE}}) \tag{S7}$$

with

$$C_{\mathrm{PRE}} = \frac{\sum_i (\Gamma_i - \overline{\Gamma})(\Gamma_i^{\mathrm{expt}} - \overline{\Gamma}^{\mathrm{expt}})}{\sqrt{\sum_i (\Gamma_i - \overline{\Gamma})^2 (\Gamma_i^{\mathrm{expt}} - \overline{\Gamma}^{\mathrm{expt}})^2}} \tag{S8}$$

where w_{PRE} is an energy scaling constant, Γ_i the PRE value computed from the coordinates, $\overline{\Gamma}$ the associated mean value, Γ_i^{expt} the observed PRE value, and $\overline{\Gamma}^{\mathrm{expt}}$ the corresponding mean value. After the structures were calculated the calculated PREs were rescaled to optimize the absolute fit between calculated and observed PREs. During the first molecular dynamics and simulated annealing portions of the calculation, the energy scale on this term was ramped from 10 to 1000, and from 10 to 2000 in the simulated annealing refinement stage.

The RepelPot term was used to prevent atomic overlap in the tetramer, with tag interactions as described above with energy scale ramped from 0.004 to 4. The gyration volume term (27) acted on the central portion of the helix (residues 3-15) to bias the packing of the tetramer toward standard protein density with an energy scale ramp of 0.01 to 10. The low-resolution residue affinity contact term (28) was used to encourage hydrophobic and complimentary charge interactions with an energy scale ramp of 1 to 10. The force field was rounded out with terms acting only on the protomer: TorsionDB (29) to obtain representative torsion angle values, and covalent bond length, bond angle and improper angle terms.

The structure calculation proceeded as follows. The orientation and position of the protomer, and its non-fixed torsion angles were randomized. Then, gradient minimization was used to optimize first the packing (gyration volume and residue affinity) and nonbonded terms, followed by optimizing the total energy. This second gradient optimization proceeded in three stages with the energy scale on the PRE term progressively taking the values 10^{-20}, 10^{-10} and 10^{-5}. This gradient docking procedure was repeated five times with different random initial coordinates, and the lowest total energy structure used for the subsequent simulated annealing calculation.

The docking procedure was followed by a round of simulated annealing refinement, consisting of the calculation of ten structures from each of the ten lowest energy structures of the docking stage. In refinement, the final scale on the PRE energy was increased from 1000 to 2000, but the simulated annealing protocol was otherwise identical to that used for docking. The final 10 lowest energy structures were used for analysis.

Validation of PRE-based structures. Independent validation of the dimer unit of the PRE-based structures was carried out by comparing mean distances between S15C-R1 spin labels within the dimer unit with those obtained from analysis of experimental DEER EPR data on dimeric GB1-htt^{NT}Q$_7$ (S15C-R1). The latter forms a dimer and does not progress to a tetramer (see Fig. 10B and C and Fig. S15). Since we have no experimental data on the conformations of surface exposed side chains within the tetramer, the side chain coordinates were removed beyond the Cβ position and the program MMMv2013.2 (30) was used to generate rotamer distribution for each spin-label pair from which predicted $P(r)$ distributions and mean distances were obtained.

SI References

1. Aiken CT*, et al.* (2009) Phosphorylation of threonine 3: implications for Huntingtin aggregation and neurotoxicity. *J Biol Chem* 284(43):29427-29436.
2. Lundström P*, et al.* (2007) Fractional ^{13}C enrichment of isolated carbons using [1-^{13}C]- or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone Cα and side-chain methyl positions in proteins. *J Biomol NMR* 38(3):199-212.
3. Anthis NJ & Clore GM (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci* 22(6):851-858.
4. Cavanagh J, Fairbrother W, Palmer AI, & Skelton N (2018) *Protein NMR Spectroscopy : Principles and Practice* (Elsevier, Waltham, MA) 3rd edition.
5. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277-293.
6. Lakomek N-A, Ying J, & Bax A (2012) Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* 53(3):209-221.
7. Iwahara J, Tang C, & Marius Clore G (2007) Practical aspects of ^{1}H transverse paramagnetic relaxation enhancement measurements on macromolecules. *J Magn Reson* 184(5):185-195.
8. Hansen DF, Vallurupalli P, & Kay LE (2008) An improved ^{15}N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* 112(19):5898-5904.
9. Hansen DF, Vallurupalli P, Lundström P, Neudecker P, & Kay LE (2008) Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: How well can we do? *J Am Chem Soc* 130(8):2667-2675.
10. Vuister GW & Bax A (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broad-band ^{13}C decoupling. *J Magn Reson* 98(2):428-435.
11. Kontaxis G, Clore GM, & Bax A (2000) Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J Magn Reson* 143(1):184-196.
12. Jeschke G (2012) DEER distance measurements on proteins. *Annu Rev Phys Chem* 63419-446.

13. Jeschke G*, et al.* (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *App Magn Reson* 30(3-4):473-498.
14. Schmidt T, Ghirlando R, Baber J, & Clore GM (2016) Quantitative resolution of monomer-dimer populations by inversion modulated DEER EPR spectroscopy. *ChemPhysChem* 17(19):2987-2991.
15. Altenbach C, Lopez CJ, Hideg K, & Hubbell WL (2015) Exploring structure, dynamics, and topology of nitroxide spin-labeled proteins using continuous-wave electron paramagnetic resonance spectroscopy. *Methods Enzymol* 564:59-100.
16. Zhao H, Brautigam CA, Ghirlando R, & Schuck P (2013) Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* Chapter 20, Unit20 12.
17. Zhao H*, et al.* (2013) Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. *Anal Biochem* 437(1):104-108.
18. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78(3):1606-1619.
19. Cole JL, Lary JW, T PM, & Laue TM (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol* 84:143-179.
20. Schuck P (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* 320(1):104-124.
21. Ghirlando R (2011) The analysis of macromolecular interactions by sedimentation equilibrium. *Methods* 54(1):145-156.
22. Sekhar A, Bain AD, Rumfeldt JA, Meiering EM, & Kay LE (2016) Evolution of magnetization due to asymmetric dimerization: theoretical considerations and application to aberrant oligomers formed by apoSOD1(2SH). *Phys Chem Chem Phys* 18(8):5720-5728.
23. Rennella E, Sekhar A, & Kay LE (2017) Self-assembly of human Profilin-1 detected by Carr–Purcell–Meiboom–Gill nuclear magnetic resonance (CPMG NMR) spectroscopy. *Biochemistry* 56(5):692-703.
24. McConnell HM (1958) Reaction rates by nuclear magnetic resonance. *J Chem Phys* 28(3):430-431.
25. Iwahara J, Schwieters CD, & Clore GM (2004) Ensemble approach for NMR structure refinement against ^{1}H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J Am Chem Soc* 126(18):5879-5896.
26. Schwieters CD, Bermejo GA, & Clore GM (2018) Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci* 27(1):26-40.
27. Schwieters CD & Clore GM (2008) A pseudopotential for improving the packing of ellipsoidal protein structures determined from NMR data. *J Phys Chem B* 112(19):6070-6073.
28. Ryabov Y, Suh JY, Grishaev A, Clore GM, & Schwieters CD (2009) Using the experimentally determined components of the overall rotational diffusion tensor to restrain molecular shape and size in NMR structure determination of globular proteins and protein-protein complexes. *J Am Chem Soc* 131(27):9522-9531.
29. Bermejo GA, Clore GM, & Schwieters CD (2012) Smooth statistical torsion angle potential derived from a large conformational database via adaptive kernel density estimation improves the quality of NMR protein structures. *Protein Sci* 21(12):1824-1836.
30. Polyhach Y, Bordignon E, & Jeschke G (2011) Rotamer libraries of spin labelled cysteines for protein studies. *Phys Chem Chem Phys* 13(6):2356-2366.
31. Kjaergaard M, Brander S, & Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J Biomol NMR* 49(2):139-149.

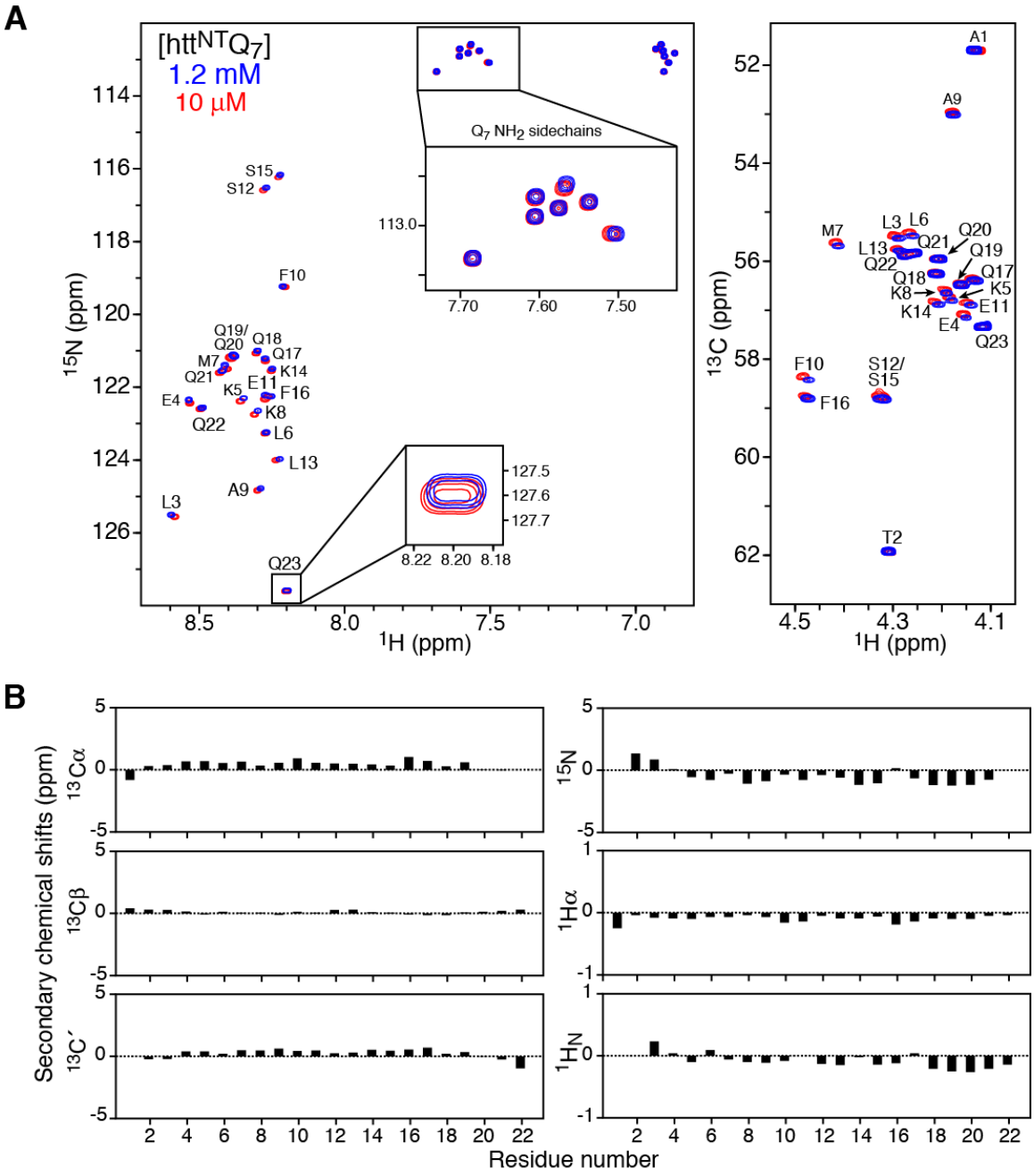


Figure S1. Spectral characterization of htt^{NT}Q$_7$. ^{1}H-^{15}N and ^{1}H-^{13}Cα correlation spectra of htt^{NT}Q$_7$. (A) 900 MHz ^{1}H-^{15}N HSQC (left) and ^{1}H-^{13}C constant time HSQC (right) spectra of 10 µM (red) and 1.2 mM (blue) htt^{NT}Q$_7$ at 10 °C in 20 mM sodium phosphate, pH 6.5, 50 mM NaCl. Insets in the ^{1}H-^{15}N HSQC spectrum highlight small concentration-dependent, exchange induced shifts for Q23 and the absence of exchange-induced shifts for the cross-peaks of most of the glutamine side chain NH_2 groups which were therefore used as an internal control to monitor exchange-induced shifts as a function of concentration. (B) Secondary backbone shifts. Reference random coil shifts are from(31).

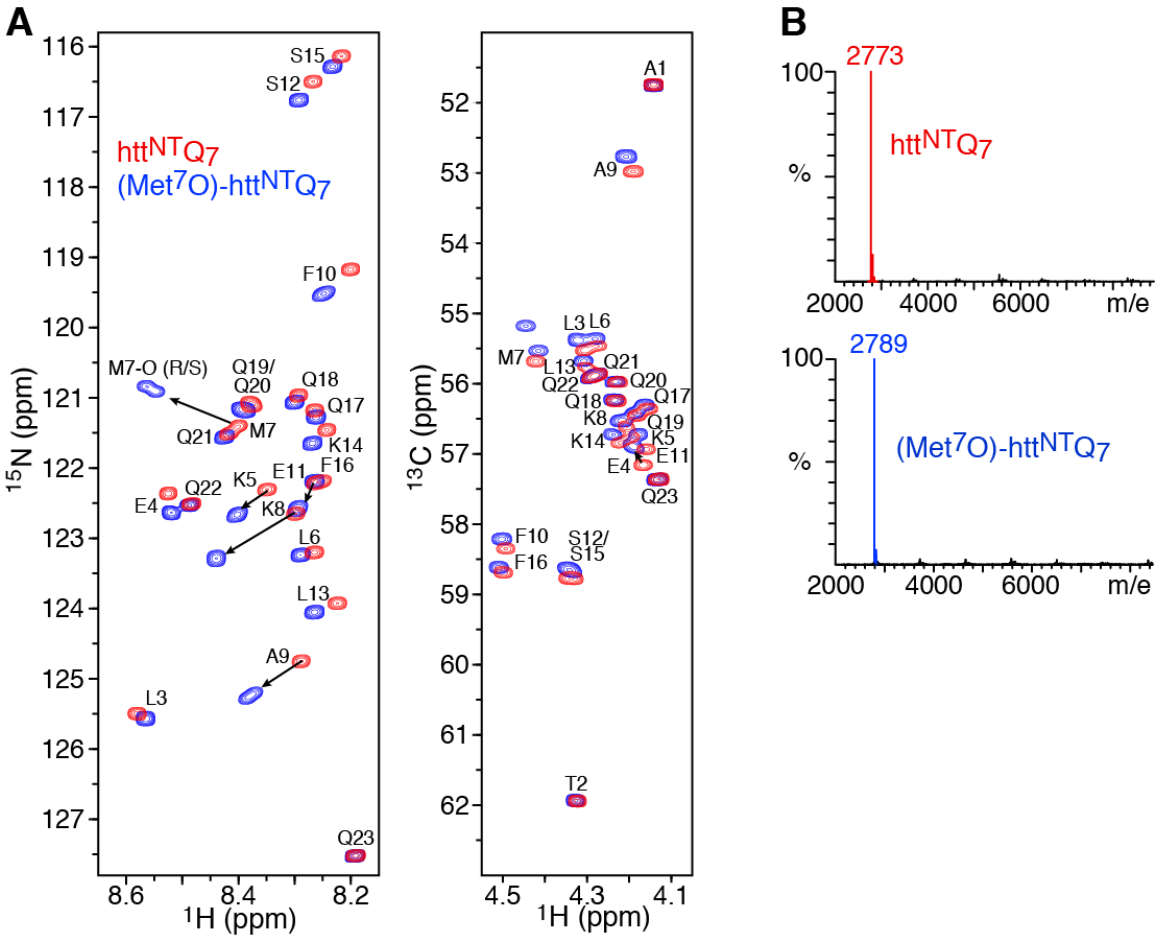


Figure S2. Comparison of native and methionine sulfoxide forms of htt$^{NT}Q_7$. (A) ^{1}H-^{15}N (left) and ^{1}H-$^{13}C\alpha$ (right) correlation spectra at 900 MHz, and (B) mass spectra obtained before and after the addition of 0.5% H_2O_2. Native htt$^{NT}Q_7$ is displayed in red and (Met^{7}O)-htt$^{NT}Q_7$ in blue. The peptide concentration used for the NMR spectra was 1 mM (in 20 mM sodium phosphate, pH 6.5, 50 mM NaCl), and the temperature was 10°C.

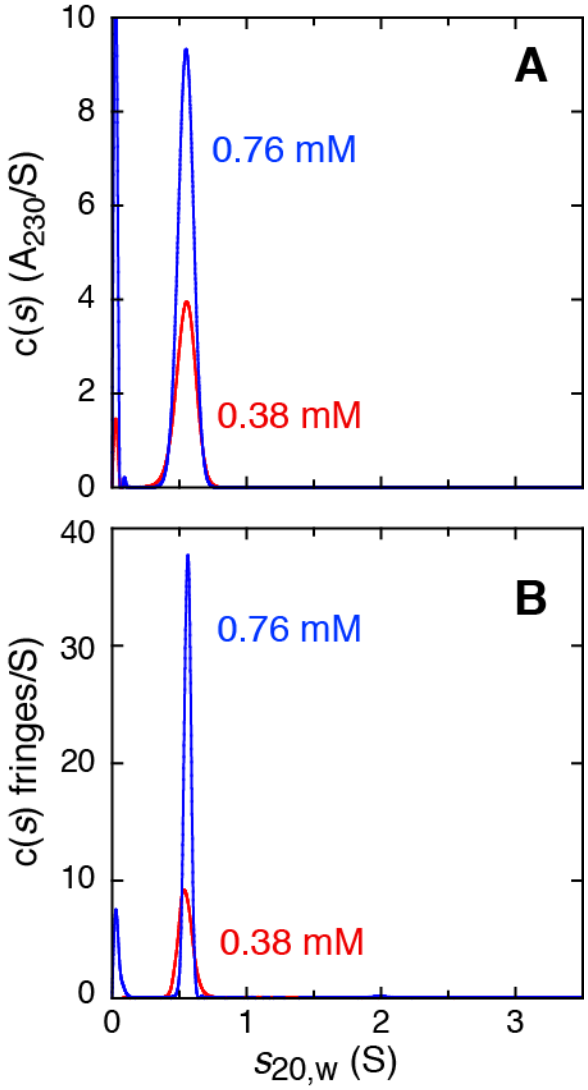


Figure S3. Sedimentation velocity analytical ultracentrifugation characterization of ^{15}N-labeled htt^{NT}Q$_7$. (A) Sedimentation velocity absorbance c(*s*) profiles for ^{15}N-labeled htt^{NT}Q$_7$ at 310 (red) and 620 μM (blue); (B) the corresponding interference c(*s*) profiles. The c(*s*) profiles show the presence of a single species at 0.55 S with an estimated molar mass of 3.0 kDa, confirming the presence of a monomer. This species accounts for ~99.8% of the absorbance signal, and both methods of detection do not show evidence for significant quantities of dimers or higher order oligomers. Data were recorded at 10 °C in NMR buffer (20 mM sodium phosphate, pH 6.5, 50 mM NaCl and 10% v/v D_2O).

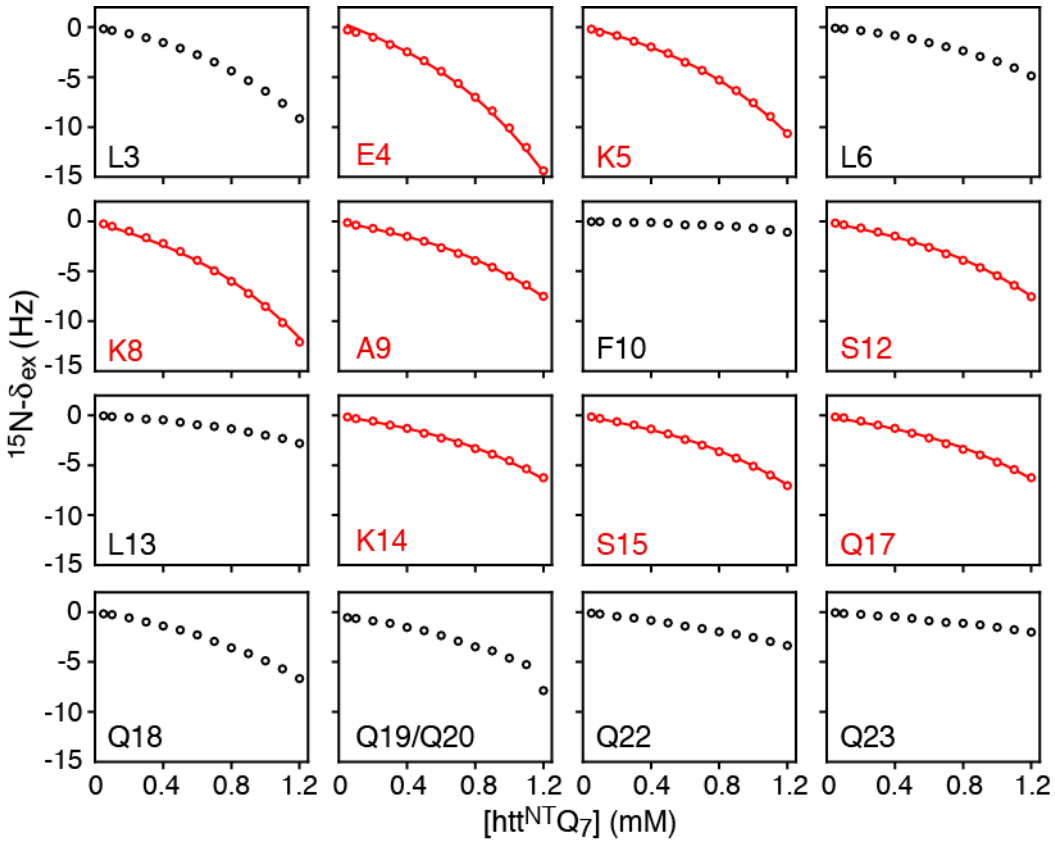


Figure S4. Concentration-dependent ^{15}N exchange-induced chemical shifts observed for uniformly ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ at 5 °C and 800 MHz. The panels highlighted in red indicate data used in the global kinetic analysis and fits, and the solid lines represent the best-fit curves. Note the ^{1}H-^{15}N cross-peaks of Q19 and Q20 are overlapped.

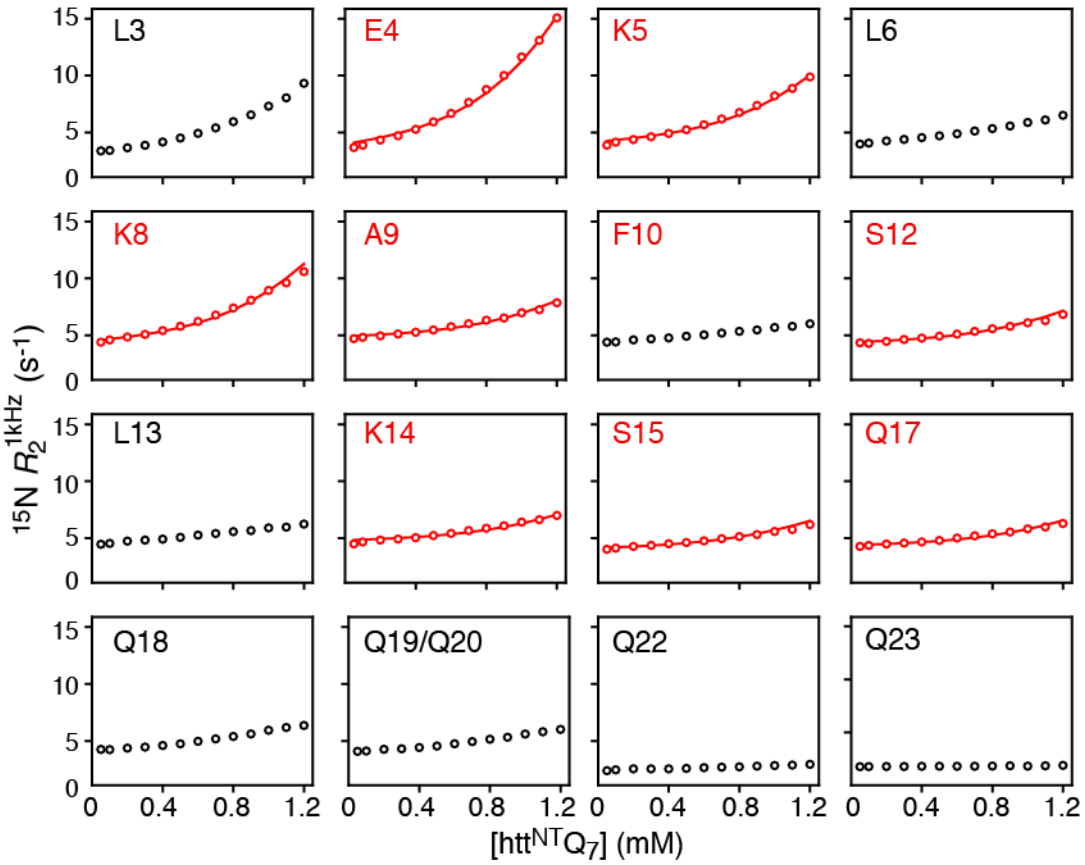


Figure S5. Concentration-dependence of ^{15}N-$R_2^{1\mathrm{KHz}}$ values observed for uniformly ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ at 5 °C and 800 MHz. The panels highlighted in red indicate data used in the global kinetic analysis and fits, and the solid lines represent the best-fit curves. Note the ^{1}H-^{15}N cross-peaks of Q19 and Q20 are overlapped. There is no measurable change in sample viscosity from the lowest (50 μM) to highest (1.2 mM) concentrations of htt^{NT}Q$_7$ employed, as evidenced by the observation that the ^{15}N-$R_2^{1\mathrm{KHz}}$ values measured for a 10 μM htt^{NT}Q$_7$ sample are the same as the corresponding values measured for a 1 mM (Met^{7}O)-htt^{NT}Q$_7$ sample (green versus blue circles, respectively, in Fig. 2B of the main text).

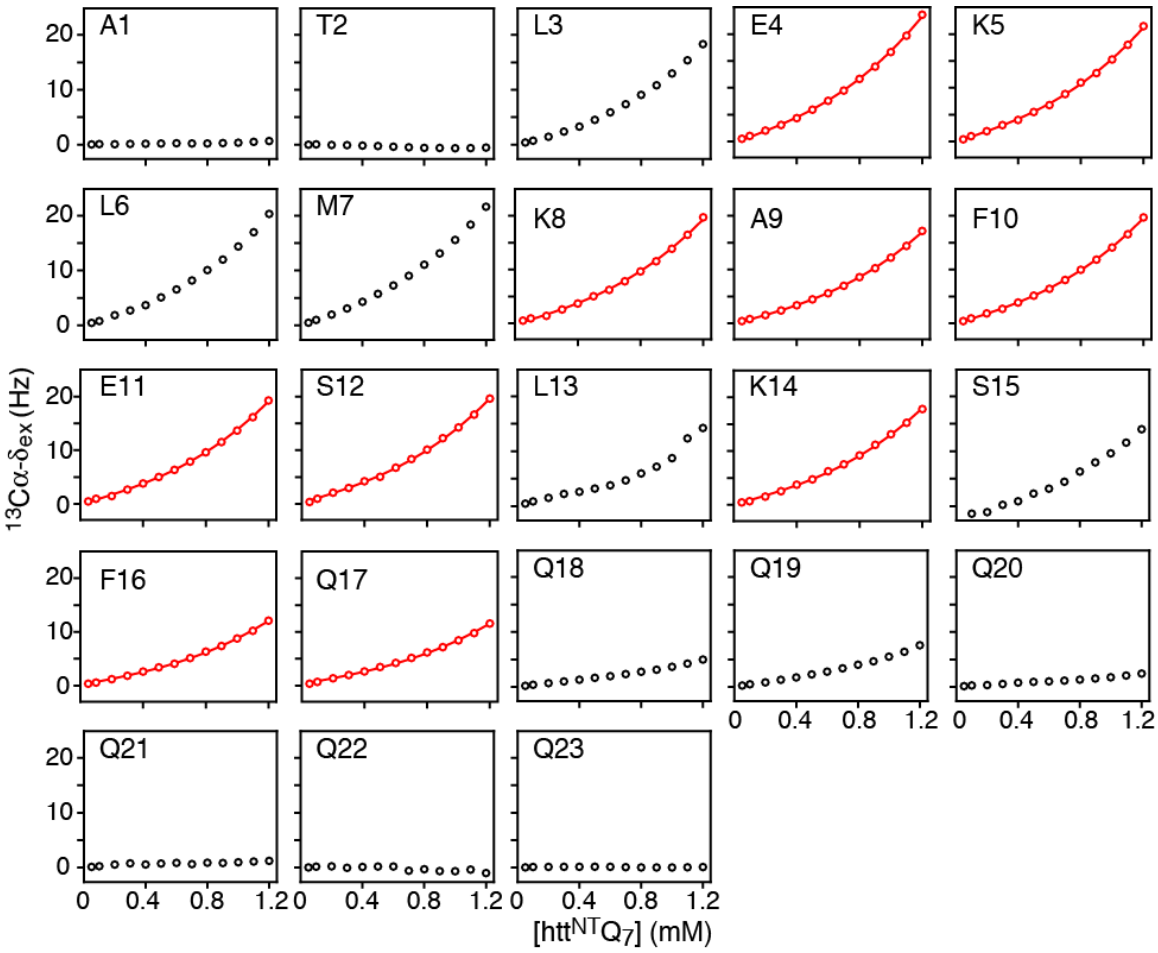


Figure S6. Concentration-dependent $^{13}C\alpha$ exchange-induced chemical shifts observed for uniformly $^{13}C/^{15}N$-labeled $htt^{NT}Q_7$ at 5 °C and 800 MHz. The panels highlighted in red indicate data used in the global kinetic analysis and fits, and the solid lines represent the best-fit curves.

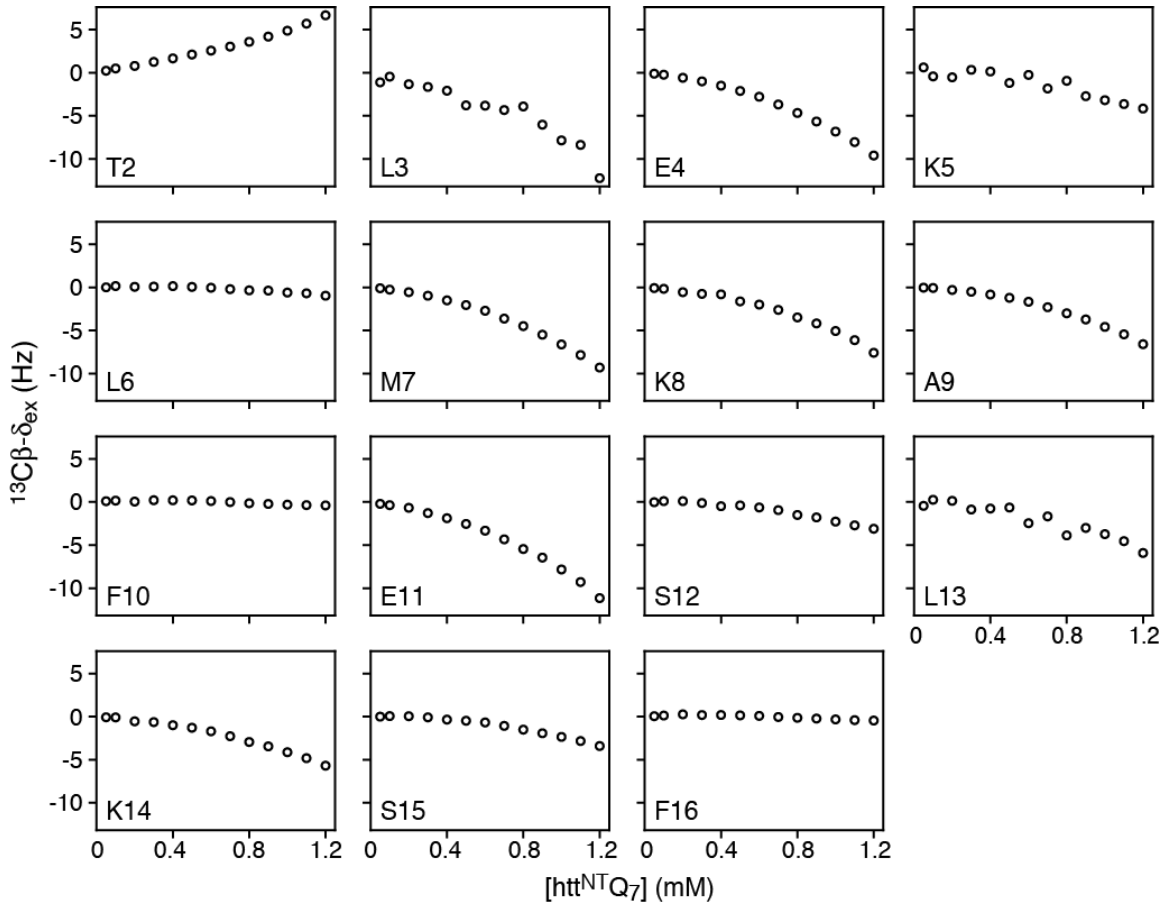


Figure S7. Concentration-dependent ^{13}Cβ exchange-induced chemical shifts observed for uniformly ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ at 5 °C and 800 MHz.

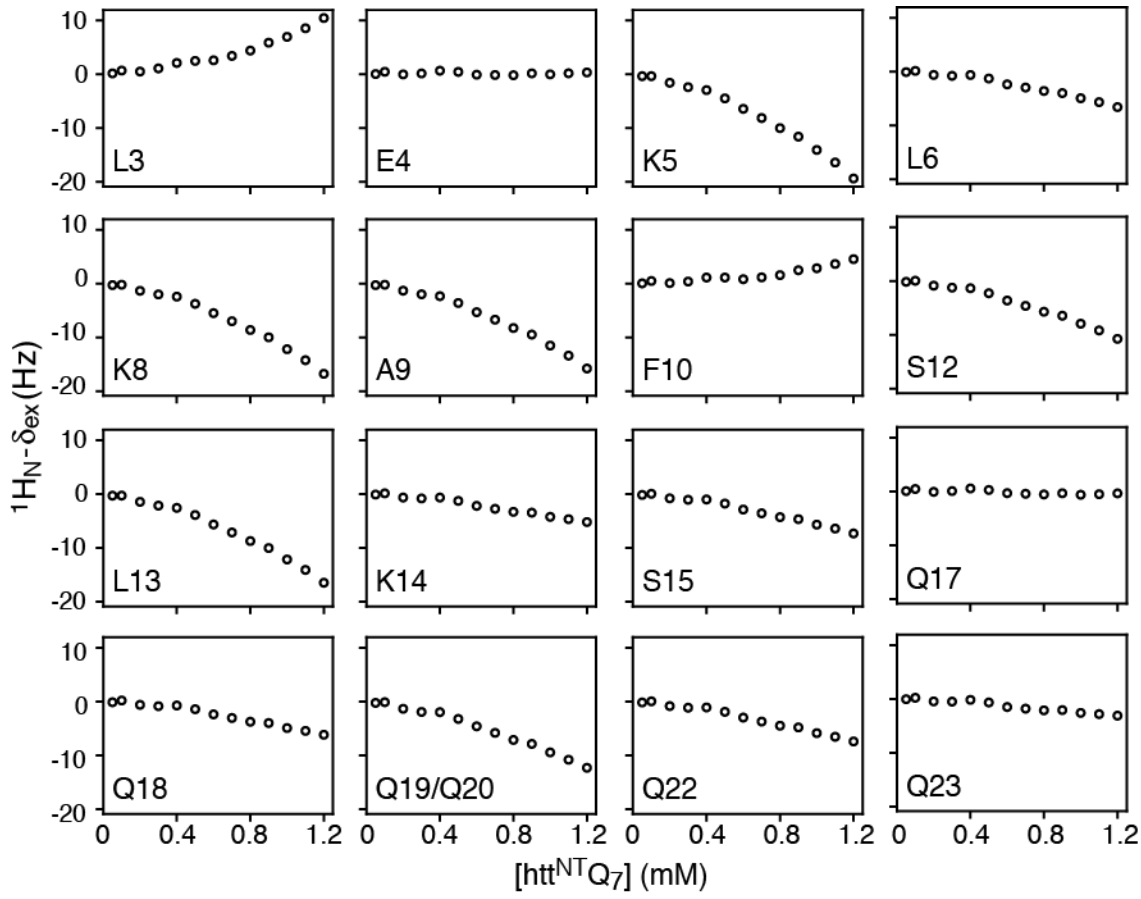


Figure S8. Concentration-dependent ^1H_N exchange-induced chemical shifts observed for uniformly $^{13}C/^{15}N$-labeled htt^{NT}Q$_7$ at 5 °C and 800 MHz.

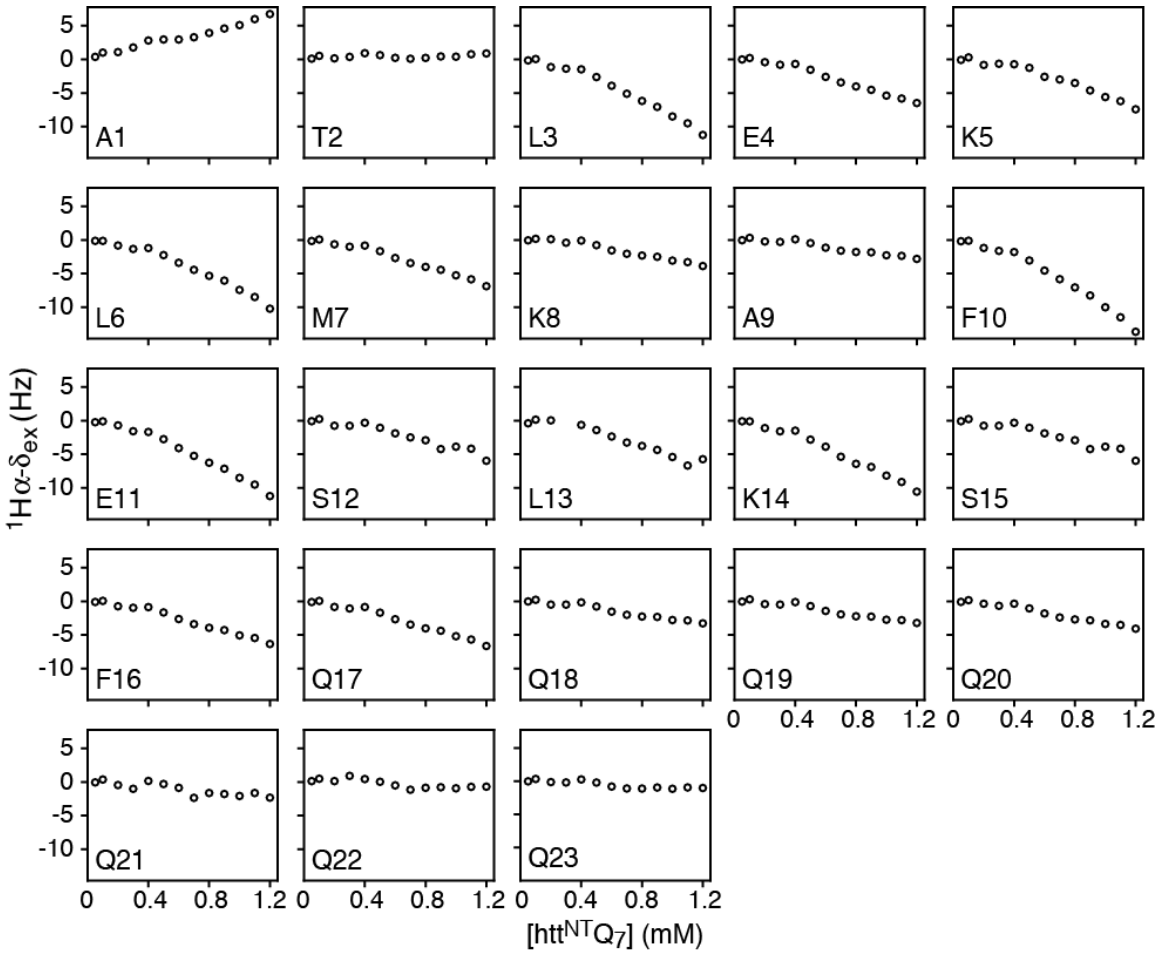


Figure S9. Concentration-dependent ^{1}Hα exchange-induced chemical shifts observed for uniformly ^{13}C/^{15}N-labeled htt^{NT}Q$_7$ at 5 °C and 800 MHz.

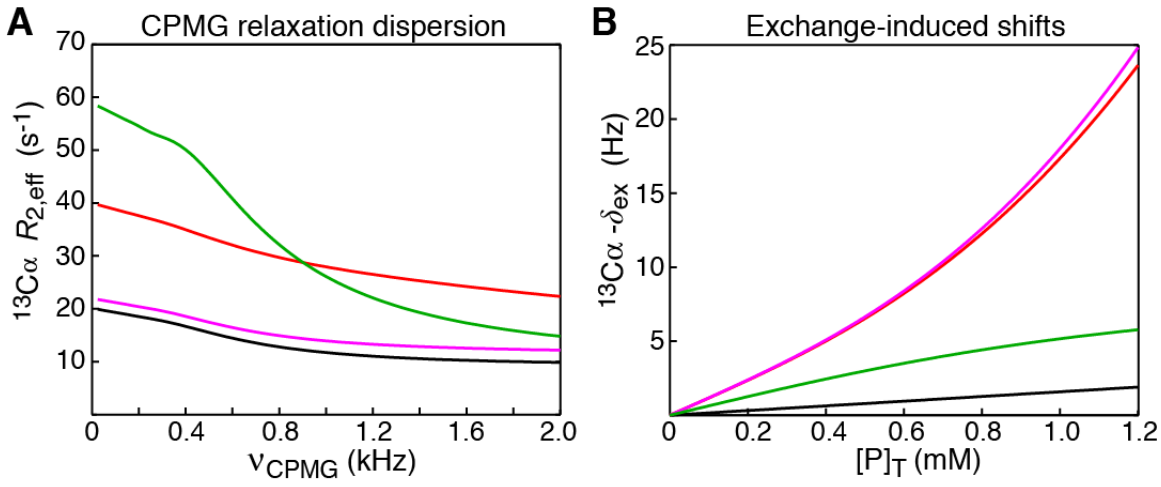


Figure S10. Comparison of $^{13}C\alpha$ CPMG relaxation dispersion profiles and $^{13}C\alpha$ exchange-induced chemical shifts, δ_{ex}, simulated for several limiting cases of the kinetic model shown in Scheme S1. (A) $^{13}C\alpha$ CPMG relaxation dispersion profiles (900 MHz) for an (isolated) two-state dimerizing system (B ↔ A only) calculated for a total protein concentration $[P]_T$ = 1.0 mM with k_3 = 1000 $M^{-1}s^{-1}$ and k_{-3} = 2,500 s^{-1} (p_B = 0.8%; $\Delta\omega_B$ = 2.0 ppm; R_2^A = 9 s^{-1}) is shown in *black*; the same profile calculated for the full four-state exchanging system (B ↔ A ↔ C ↔ D; Scheme S1) with k_1 = 3 x 10^5 $M^{-1}s^{-1}$, k_{-1} = 40,000 s^{-1}, k_2 = 2 x 10^9 $M^{-1}s^{-1}$ and k_{-2} = 45,000 s^{-1} (p_C =1.5%; p_D = 1%; $\Delta\omega_C$ = $\Delta\omega_D$ = 3.0 ppm) is shown in *red*; the *magenta* and *green* curves show profiles for the same full four-state scheme as the red curve, but with the tetramerization process (A ↔ C ↔ D) either accelerated or slowed-down ~10-fold, respectively (i.e. k_1, k_{-1}, k_2 and k_{-2} are multiplied or divided by 10 to preserve the values of p_C and p_D). (B) $^{13}C\alpha$ exchange-induced chemical shifts δ_{ex} (Hz) as a function of total protein concentration $[P]_T$ (mM) (900 MHz) with the same rate constants and color coding as in panel A.

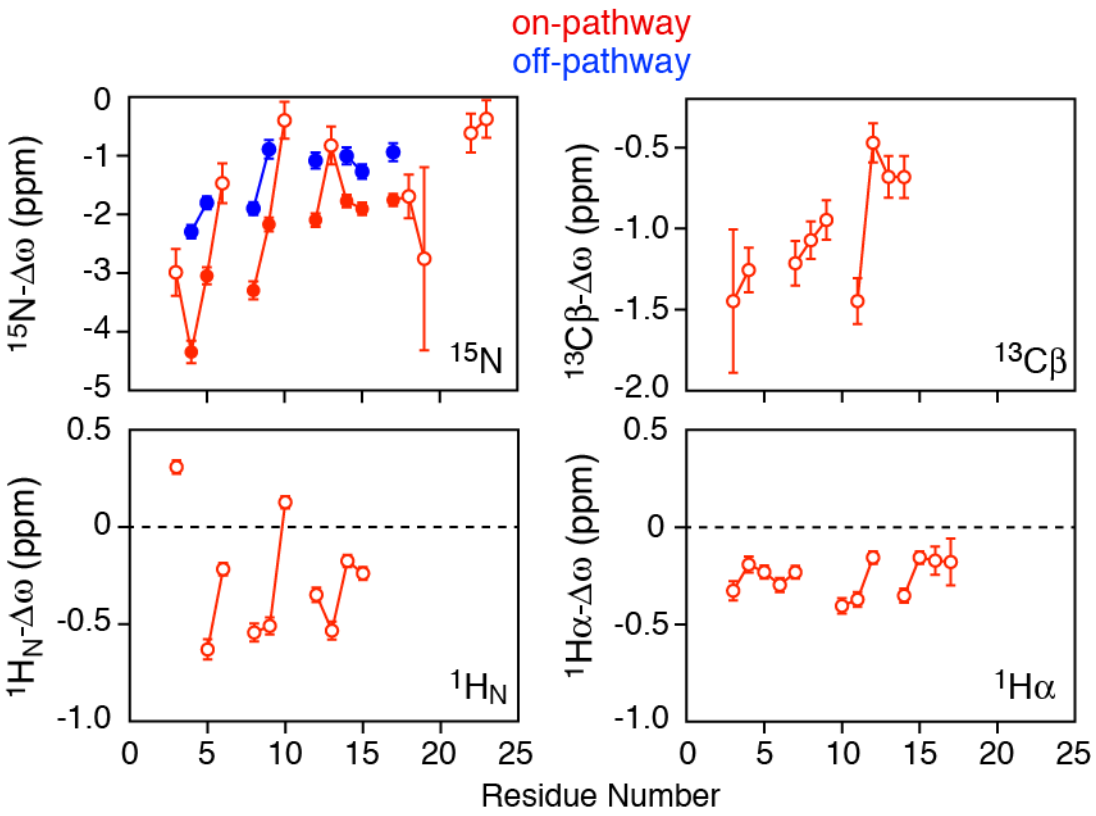


Figure S11. Calculated differences in chemical shifts (Δω) between the minor oligomeric species and the major monomeric state of $htt_{NT}Q_7$ at 5°C. Filled-in circles indicate residues and data used in the global fits with the on-pathway states (dimer and tetramer with shifts assumed to be equal) in red, and the off-pathway dimer in blue; open circles are shifts for the on-pathway states (dimer and tetramer with shifts assumed to be equal) calculated from the exchange-induced shift data as a function of $htt^{NT}Q_7$ concentration using the rate constants determined in the global fits.

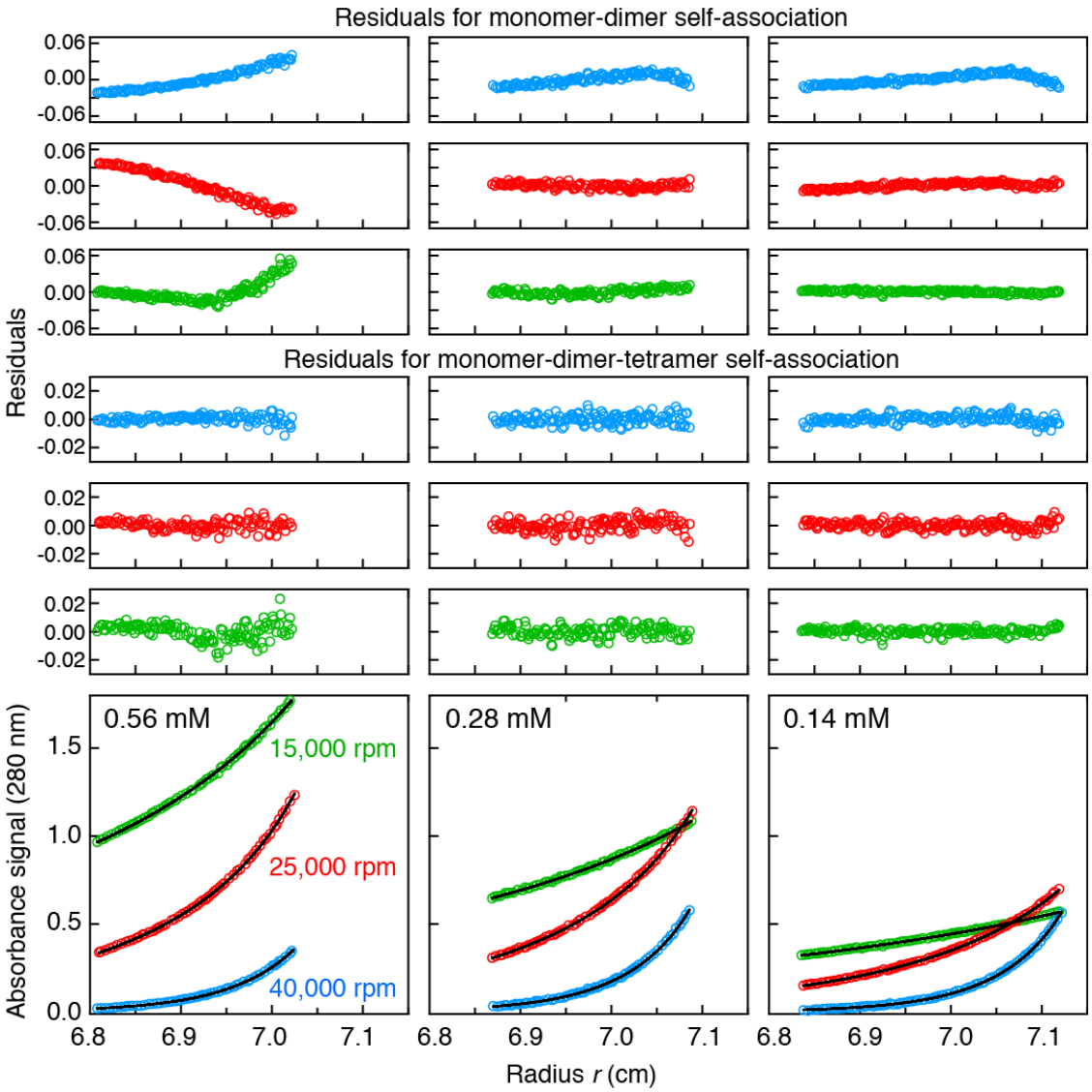


Figure S12. Absorbance sedimentation equilibrium AUC analysis of GB1-htt^{NT}Q7. Absorbance sedimentation equilibrium profiles are shown in terms of absorbance (at 280 nm) versus radius for data collected at concentrations of 0.56, 0.28 and 0.14 mM and at spinning speeds of 15,000 (green), 25,000 (red) and 40,000 (blue) rpm. Data were analyzed globally, along with lower concentration samples as well as data recorded at 250 nm absorbance, in terms of reversible monomer-dimer- tetramer and monomer-dimer self-association. Best-fits to the reversible monomer-dimer-tetramer model are shown as black lines through the experimental points. The corresponding distributions of the residuals for the monomer-dimer and monomer-dimer-tetramer models are shown in the plots above, from which it is clear that there are large systematic deviations in the residuals for the monomer-dimer model. For the monomer-dimer-tetramer model, excellent fits were obtained with r.m.s.d. values ranging from 0.0022 to 0.0042 absorbance units, and values for the monomer-dimer and dimer-tetramer equilibrium dissociation constants of 9^{+3}_{-3} mM and 8^{+14}_{-2}μM, respectively. Data were recorded at 20°C in 50 mM sodium phosphate, pH 6.5, 100 mM NaCl.

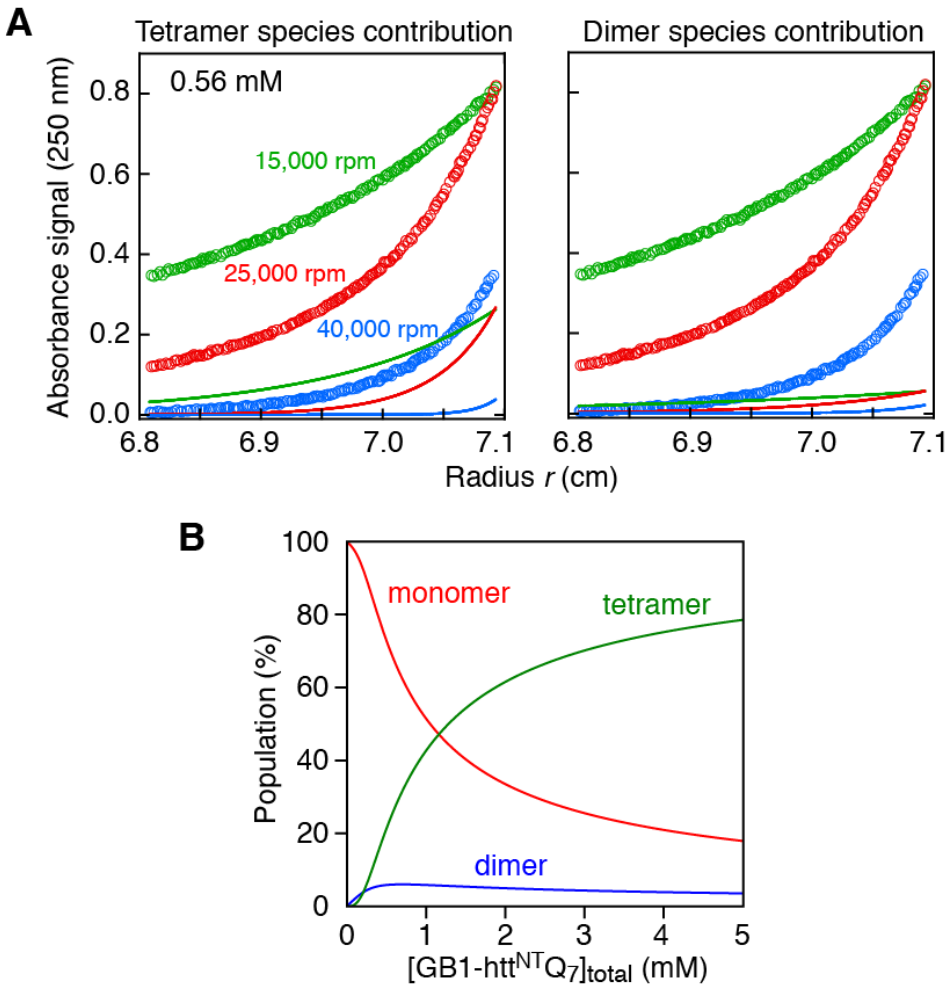


Figure S13. Contribution of tetramer and dimer species of GB1-htt$^{NT}Q_7$ to sedimentation equilibrium AUC profiles and concentration dependence of species populations. (A) Absorbance sedimentation equilibrium profiles (circles) for 0.56 mM GB1-htt$^{NT}Q_7$ shown in terms of absorbance at 250 nm versus radius together with the expected distribution of tetramer (left panel) and dimer (right panel) shown as continuous lines and based on the global best-fit of a reversible monomer-dimer-tetramer self-association model to the sedimentation equilibrium profiles (see Fig. S12). Note that the model used to analyze the sedimentation equilibrium data does not account for dimer that is incompetent for tetramerization but clearly shows the presence of higher-order oligomer (i.e. tetramer) at the highest loading concentration (0.56 mM) studied. (B) Simulation of the species populations (in monomer units) as a function of concentration calculated from the equilibrium dissociation constants obtained from sedimentation equilibrium AUC (Fig. S12 and Table 2 in main text)

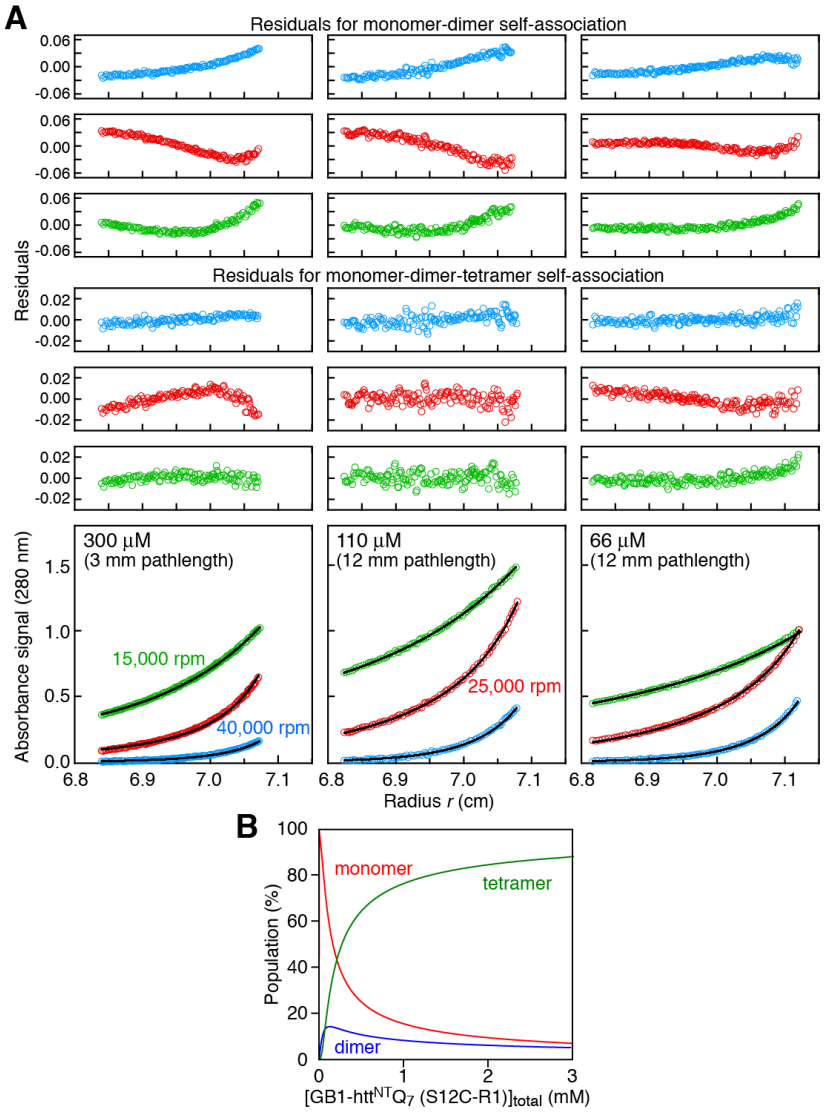


Figure S14. Sedimentation equilibrium AUC analysis of GB1-htt^{NT}Q$_7$ (S12C-R1). (A) Absorbance sedimentation equilibrium profiles shown in terms of absorbance versus radius for data collected at 300 (left), 110 (middle) and 66 (right) µM and 15,000 rpm (green), 25,000 rpm (red) and 40,000 rpm (blue). Data shown were collected at 280 nm in 3 mm (left) or 12 mm (middle and right) pathlength cells. Data collected at 250 nm and concentrations of 32 and 16 µM are not shown. Data were analyzed globally in terms of reversible monomer-dimer-tetramer and monomer-dimer self-association. Best-fits to the reversible monomer-dimer-tetramer model are shown as black lines. The corresponding distributions of the residuals for the monomer-dimer and monomer-dimer-tetramer models are shown in the plots above, from which it is clear that there are large systematic deviations in the residuals for the monomer-dimer model. For the monomer-dimer-tetramer model, excellent fits were obtained with r.m.s.d. values ranging from 0.0022 to 0.0042 absorbance units, and values for the monomer-dimer and dimer-tetramer equilibrium dissociation are 0.58 ± 0.13 mM and 9 ± 3 µM, respectively. Data were recorded at 20°C in 50 mM sodium phosphate (pH 6.5) and 100 mM NaCl. (B) Simulation of the species populations (in monomer units) as a function of concentration calculated from the equilibrium dissociation constants obtained from sedimentation equilibrium AUC.

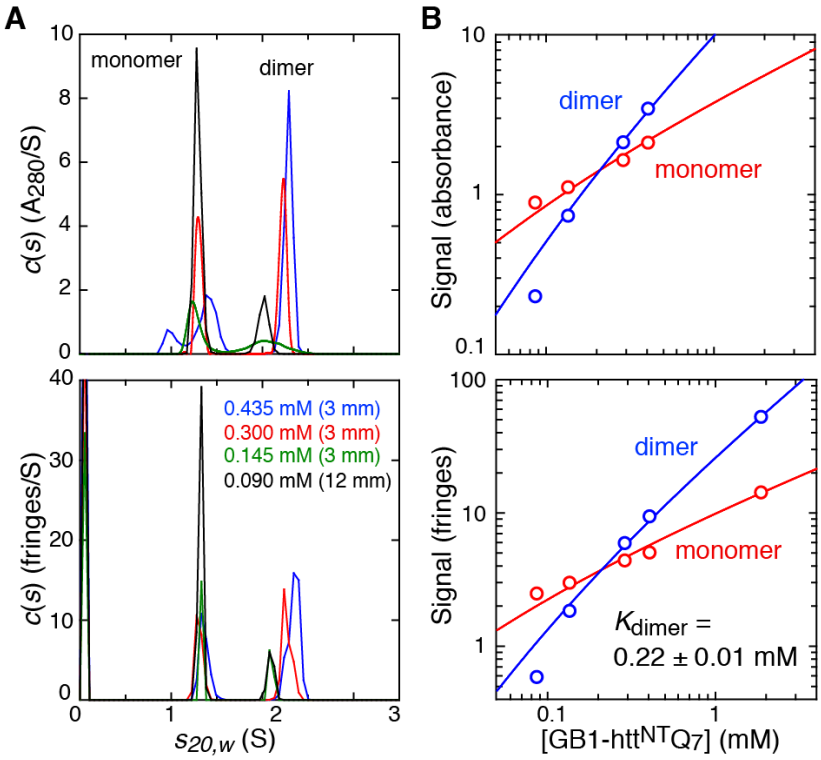


Figure S15. Sedimentation velocity AUC characterization of GB1-htt$^{NT}Q_7$ (S15C-R1) is indicative of monomer-dimer self-association. (A) Sedimentation velocity absorbance (280 nm) and interference c(s) profiles are shown in the top and bottom panels, respectively. As the concentration is increased the proportion of monomer at 1.31 S decreases, and the proportion of dimer at 2.24 S increases. (Note at the lowest concentration, a 12 mm pathlength cell was used; for the other concentrations a 3 mm cell was employed. Interference data for the highest concentration of 1.87 mM is not shown). Peak integration provides the signal for each species. (B) Global fitting of the signal for each species (open circles), obtained by peak integration of the absorbance and interference(s) profiles, to a monomer-dimer model yields binding isotherms (continuous lines) with K_{dimer} = 0.22 ± 0.01 mM, consistent with the value obtained by monitoring the modulation depth as a function of concentration in DEER experiments (Fig. 8C, main text). Data were recorded at 20 °C in 50 mM sodium phosphate (pH 6.5) and 100 mM NaCl.

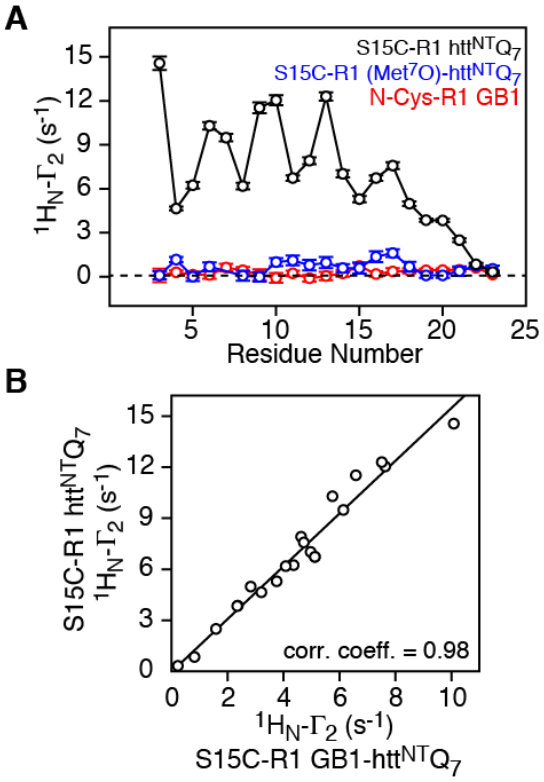


Figure S16. Control intermolecular PRE experiments. (A) Experimental PRE profiles (^1H_N-Γ_2) measured on 150 μM ^{15}N-labeled htt^{NT}Q$_7$ in the presence of 150 μM nitroxide spin-labeled htt^{NT}Q$_7$ (S15C-R1) (black), (Met^{7}O)-htt^{NT}Q$_7$ (S15C-R1) (blue) and N-Cys-R1 GB1 (red). Error bars: 1 SD. (B) Correlation between experimental ^1H_N-Γ_2 values measured on ^{15}N-labeled htt^{NT}Q$_7$ in the presence of spin-labeled htt^{NT}Q$_7$ (S15C-R1) [*y* axis; 1:1 ratio of ^{15}N-labeled to S15C-R1-labeled htt^{NT}Q$_7$, with a total peptide concentration of 300 μM) or GB1-htt^{NT}Q$_7$ (S15C-R1) [*x* axis; 40:1 ratio of ^{15}N-labeled htt^{NT}Q$_7$ to S15C-R1 labeled GB1-htt^{NT}Q$_7$ with a total peptide concentration of 600 μM]. Data were recorded at 10°C at a spectrometer frequency of 600 MHz.

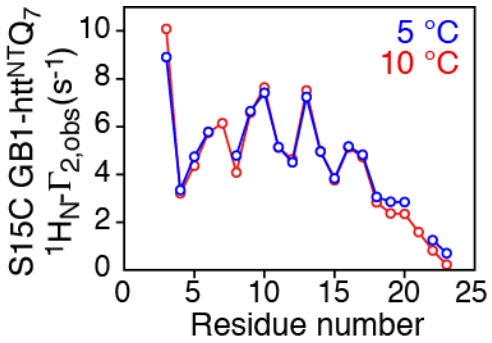


Figure S17. Comparison of experimental ^1H_N-Γ_2 profiles recorded at 5 °C (blue) and 10 °C (red) observed on ^{15}N-labeled htt^{NT}Q$_7$ in the presence of nitroxide spin-labeled GB1-htt^{NT}Q$_7$ (S15C-R1) at a 40:1 molar ratio of ^{15}N-labeled to nitroxide-labeled peptide (total peptide concentration: 600 μM).

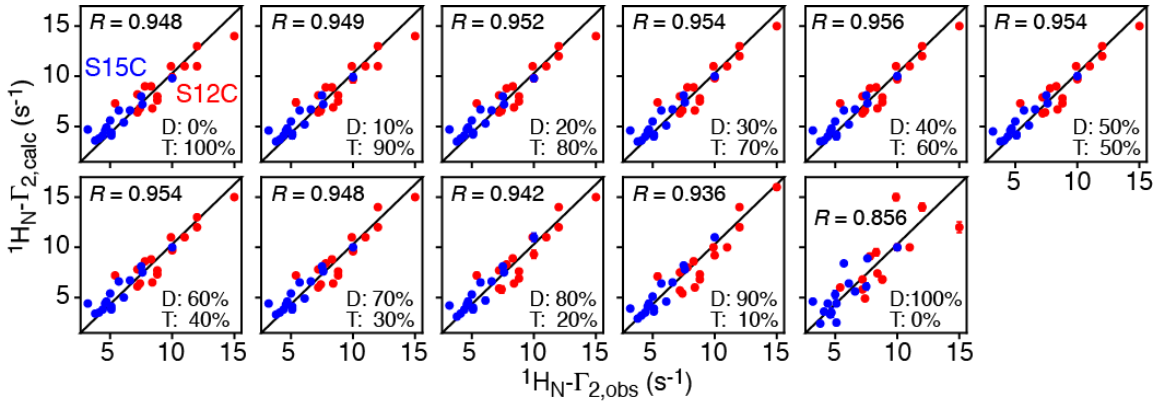


Figure S18. Correlation between observed and calculated PREs for structures calculated with different proportions of dimer and tetramer ranging from 100% tetramer to 100% dimer. The experimental ^1H_N-Γ_2 values (*x*-axis) were measured on ^{15}N-labeled htt$^{NT}Q_7$ in the presence of nitroxide spin-labeled GB1-htt$^{NT}Q_7$ (S15C-R1) at a 40:1 molar ratio of ^{15}N-labeled to nitroxide-labeled peptide; the total peptide concentration was 600 μM and the data were recorded at 10 °C. The mean ^1H_N-Γ_2 calculated values (*y*-axis) are taken from the 10 lowest energy simulated annealing structures calculated at different ratios of dimer to tetramer.

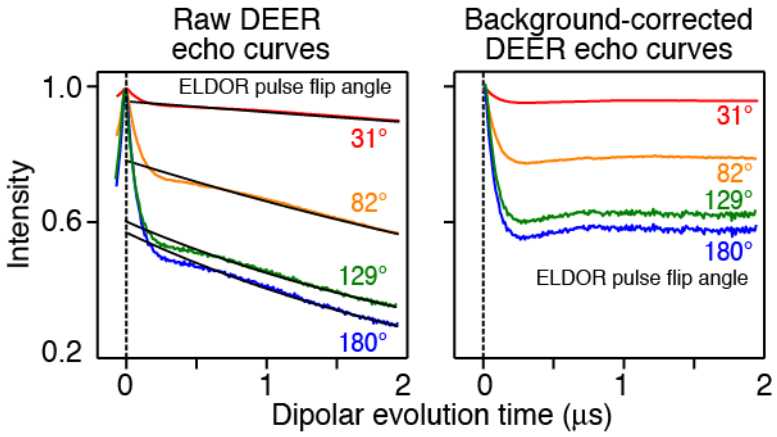


Figure S19. Examples of raw (left) and background-corrected (right) Q-band DEER echo curves obtained on GB1-htt$^{NT}Q_7$ (S15C-R1) with different ELDOR pulse flip angles in the IM-DEER experiment. The best-fit exponential backgrounds are displayed in the left panel as black lines. Calculations were carried out using DeerAnalysis 2016 (13).

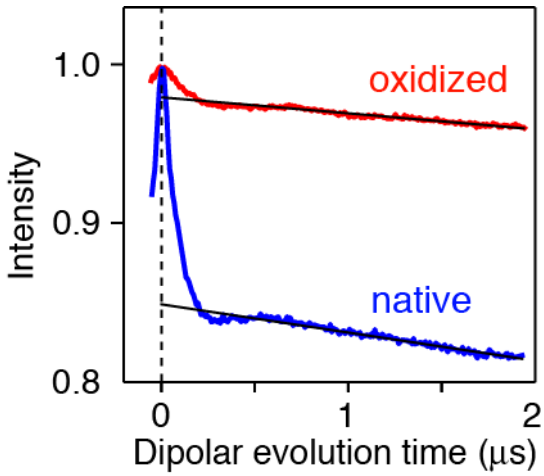


Figure S20. Q-band DEER echo curves recorded on 100 μM GB1-htt^{NT}Q$_7$ (S15C-R1) in native (reduced, blue) and oxidized (Met^7-sulfoxide, red) states at 50 K. The black lines represent the best-fit exponential backgrounds obtained using DeerAnalysis 2016 (13). The modulation depth, given by the difference in intensity between the DEER echo curve and the background at zero dipolar evolution time, is hugely reduced in the Met^{7}O state (obtained by mild treatment with H_2O_2), indicating that the amount of dimer present in the Met^{7}O oxidized sample is minimal (and any residual dimer present is due to incomplete Met^7 oxidation).

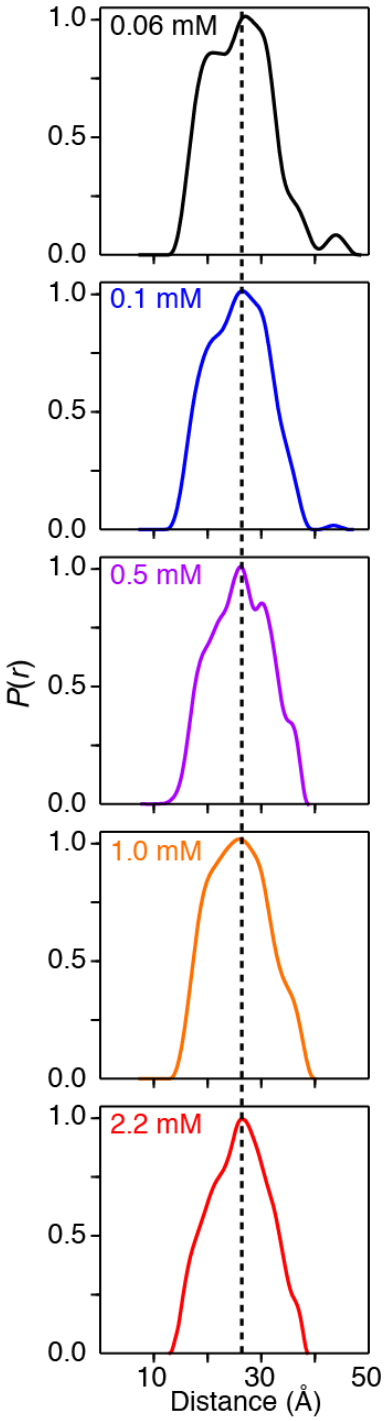


Figure S21. DEER-derived $P(r)$ distance distributions between nitroxide labels in the GB1-htt^{NT}Q$_7$ (S15C-R1) dimer as a function of peptide concentration from 0.06 to 2.2 mM. The dashed line indicates the mean distance which remains invariant. The $P(r)$ distributions were obtained from Q-band DEER measurements at 50 K by validated Tikhonov regularization using DeerAnalysis 2016 (13). The DEER echo curve for the 2.2 mM sample is shown in the main text (Fig. 6).

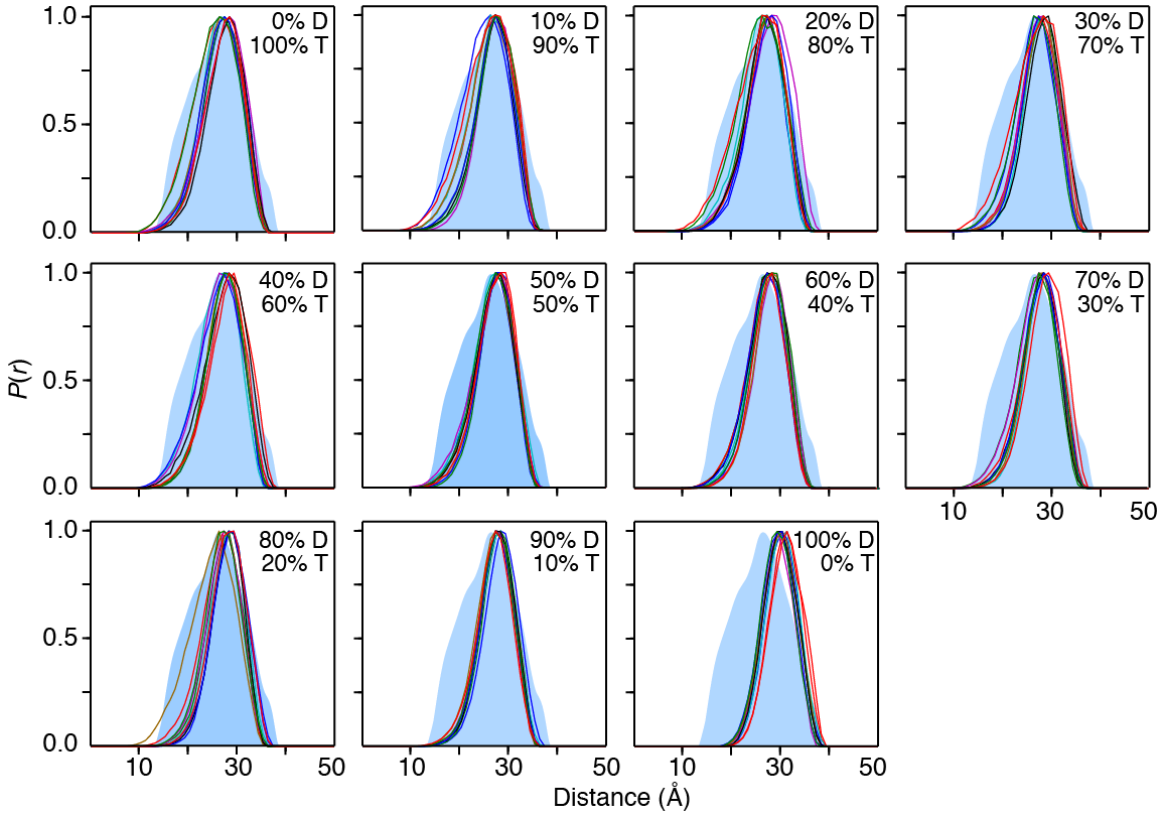


Figure S22. Comparison of the experimental DEER-derived $P(r)$ distribution between spin labels in the GB1-htt$^{NT}Q_7$ (S15C-R1) dimer (shaded light blue area) with those between S15C-R1 spin labels in the dimer unit from the PRE-based simulated annealing structures calculated with varying proportions of dimer (D) to tetramer (T) (continuous lines). (Note the tetramer is a dimer of dimers, see Figure 8A of main text). The Q-band DEER echo curve is shown in the main text (Figure 10, main text) and the experimental $P(r)$ distribution (light blue shaded region) is obtained by validated Tikhonov regularization using DeerAnalysis 2016 (13). The calculated $P(r)$ distributions within the dimer unit of the simulated annealing structures were obtained using the program MMM2013.2 (30) which generates a probability weighted ensemble of all possible allowable conformations of the Cys-R1 side chain using a conformational database potential of mean force. The calculated $P(r)$ distributions for the lowest 10 energy structures for each dimer to tetramer ratio are displayed.