PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY CHEMISTRY

Quantitative NMR analysis of the kinetics of prenucleation oligomerization and aggregation of pathogenic huntingtin exon-1 protein

Alberto Ceccon[a], Vitali Tugarinov[a], Francesco Torricella[a,1], and G. Marius Clore[a,2]

Contributed by G. Marius Clore; received May 3, 2022; accepted June 3, 2022; reviewed by Hashim Al-Hashimi and David Weber

The N-terminal region of the huntingtin protein, encoded by exon-1 (htt^{ex1}) and containing an expanded polyglutamine tract, forms fibrils that accumulate in neuronal inclusion bodies, resulting in Huntington's disease. We previously showed that reversible formation of a sparsely populated tetramer of the N-terminal amphiphilic domain, comprising a dimer of dimers in a four-helix bundle configuration, occurs on the microsecond timescale and is an essential prerequisite for subsequent nucleation and fibril formation that takes place orders of magnitude slower on a timescale of hours. For pathogenic htt^{ex1}, such as htt^{ex1}Q$_{35}$ with 35 glutamines, NMR signals decay too rapidly to permit measurement of time-intensive exchange-based experiments. Here, we show that quantitative analysis of both the kinetics and mechanism of prenucleation tetramerization and aggregation can be obtained simultaneously from a series of ^{1}H-^{15}N band–selective optimized flip-angle short-transient heteronuclear multiple quantum coherence (SOFAST-HMQC) correlation spectra. The equilibria and kinetics of tetramerization are derived from the time dependence of the ^{15}N chemical shifts and ^{1}H-^{15}N cross-peak volume/intensity ratios, while the kinetics of irreversible fibril formation are afforded by the decay curves of ^{1}H-^{15}N cross-peak intensities and volumes. Analysis of data on htt^{ex1}Q$_{35}$ over a series of concentrations ranging from 200 to 750 μM and containing variable (7 to 20%) amounts of the Met^{7}O sulfoxide species, which does not tetramerize, shows that aggregation of native htt^{ex1}Q$_{35}$ proceeds via fourth-order primary nucleation, consistent with the critical role of prenucleation tetramerization, coupled with first-order secondary nucleation. The Met^{7}O sulfoxide species does not nucleate but is still incorporated into fibrils by elongation.

huntingtin exon-1 | short-lived excited states | aggregation kinetics | nucleation | NMR spectroscopy

CAG expansion within exon 1 of the huntingtin gene (*HTT*) is responsible for Huntington's disease, a fatal, autosomal dominant neurodegenerative condition (1–3). *HTT* exon-1 encodes a polypeptide, htt^{ex1}Q$_n$, comprising a 16-residue N-terminal amphiphilic domain (NT), a polyglutamine (polyQ) repeat of variable length n, and a proline rich domain (PRD) comprising two polyproline repeats separated by 17 residues (3). Extension of the polyQ repeat beyond about 35 residues results in clinical disease with an age of onset that is approximately correlated with the length of the polyQ repeat (4–6). Proteolysis of the huntingtin protein (7) or incomplete messenger RNA splicing of *HTT* (8) results in N-terminal fragments with long polyQ repeats that aggregate and fibrillize to form neuronal inclusion bodies (9, 10). Solid-state NMR studies have shown that htt^{ex1} fibrils comprise a central rigid β-hairpin/β-sheet polyQ core that stacks from one layer of the fibril to the next via interdigitated hydrogen bonds between the glutamine side chains (11, 12), with NT helices of intermediate dynamics and highly mobile PRD domains on the outside of the fibril (13–19). This structural picture of htt^{ex1} fibrils is also supported by cryo-electron tomography (20, 21). In a series of quantitative exchange-based NMR studies on constructs of htt^{ex1}, we showed that htt^{ex1}Q$_7$ (22, 23) as well as a shorter construct, htt^{NT}Q$_7$, lacking the PRD domain (24), undergo transient tetramerization of the NT region on a timescale of 50 to 70 μs to form a sparsely populated four-helix bundle with D_2 symmetry comprising a dimer of dimers (24). The dimer consists of an anti-parallel helical coiled-coil with C_2 symmetry formed by residues 3 to 17 of the NT region (24). The NT tetramer increases the local concentration of the polyQ tracts and provides a template that promotes nucleation of the polyQ region. Although aggregation/fibril formation occurs on a timescale of minutes to hours, the importance of submillisecond prenucleation tetramerization for subsequent fibril formation is evidenced by the fact that reduction of the population of tetramers by a number of mechanisms effectively blocks fibril formation (22, 24–27), thereby providing a link between two processes whose timescales are

Significance

Huntington's disease is a fatal neurodegenerative condition caused by polyglutamine expansion (≥35) in the N-terminal region of the huntingtin protein encoded by exon-1 (htt^{ex1}), resulting in fibril accumulation within neuronal inclusion bodies. Microsecond-timescale reversible oligomerization to generate sparsely populated tetramers of the N-terminal amphiphilic domain of htt^{ex1} is critical for nucleation and elongation on the hours timescale to form fibrils comprising a polyglutamine core. Here, we develop an NMR approach to simultaneously quantify the kinetics of transient tetramerization and irreversible fibril formation of a pathogenic htt^{ex1} construct with 35 glutamines. Native htt^{ex1} undergoes tetramer-dependent primary nucleation and monomer-dependent elongation and secondary nucleation, whereas the Met^{7} sulfoxide form, which does not tetramerize or nucleate, is still incorporated into fibrils via elongation.

Reviewers: H.A.-H., Duke University Hospital; and D.W., University of Maryland.

The authors declare no competing interest.

[1]Permanent address: Magnetic Resonance Center, University of Florence, Sesto Fiorentino, Florence 50019, Italy.

[2]To whom correspondence may be addressed. Email: mariusc@mail.nih.gov.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2207690119/-/DCSupplemental.

Published July 12, 2022.

separated by many orders of magnitude. Inhibition of fibril formation can be achieved by three mechanisms: sequestration of the htt^{ex1} monomer by binding of chaperonins to the NT region (26), thereby decreasing the concentration of free monomer below the critical concentration required for nucleation [since the population of tetramers is proportional to the third power of the free monomer concentration (24)]; modification of NT residues at the dimer and/or tetramer interface, such as oxidation of Met^7 to a sulfoxide, that inhibits productive dimerization or tetramerization (24, 25); and lastly, long-range allosteric inhibition of the monomer ↔ dimer and/or dimer ↔ tetramer transitions by binding of cellular factors, such as profilin and the Fyn SH3 domain, respectively, to the PRD domain (22, 27). These in vitro data are supported by in vivo observations that deletion of the NT region or introduction of mutations within the NT region that are predicted to decrease helical propensity, reduce the formation of cellular inclusion bodies, and decrease neuronal toxicity (28).

In our previous work on $htt^{ex1}Q_n$ tetramerization, we made use of a construct comprising only seven glutamine repeats that remained monomeric over a prolonged period of time, only forming fibrils over a period of weeks, thereby permitting detailed analysis of the exchange dynamics of tetramerization by time-intensive NMR experiments, including concentration dependence of chemical shifts, transverse relaxation in the rotating frame ($R_{1\rho}$) at multiple spin-lock fields, and relaxation dispersion measurements (22–24). For $htt^{ex1}Q_{35}$, however, such NMR experiments are not feasible since NMR signals arising from the monomer decay rapidly as a result of aggregation over a period of hours, even at a temperature of 5 °C (27). Here, we show that quantitative analysis of both the kinetics and mechanism of aggregation/fibril formation and tetramerization occurring on the timescale of hours and microseconds, respectively, can be obtained simultaneously from a series of ^1H-^{15}N band-selective optimized flip-angle short-transient heteronuclear multiple quantum coherence (SOFAST-HMQC) correlation spectra (29, 30), recorded as a function of time. The kinetics of fibril formation is afforded by quantitative analysis of the decay curves of ^1H-^{15}N cross-peak intensities and volumes, while the kinetics and equilibria involving the monomer ↔ dimer ↔ tetramer transitions are derived from the time dependence of the ^{15}N chemical shifts and cross-peak volume/intensity (*V/I*) ratios (Fig. 1). We show that tetramerization of native $htt^{ex1}Q_{35}$ occurs on the same microsecond timescale as the shorter $htt^{ex1}Q_7$ construct, but the stability of the $htt^{ex1}Q_{35}$ tetramer is enhanced fourfold relative to that of $htt^{ex1}Q_7$ (23), thereby reducing the critical concentration of $htt^{ex1}Q_{35}$ monomer required for nucleation. We further show, using the formalism developed in ref. 31, that fibrillization of native $htt^{ex1}Q_{35}$ involves a fourth-order primary nucleation process coupled to monomer-dependent first-order secondary nucleation. The Met^7O sulfoxide form of $htt^{ex1}Q_{35}$, on the other hand, does not tetramerize and does not undergo either primary or secondary nucleation, but it can still be incorporated into fibrils formed via nucleation of native $htt^{ex1}Q_{35}$.

Results and Discussion

Experimental Design. Even at 5 °C, aggregation of $htt^{ex1}Q_{35}$ occurs in a matter of hours at the concentrations appropriate for NMR experiments (~0.1 to 1 mM). As a result, it is not feasible to record triple-resonance backbone correlation experiments for assignment of backbone resonances; however, assignment of the $htt^{ex1}Q_{35}$ ^1H-^{15}N correlation spectrum (*SI Appendix*, Fig. S1*A* and Table S1) could easily be made with reference to the previously published $^1H_N/^{15}N$ assignments of the shorter $htt^{ex1}Q_7$ construct, which remains monomeric for many days (22), as there are only minor chemical shift differences between the two constructs for residues in the NT and PRD domains. $^1H_N/^{15}N$ cross-peaks that belong to the glutamines of the polyQ domain display extensive chemical shift overlap, with the exception of Gln^{50} and Gln^{51}, which are immediately adjacent to the N-terminal end of the first polyproline tract of the PRD domain.

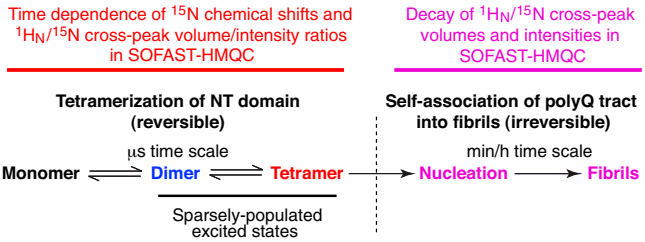

Fig. 1. Simultaneous determination of the kinetics and mechanism of tetramerization and fibril formation of $htt^{ex1}Q_{35}$ from serially recorded ^1H-^{15}N SOFAST-HMQC spectra. The kinetics and equilibria of tetramerization are derived from the time dependence of the ^{15}N chemical shifts and $^1H_N/^{15}N$ cross-peak *V/I* ratios, which provide the dependence of these spectral parameters on the concentration of free monomer, while the kinetics and mechanism of fibril formation are determined from analysis of the decay of $^1H_N/^{15}N$ cross-peak intensities and volumes. Inhibition of tetramer formation, either directly by modification of residues within the NT region (24, 25) or sequestration of the monomer by binding of the NT region to chaperones (26) or indirectly via an allosteric mechanism by binding of cellular proteins such as profilins and SH3 domains to the C-terminal polyproline rich domain (22, 27), blocks fibril formation.

Likewise, the relaxation-based NMR experiments previously used for $htt^{ex1}Q_7$ to probe submillisecond processes at equilibrium (22), including relaxation dispersion and concentration dependence of transverse relaxation in the rotating frame (23), can also not be applied.

Our strategy to quantitatively analyze the kinetics of tetramerization and fibrillization of $htt^{ex1}Q_{35}$ simultaneously therefore entailed following the intensity, volume, and chemical shifts of cross-peaks arising from NMR-visible, monomeric $htt^{ex1}Q_{35}$ as a function of time using serially acquired SOFAST-HMQC correlation spectra (29, 30), where each spectrum could be readily recorded with reasonable signal-to-noise over a period of 6 min (*SI Appendix*, Fig. S1*B*). The SOFAST-HMQC experiment permits high repetition rates with short interscan delays by combining a small number of radiofrequency pulses, Ernst-angle excitation (32), and longitudinal relaxation optimization (33, 34). Cross-peaks from immobilized residues within $htt^{ex1}Q_{35}$ fibrils are not visible, however, as they are too broad and are therefore buried within the baseline (owing to their very large transverse relaxation rates on account of their high molecular weight and hence slow rotational tumbling).

Qualitative Picture of Aggregation and Tetramerization Monitored by NMR. Fig. 2*A* shows examples of the time dependence of cross-peak intensities and/or volumes from residues in the NT, the last residue of the polyQ repeat, and four residues within the PRD, for a sample comprising a total initial concentration of 750 μM $htt^{ex1}Q_{35}$ at 5 °C. Aggregation was complete in about 10 h, at which point no cross-peaks were visible for residues in the NT and polyQ regions (Fig. 2*B* and *SI Appendix*, Fig. S1*B*). Residual cross-peak intensity, however, was apparent for residues within the PRD, with increasing intensity toward the C terminus. Thus, one can conclude that while the NT and polyQ regions were fully immobilized in the

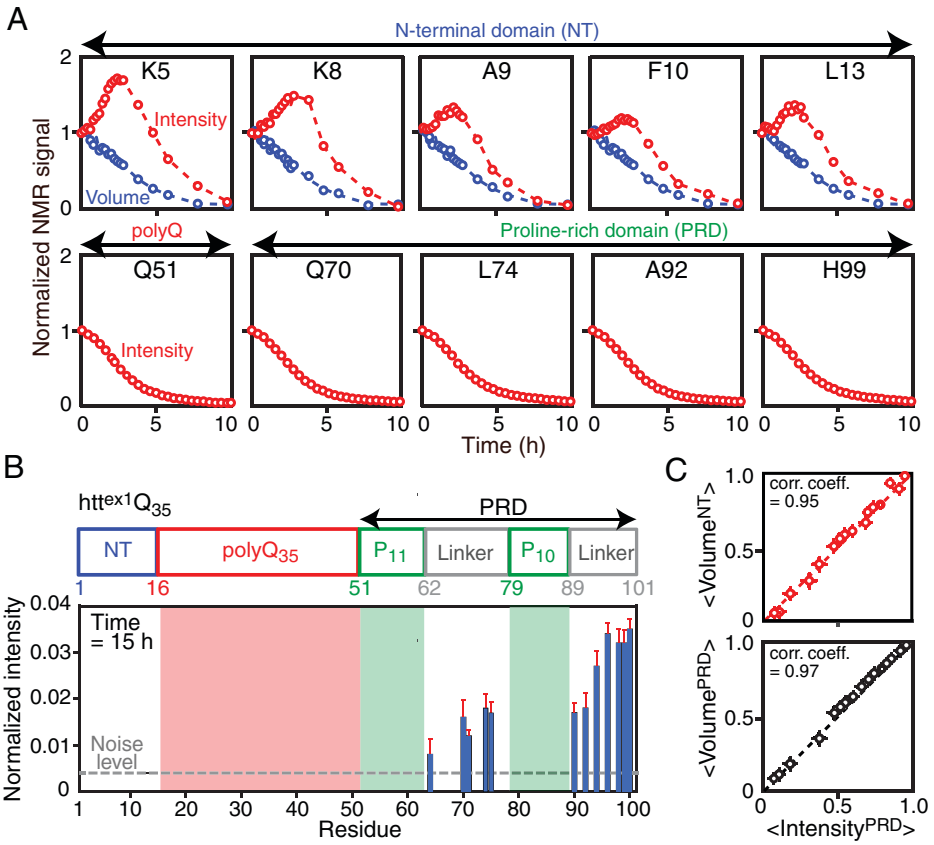

Fig. 2. Overall aggregation profiles of 750 μM ^{15}N-labeled htt^{ex1}Q$_{35}$ monitored by acquisition of serial ^{1}H-^{15}N SOFAST-HMQC spectra at 5 °C and 800 MHz. (*A*) Time courses of normalized cross-peak intensities (red circles) and volumes (blue circles) for residues in the NT, polyQ, and PRD domains. The data are normalized to the first (reference) experiment recorded immediately upon dissolution and pH adjustment (t ~0 h). (*B*) Domain organization of htt^{ex1}Q$_{35}$ (*Top*) and residual ^{1}H-^{15}N cross-peak intensities after 15 h (*Bottom*). No cross-peaks are detectable above the noise level for residues in either the NT or polyQ regions, indicating that these regions are fully immobilized within the htt^{ex1}Q$_{35}$ fibrils. ^{1}H-^{15}N cross-peaks attributable to residues in the PRD, however, are present with the intensity increasing toward the C terminus, indicative of significant mobility/disorder of the PRD within the htt^{ex1}Q$_{35}$ fibrils. The error bars represent the SDs for the cross-peak intensities determined from three ^{1}H-^{15}N SOFAST-HMQC spectra recorded at 14.8, 15.0, and 15.2 h. Note that no cross-peaks can be detected for the two proline repeats (shaded in green) in a ^{1}H-^{15}N correlation spectrum since the proline nitrogen atom is not bonded to a proton. (*C*) Correlation between the normalized average ^{1}H-^{15}N cross-peak intensities for residues within the PRD (x axis) and the average normalized ^{1}H-^{15}N cross-peak volumes (y axis) for residues in the NT (*Top*) and PRD (*Bottom*) domains. corr. coeff., correlation coefficient.

fibrils/aggregates, residue C-terminal to the first polyproline exhibited significant mobility and hence disorder in the fibrils, consistent with previous solid-state NMR studies (16, 18).

In the absence of a concentration-dependent exchange process resulting in exchange line broadening, cross-peak intensity (height) will be linearly correlated to cross-peak volume, which is directly proportional to the concentration of monomeric, NMR-visible htt^{ex1}Q$_{35}$. The average intensity of cross-peaks within the PRD was linearly correlated to the average volume of cross-peaks within both the NT (correlation coefficient, $r = 0.95$) and PRD ($r = 0.97$) regions (Fig. 2*C*); likewise, the average intensity of the overlapping cross-peaks comprising the polyQ region was linearly correlated to the average volume of the NT cross-peaks ($r = 0.98$) and average intensity of the PRD cross-peaks ($r = 0.97$) (*SI Appendix*, Fig. S2). One can therefore conclude that fibril formation occurs cooperatively and that all residues of htt^{ex1}Q$_{35}$ follow the same aggregation kinetics.

For residues in the NT region, it is clear that cross-peak intensity and volume do not track initially; specifically, while cross-peak volume decreases from $t = 0$, cross-peak intensities increase, reaching a maximum at 2 to 3 h, depending on the residue, before decreasing (Fig. 2*A*). This phenomenon can be attributed to chemical exchange line broadening arising from submillisecond tetramerization accompanied by significant ^{15}N and ^{1}H$_N$ chemical shift differences between monomeric and transient dimeric/tetrameric species (22, 24). The fractional population of tetramers (E$_4$) is proportional to the third power of the monomer concentration. As the concentration of monomer decreases initially, exchange line broadening is reduced and the concomitant increase in cross-peak intensity due to line narrowing is larger than the decrease in intensity resulting from the reduction in monomer concentration.

Modification of Met^7 to a Sulfoxide Blocks Tetramerization and Nucleation But Stills Allows Incorporation into Fibrils. During expression and purification of htt^{ex1}Q$_{35}$, there is invariably a small fraction in which Met^7 is oxidized to a sulfoxide (Met^{7}O). Out of the 16 residues comprising the NT domain of htt^{ex1}Q$_{35}$, the only visible cross-peaks arising from the Met^{7}O sulfoxide form belonged to Ser^{12} and Ser^{15} (Fig. 3*A*). Ser^{12} had a distinct and very well–isolated cross-peak for the Met^{7}O species, suitable for quantitative kinetic analysis, while the ^{1}H$_N$/^{15}N cross-peaks for Ser^{15} in the reduced and oxidized species were partially overlapped (Fig. 3*A*). In the batch used to make the 750-μM sample, 20 (±2)% was present in the Met^{7}O sulfoxide form (Fig. 3*A*). In all the other samples (200 to 500 μM) used in the current work, obtained from a separate preparation, the fraction Met^{7}O form was much reduced and comprised only 7 to 8%. Fully Met^{7}O-oxidized htt^{ex1}Q$_{35}$ did not

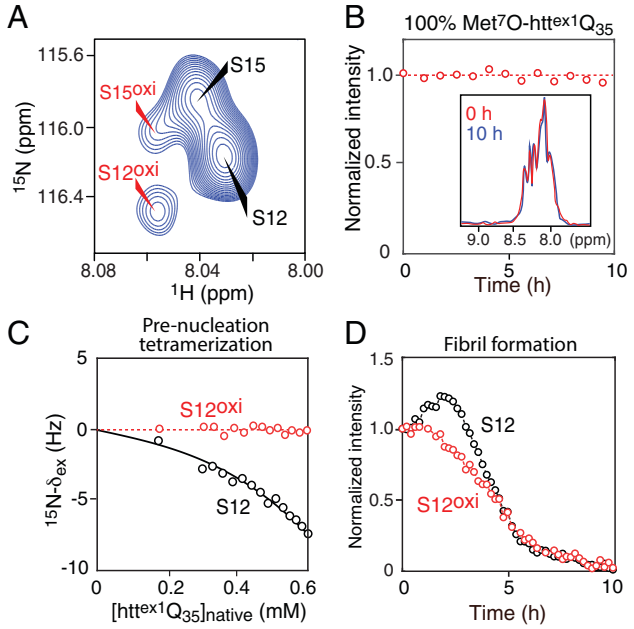

Fig. 3. Impact of the oxidation of Met^{7} to a sulfoxide (Met^{7}O) on prenucleation tetramerization and fibril formation. (*A*) Ser^{12}/Ser^{15} region of the ^{1}H-^{15}N SOFAST-HMQC spectrum of 750 μM ^{15}N-labeled htt^{ex1}Q$_{35}$. Cross-peaks arising from native (reduced) and Met^{7}O oxidized forms are labeled in black and red, respectively. In this sample, the fraction of the Met^{7}O sulfoxide form is 20 ± 2%. (*B*) 100% Met^{7}O oxidized htt^{ex1}Q$_{35}$ (150 μM) does not aggregate as monitored from the time course of the integrated intensity (from 7.9 to 8.5 ppm) of the amide proton envelope (obtained from the first increment of a ^{1}H-^{15}N SOFAST-HMQC spectrum. The inset shows a one-dimensional spectrum of the amide proton envelope recorded at 0 (red) and 10 (blue) h. (*C*) Concentration dependence of ^{15}N exchange-induced shifts (^{15}N-δ_{ex}) for Ser12 in the native reduced (black) and Met^{7}O oxidized (red) forms of htt^{ex1}Q$_{35}$ referenced to the shifts of a 10-μM sample. The total concentration of htt^{ex1}Q$_{35}$ at t = 0 h was 750 μM, corresponding to 600-μM native and 150-μM Met^{7}O oxidized forms, and the shifts are followed by serial ^{1}H-^{15}N SOFAST-HMQC spectra as the concentration of monomeric htt^{ex1}Q$_{35}$ decreases. The concentration is determined from the time dependence of the average cross-peak intensities for residues within the PRD domain. (*D*) Time dependence of the ^{1}H-^{15}N cross-peak intensities for Ser^{12} in the native reduced (black) and Met^{7}O oxidized (red) forms observed for a 750-μM sample of ^{15}N-labeled htt^{ex1}Q$_{35}$ with 20% in the Met^{7}O oxidized form. All NMR data were recorded at 5 °C and 800 MHz.

aggregate (Fig. 3*B*), consistent with our previous work on the two shorter constructs, htt^{NT}Q$_{7}$ and htt^{NT}Q$_{10}$, lacking the PRD domain (35). Further, cross-peaks within the NT region attributable to the Met^{7}O form did not exhibit exchange-induced shifts as a function of concentration, in contrast to the corresponding cross-peaks of the native reduced form (Fig. 3*C*), indicating that the Met^{7}O form did not undergo transient homo- or hetero-tetramerization, even in the presence of the native reduced form. However, the cross-peaks attributable to the Met^{7}O form decreased with time as the sample fibrillized (Fig. 3*D*), indicating that while the Met^{7}O form did not undergo primary nucleation, it was incorporated into fibrils formed via nucleation of the native reduced form. Further, the half-life ($t_{1/2}$) for incorporation of the Met^{7}O form into the fibrils was ~4 h, about ~30% slower than for the native reduced form ($t_{1/2}$ ~3 h) (Fig. 3*D*).

Transient Tetramerization of the Native Reduced Form of htt^{ex1}Q$_{35}$. The concentration dependence of the ^{15}N exchange-induced shifts (Fig. 4*A*) and *VII* ratios (Fig. 4*B*) for cross-peaks of residues within the NT region is shown in Fig. 3 *A* and *B*, respectively. The concentration of monomeric native (reduced) htt^{ex1}Q$_{35}$ as a function of time, $m_{red}(t)$, was readily obtained by directly recasting the average normalized cross-peak intensity ($<I_{PRD}>$) of the PRD residues (Fig. 4*A*) as $m_{red}(t) = fm_{tot}[< I_{PRD}(t) > - < I_{PRD}(\infty) >]/[< I_{PRD}(0) > - < I_{PRD}(\infty) >]$, where m_{tot} is the total monomer concentration at $t = 0$ and f is the fraction of sample in the native reduced form (f = 0.8 for the 750-μM sample). These data display the typical curvature expected for submillisecond tetramerization previously observed and quantitatively described for htt^{ex1}Q$_{7}$ and htt^{NT}Q$_{7}$ (22–24).

We previously showed that prenucleation oligomerization involves a branched pathway (22, 24): an on-pathway leading to a tetramer (dimer of dimers) via a productive coiled-coil helical dimer of the NT region associated with large secondary shifts within the NT region, and an off-pathway branch resulting in the formation of a nonproductive, partially helical NT dimer (with much smaller secondary shifts) that does not undergo further self-association and probably comprises an ensemble of states in the incorrect register. The contribution of off-pathway dimerization to both exchange-induced shifts and exchange-line broadening is small compared to that of on-pathway dimerization and tetramerization (23) and, for the purposes of the present analysis, can be ignored (see *Experimental Procedures*). While the exchange-induced shifts report solely on the on-pathway monomer ↔ dimer ↔ tetramer equilibria, the *VII* ratios also contain information on the rates of these processes.

The calculation of *VII* ratios in our analysis, apart from accounting for the reduction of NMR signal from monomeric htt^{ex1}Q$_{35}$ as a result of aggregation, also eliminated the effects of homonuclear ^{1}H-^{1}H couplings (predominantly three-bond $^{3}J_{HN\alpha}$ couplings) in the acquisition period of the ^{1}H-^{15}N SOFAST-HMQC experiment, as both volumes and intensities (heights) are expected to be affected by these couplings in the same manner. Under conditions where $e^{-R^{eff}_{2,H}t_2^{max}} \ll 1$ and $R^{eff}_{2,DQ/ZQ}t_1^{max}/2 \ll 1$, and hence $\left(1 - e^{-R^{eff}_{2,DQ/ZQ}t_1^{max}/2}\right) \approx R^{eff}_{2,DQ/ZQ}t_1^{max}/2$, the *VII* ratio in a ^{1}H-^{15}N SOFAST-HMQC experiment, processed with no apodization of the time-domain data, can be expressed as

$$\frac{V}{I} \approx \alpha' \left[\frac{R^{eff}_{2,H}}{2 + e^{-R^{eff}_{2,DQ}t_1^{max}} + e^{-R^{eff}_{2,ZQ}t_1^{max}}}\right], \qquad [1]$$

where α′ is a fitting parameter, t_1^{max} and t_2^{max} are the acquisition times in the indirect (^{15}N) and direct (^{1}H) dimensions, respectively; the effective transverse relaxation rates $R^{eff}_{2,DQ}$, $R^{eff}_{2,ZQ}$, and $R^{eff}_{2,H}$ are equal to ($R_{2,DQ}$ + $R_{ex,DQ}$), ($R_{2,ZQ}$ + $R_{ex,ZQ}$), and ($R_{2,H}$ + $R_{ex,H}$), respectively; $R_{2,DQ}$, $R_{2,ZQ}$, and $R_{2,H}$ are ^{15}N-^{1}H$_{N}$ double-quantum (DQ), zero-quantum (ZQ), and amide proton single-quantum (H) transverse relaxation rates, respectively; and $R_{ex,n}$ values are the contributions to the rate of each type of coherence n ($n \in$ 'DQ', 'ZQ', 'H') from chemical exchange, calculated from the largest (real) eigenvalue of the appropriate NMR Liouvillian given by the Bloch-McConnell equations (36). The derivation of Eq. **1** is provided in *SI Appendix*.

A global fit of the kinetic scheme for tetramerization (Fig. 4*C*) to the ^{15}N exchange-induced shift data and *VII* ^{1}H-^{15}N cross-peak ratios shown in Fig. 3 *A* and *B*, assuming the same ^{15}N chemical shift differences between monomer and dimer/tetramer for htt^{ex1}Q$_{7}$ (*SI Appendix, SI Methods*), yield values of 0.06 ± 0.01 M and 11 ± 5 μM for the equilibrium dissociation constants from dimer to monomer (K_1^{diss}) and tetramer to dimer (K_2^{diss}), respectively, and values of ~3.8 (±1.0) × 10^{4} s^{-1} and 6.2 (±1.8) × 10^{5} M^{-1}s^{-1} for the dissociation and association rate constants, respectively, for the monomer ↔ dimer transition. In the case of the dimer ↔ tetramer

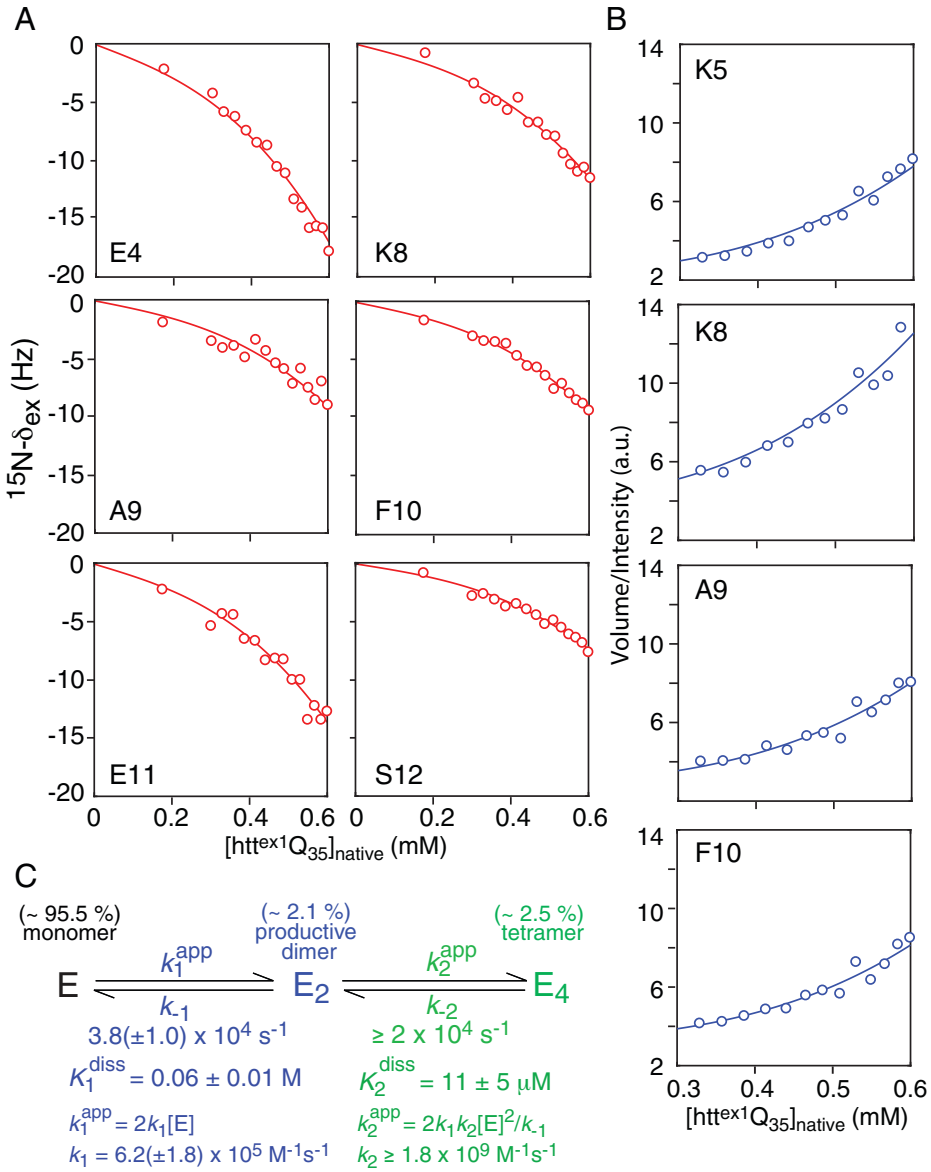

Fig. 4. Quantitative analysis of prenucleation tetramerization of htt^{ex1}Q$_{35}$ from a global fit to concentration-dependent ^{15}N exchange–induced shifts (^{15}N-δ_{ex}) and ^{1}H-^{15}N cross-peak *V/I* ratios. (*A–C*) ^{15}*N*-δ_{ex} (A) and ^{1}H-^{15}N cross-peak (*B*) *V/I* ratios as a function of the concentration of native monomeric htt^{ex1}Q$_{35}$ at 5 °C and 800 MHz. The concentrations are derived from the time dependence of the average cross-peak intensity for residues within the PRD. The total sample concentration is 750 μM, but only the native reduced form, corresponding to 600 μM, undergoes prenucleation oligomerization. The experimental data are shown as circles, and the best fits to the tetramerization scheme in (*C*) are represented by the continuous solid lines. The ^{15}N-δ_{ex} data report only on the monomer ↔ dimer and dimer ↔ tetramer equilibria, while the cross-peak *V/I* ratios also report on the dissociation rate constant from dimer to monomer (k_{-1}) and provide a lower limit on the dissociation rate constant from tetramer to dimer (k_{-2}). (*C*) Kinetic scheme with values of equilibria and rate constants (E, monomer; E$_2$, dimer; E$_4$, tetramer). The populations listed above the species correspond to those at the highest concentration (600 μM at t = 0 h) of native htt^{ex1}Q$_{35}$. For errors of 0.5 Hz and 1 a.u. for ^{15}N-δ_{ex} and ^{1}H-^{15}N *V/I*, respectively, the value of the reduced χ^2 for the global fit is 0.86. a.u., arbitrary units.

transition, only lower limits of k_{-2} (≥2 × 10^4 s^{-1}) and therefore k_2 (≥1.8 × 10^9 M^{-1}s^{-1}) can be established (*SI Appendix*, Fig. S3). The values of the equilibrium and rate constants (K_1^{diss}, k_1 and k_{-1}) for the monomer ↔ dimer transition are the same within error as those reported previously for the shorter htt^{ex1}Q$_7$ construct (22, 23). The same is true for the lower limit for k_{-2}. However, the equilibrium dissociation constant K_2^{diss} for the dimer ↔ tetramer transition is two- to threefold smaller for htt^{ex1}Q$_{35}$ than htt^{ex1}Q$_7$ (22, 23), corresponding to an increase of ~0.4 to 0.6 kcal.mol^{-1} in stability of the htt^{ex1}Q$_{35}$ tetramer relative to the htt^{ex1}Q$_7$ tetramer. At the highest concentration (0.6 mM, close to the solubility limit) of the native reduced form of htt^{ex1}Q$_{35}$ employed here, the populations of dimer and tetramer are ~2% and 2.5%, respectively. Extrapolated to 1.2 mM, the limit of solubility for the shorter htt^{ex1}Q$_7$ construct (22), the population of htt^{ex1}Q$_{35}$ dimer and tetramer would be ~3.5% and 16.5%, respectively, compared to ~3% and 5%, respectively, for htt^{ex1}Q$_7$.

Quantitative Analysis of the Kinetics and Mechanism of htt^{ex1}Q$_{35}$ Aggregation. The time dependence of the disappearance of monomeric htt^{ex1}Q$_{35}$ at initial concentrations of monomer ranging from 200 to 750 μM is shown in Fig. 5*A*. The top panel displays the average cross-peak intensity of residues within the PRD, while the bottom panel displays the cross-peak volume and intensity of Ser^{12} in the native reduced and Met^{7}O oxidized forms, respectively, for the 750-μM sample, which contained 20% of the oxidized form. The quantitative global kinetic analysis of the data in Fig. 5*A* follows the formalism described in refs. 31 and 37, with the differences that we

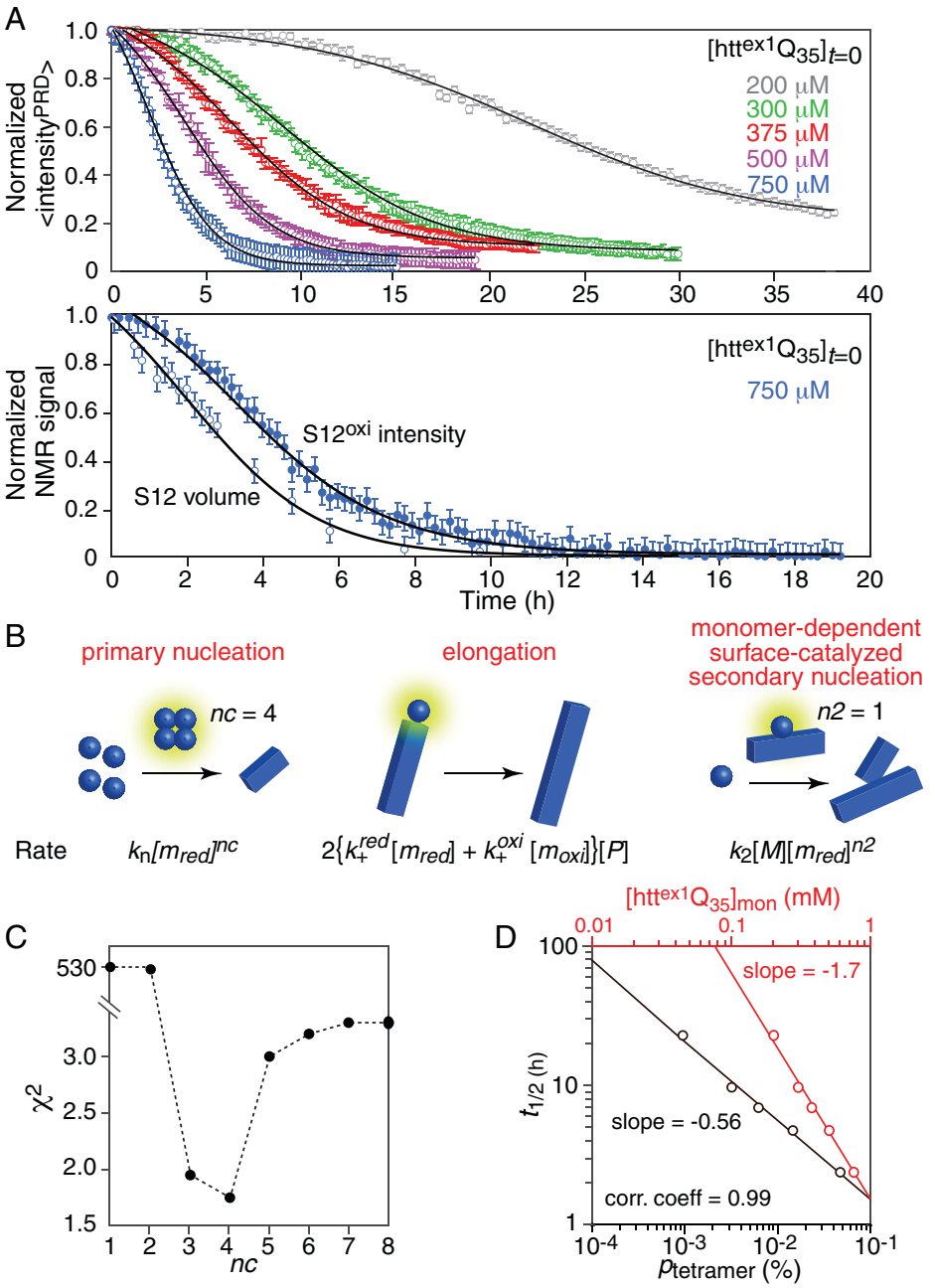

Fig. 5. Quantitative analysis of the mechanism and kinetics of htt^{ex1}Q$_{35}$ aggregation probed by serially acquired ^{1}H-^{15}N SOFAST-HMQC spectra. (*A*) Time dependence of the average ^{1}H-^{15}N cross-peak intensities for residues in the PRD domain (*Top*) and of the ^{1}H-^{15}N cross-peak volume and intensity of Ser^{12} in the native reduced and Met^{7}O oxidized forms, respectively, (*Bottom*) of htt^{ex1}Q$_{35}$. The total concentrations of monomeric htt^{ex1}Q35 at t = 0 are indicated. For the 750-μM sample, the percentage of the Met^{7}O oxidized form is 20 (±2)%; for the other samples, the percentage of the Met^{7}O oxidized form is 7 to 8%. The experimental data (shown as circles, with the bars equal to 1 S.D.) were recorded at 5 °C and 800 MHz and normalized to the first time point (at t ~0 h). The best-fit curves are shown as black continuous lines and were obtained from a global fit to the kinetic scheme described by the set of differential equations in Eq. **4**, which incorporates fourth-order primary nucleation (nc = 4), elongation, and first-order secondary nucleation ($n2$ = 1) for the native reduced form and elongation for the Met^{7}O oxidized form. (*B*) Schematic depiction of primary nucleation, elongation, and monomer-dependent surface-catalyzed secondary nucleation. [m_{red}] and [m_{oxi}] are the concentrations of free monomer in the native reduced and Met^{7}O oxidized forms, respectively; [M] is the total fibril mass in monomer units; [P] is the number concentration of extendable ends of the fibril chain; k_n, k_+, and k_2 are the rate constants for primary nucleation, elongation, and secondary nucleation, respectively; and nc and $n2$ are the order of primary and secondary nucleation, respectively. (*C*) Grid search showing the dependence of the reduced χ^2 on the order of primary nucleation (nc). The minimum χ^2 is obtained for nc = 4, consistent with prenucleation tetramerization. Although the value of χ^2 for nc = 3 is only slightly higher than that for nc = 4, nc = 3 can be excluded as unrealistically high values of the initial fibril number concentration, $P(0)$, about 20-fold higher than for nc = 4, are required to fit the data (*SI Appendix*, Fig. S4). (*D*) Log-log plot of aggregation $t_{1/2}$ versus concentration of htt^{ex1}Q$_{35}$ monomer (red) and population ($p_{tetramer}$) of tetramer (black, calculated using the equilibrium dissociation constants listed in Fig. 4*C*). corr. coeff., correlation coefficient.

followed the disappearance of the monomer species rather than the appearance of fibrils and that the incorporation of monomeric htt^{ex1}Q$_{35}$ into fibrils had to be treated separately for the native reduced and Met^{7}O oxidized forms.

The NMR signal intensities of native reduced (I^{red}_{calc}) and Met^{7}O oxidized (I^{oxi}_{calc}) forms of htt^{ex1}Q$_{35}$ are calculated from

$$I^{red}_{calc} = A_0[1 - \alpha M_{red}(t)/fm_{tot}] \quad [2.1]$$

$$I^{oxi}_{calc} = A_0[1 - \alpha M_{oxi}(t)/(1-f)m_{tot}], \quad [2.2]$$

where f is the fraction of reduced species and $(1 - f)$ is the fraction of oxidized species in the NMR samples; the super/subscripts "red" and "oxi" pertain to reduced and oxidized

forms of htt$^{ex1}Q_{35}$, respectively; α is a scaling factor that takes into account that NMR intensities for the PRD do not decay to zero for fully aggregated samples; and A_0 is an overall scaling factor that accounts for possible errors in normalization of NMR intensities. Since the cross-peaks for the reduced and oxidized forms of the PRD are completely superimposed, the total, normalized calculated NMR signal intensity for the PRD domain, I_{calc}^{PRD}, is given by

$$I_{calc}^{PRD} = fI_{calc}^{red} + (1-f)I_{calc}^{oxi}, \quad [3]$$

while for Ser12, the cross-peak volume of the reduced species and the cross-peak intensity of the oxidized species are given by Eqs. **2.1** and **2.2**, respectively, with α set to 1 since these curves decay to zero.

The simplest kinetic scheme that incorporates both reduced and oxidized species of htt$^{ex1}Q_{35}$ and can fit all the data in Fig. 5*A* simultaneously is a model involving primary nucleation for the reduced form only, secondary nucleation, and elongation (Fig. 5*B*) (37), described by a system of three coupled ordinary differential equations

$$\frac{dP}{dt} = k_n[fm_{tot} - M_{red}(t)]^{nc} + \{k_2^{red}[fm_{tot} - M_{red}(t)]^{n2} + k_2^{oxi}[(1-f)m_{tot} - M_{oxi}(t)]^{n2}\}[M_{red}(t) + M_{oxi}(t)], \quad [4.1]$$

$$\frac{dM_{red}}{dt} = 2k_+^{red}[fm_{tot} - M_{red}(t)]P(t), \quad [4.2]$$

$$\frac{dM_{oxi}}{dt} = 2k_+^{oxi}[(1-f)m_{tot} - M_{oxi}(t)]P(t), \quad [4.3]$$

where $M(t)$ is the (mass) concentration of the fibrils (in monomer units); $P(t)$ is the (number) concentration of extendable ends of the fibril chain; nc is the order of primary nucleation (where the optimal value is 4; see below); $n2$ is the order of secondary nucleation (assumed to be 1); k_n is the primary nucleation rate constant (in units of $M^{(1-nc)}h^{-1}$); k_2 is the secondary nucleation rate constant (in units of $M^{-n2}h^{-1}$); k_+ is the chain elongation rate constant (in units of $M^{-1}h^{-1}$); and m_{tot} is the total concentration of monomers (of both oxidized and reduced forms). The first term in Eq. **4.1** describes the process of primary nucleation of the native reduced species (as discussed above, the Met^{7}O oxidized species does not undergo primary nucleation; see Fig. 3); the second and third terms in Eq. **4.1** describe first-order secondary nucleation involving reduced (k_2^{red}) and oxidized (k_2^{oxi}) species; and Eqs. **4.2** and **4.3** describe chain elongation by addition of a reduced and oxidized subunit, respectively. First-order secondary nucleation is dependent on the concentration $M(t)$ of fibrils (aggregates), as well as the monomer concentration, $m_{tot} - M_{red}(t)$, and hence assumes some interactions between the monomers and fibrils as described in ref. 31 and refs. 37–42. Specifically, secondary nucleation from monomers is catalyzed on the fibril surface and may result in branched fibrils or new fibrils (41).

Some residual seeding, manifested by nonzero values of $P(0)$ at the start of the NMR experiments, was inevitable, and the values of $P(0)$ were therefore optimized for each value of the nucleation order nc by an initial pass of minimization with A_0 in Eq. **2** set to 1 and the values of the rate constants optimized; in the second pass of minimization, the values of $P(0)$ were held fixed, and the values of A_0 were optimized together with the rate constants. Further, in the initial calculations, the value of the secondary nucleation rate constant k_2^{oxi} for the oxidized species optimized to a small and ill-determined value and was subsequently set to zero; that is, secondary nucleation only involved the reduced species, and the oxidized species only participated in chain elongation. If the contribution of the Met^{7}O oxidized form to elongation was omitted completely (i.e., f was set to 1 in Eqs. **4.1**–**4.3**), the experimental data could no longer be adequately reproduced owing to the presence of a significant amount (20%) of the Met^{7}O species in the 750-μM sample (*SI Appendix*, Fig. S5).

A grid search of the value of the primary nucleation order, nc, yielded a minimum χ^2 value at $nc = 4$ (Fig. 5*C*). Although a solution with only a slightly higher χ^2 could be obtained with $nc = 3$, the values of $P(0)$ required to fit the data were unrealistically large (*SI Appendix*, Fig. S4), ~20-fold higher than those obtained with $nc = 4$ (at 750 μM htt$^{ex1}Q_{35}$, $P(0)$ has a value of ~60 nM for $nc = 4$ vs. ~1.2 μM for $nc = 3$). The value of $nc = 4$ is fully consistent with transient tetramerization (Fig. 5) as a prerequisite for primary nucleation. Further, a log-log plot of the $t_{1/2}$ of monomer disappearance versus the population of tetramer at $t = 0$ was linear with a slope of ~ -0.56 (Fig. 5*D*, black), corresponding to a scaling exponent $\gamma \sim -1.7$ for the initial ($t = 0$) concentration of monomer (Fig. 5*D*, red). Numerical simulations of the scaling exponent γ for different combinations of the orders of primary (nc) and secondary ($n2$) nucleation for the rate constants reported in Table 1 show that the deviation of γ from a value of -2.0 ($-nc/2$) predicted for a model with primary nucleation only (43) is indicative of (a) the presence of a lower order ($n2 < 2$) secondary nucleation process and (b) that the order of primary nucleation, nc, is higher than 2, indirectly demonstrating the optimality of the choice of $nc = 4$ in our modeling.

The kinetic model described by Eqs. **4.1**–**4.3** yields excellent agreement between observed and calculated decay curves with a reduced χ^2 of 1.8 (Fig. 5 *A* and *C* and Table 1). In contrast, the best fit to a simpler kinetic model involving only primary nucleation (the first term in Eq. **4.1**) and elongation (Eqs. **4.2** and **4.3**) (37, 43) but lacking any type of secondary nucleation (the second term in Eq. **4.1**) yields a reduced χ^2 of >16 for any choice of $P(0)$, clearly not satisfying the experimental data. Invoking more complicated kinetic models of aggregation potentially applicable under acquiescent conditions of NMR measurements, including multistep secondary nucleation (44), saturating elongation (37), or reversible elongation (with the rate proportional to $-k_{off}P(t)$, where k_{off} is the rate constant of the reverse process) (37) did not prove warranted to reproduce the experimental data for htt$^{ex1}Q_{35}$ in Fig. 4. These findings are consistent with the linear log-log plots of $t_{1/2}$ for aggregation versus concentration of monomer and population of tetramer (Fig. 5*D*) (37).

The optimized values of the rate constants are listed in Table 1. The rate constant for elongation, k_+, is about 30% slower for the Met^{7}O sulfoxide species than the native reduced form. The concentration of extendable ends of the fibril chain at $t = 0$, $P(0)$, is proportional to the total concentration of htt$^{ex1}Q_{35}$ and ranges from ~2.5 nM at the lowest concentration of htt$^{ex1}Q_{35}$ employed (200 μM) to ~60 nM at the highest concentration (750 μM) (*SI Appendix*, Fig. S4, *Left* for $nc = 4$). Plots showing the time evolution of $M(t)$, $P(t)$, and the $M(t)/P(t)$ ratio, which is proportional to fibril length, is shown in Fig. 6. Although values of the individual rate constants are obtained from the fits (Table 1), the exact values of the rate constants are dependent on the choice of $P(0)$ as a consequence of strong correlations between the rate constant for elongation, k_+, and the initial number of extendable ends, $P(0)$ (37). It is therefore also useful to report the values of

Table 1. Rate constants for aggregation of htt^{ex1}Q$_{35}$ incorporating primary nucleation, elongation, and secondary nucleation

Rate constant	Native reduced	Met^{7}O oxidized
Primary nucleation (nc = 4), k_n ($M^{-3}h^{-1}$)	3.4 (±0.4) × 10^5	0*
Elongation, k_+ ($M^{-1}h^{-1}$)	7.8 (±0.3) × 10^5	5.7 (±0.3) × 10^5
Secondary nucleation ($n2$ = 1), k_2 ($M^{-1}h^{-1}$)	0.50 ± 0.02	0†
Products of nucleation and elongation,		
k_nk_+ (M^{-4} h^{-2})	2.7 (±0.04) × 10^{11}	0
k_2k_+ (M^{-2} h^{-2})	3.9 (±0.3) × 10^5	0

The reduced χ^2 for the global fit to all the data in Fig. 4 is 1.8. The optimized values of the coefficients α in Eq. **2**, which account for the presence of residual cross-peak intensity for the PRD in the htt^{ex1}Q$_{35}$ fibrils, are 0.79 ± 0.01, 0.92 ± 0.01, 0.90 ± 0.91, 0.94 ± 0.01, and 0.98 ± 0.01 at sample concentrations of 200, 300, 375, 500, and 750 μM, respectively. As expected, the values of α tend to increase slightly with sample concentration, likely reflecting higher viscosity (and hence higher relaxation rates and lower residual NMR signal intensities) of the fibrillized/aggregated samples at higher htt^{ex1}Q$_{35}$ concentrations.
*Set to zero since the Met^{7}O oxidized form does not undergo transient homo- or hetero-tetramerization and does not form fibrils on its own (see Fig. 3).
†k_2^{oxi} optimized to a small value that was not defined by the data (0.09 ± 0.10 $M^{-1}h^{-1}$). It was therefore set to 0 in subsequent calculations.

the products k_nk_+ and k_2k_+, 2.7 (±0.04) × 10^{11} $M^{-4}h^{-2}$ and 3.9 (±0.3) × 10^5 M^{-2} h^{-2}, respectively, which are much less affected by the set of initial conditions (37). Given that the units of k_nk_+ and k_2k_+ are related as M^{-2}, the values above imply that at the sample concentrations employed in this study, the primary and secondary processes contribute approximately equally to the nucleation stage of aggregation, with secondary nucleation dominating slightly. The concentration of native monomer (m_{red}) at which the secondary nucleation rate exceeds the primary nucleation rate can be obtained from the rate expressions for these two processes shown in Fig. 5*B* by solving the cubic equation $k_n m_{red}^3 = k_2 M$, where the concentration M of fibrils (in monomer units) is given by $(m_{tot} - m_{red} - m_{ox})$, and the concentration of the Met^{7}O species (m_{ox}) is given approximately by $m_{ox} \sim m_{red}(1-f)k_+^{ox}/fk_+^{red}$. For the 750-μM sample (where f = 0.80), the cross-over point after which secondary nucleation dominates is reached when the monomer concentration falls to ~540 μM; for the 500-, 375-, 300-, and 200-μM samples (where f = 0.93), the cross-over point is reached when the monomer concentration falls to 440, 345, 240, and 195 μM, respectively.

The ratio of $M(t)/P(t)$ reaches a maximum of ~2,000 at ~8 h for the 300- to 750-μM samples and at ~20 h for the 200-μM sample (Fig. 6, *Lower Right*). Assuming that the separation between polyQ β strands along the long axis of the fibril is ~4.8 Å, each monomer unit comprises a two-stranded antiparallel β-sheet connected by a β-hairpin (12, 19), and elongation occurs from both ends of the fibril, one can deduce that the fibrils reach a limiting length of ~1 μm, consistent with the electron microscopy image shown in Fig. 6 (*Inset* in lower right panel). The limiting fibril length is likely largely determined by secondary nucleation generating either branched or new fibrils.

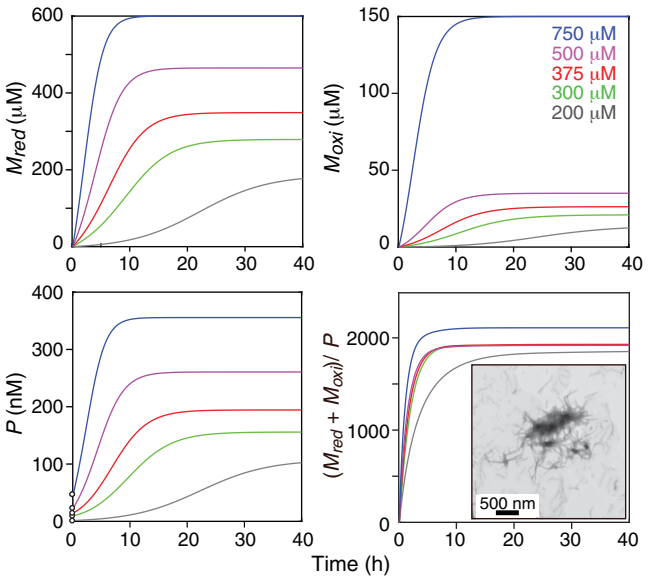

Fig. 6. Simulated time dependence of htt^{ex1}Q$_{35}$ fibrillization at 5 °C calculated using the optimized values of the rate constants given in Table 1. The fibril concentration in monomer units (M_{red} and M_{oxi}), the concentration of extendable ends (P), and the ratio of total fibril concentration (M_{red} + M_{oxi}) to extendable ends (P) are displayed. The latter ratio is proportional to fibril length. Assuming a separation of ~4.8 Å (the distance between strands in a β-sheet) between monomer units along the length of the fibril, the fibril length reaches a maximum of approximately ~1 μm, consistent with negative stain electron microscopy images of htt^{ex1}Q$_{35}$ fibrils (after 70-h incubation of 40 μM htt^{ex1}Q$_{35}$ at 5 °C), shown in the *Inset*.

Concluding Remarks. In an aggregating system where the NMR-visible monomeric species disappears completely over a period of hours, it is no longer feasible to carry out time-intensive exchange-based NMR experiments to unravel micro- to millisecond processes involving sparsely populated transient oligomeric states. Here, we show how a series of SOFAST-HMQC correlation spectra, each acquired in a few minutes, can be used to not only quantitatively analyze the mechanism and aggregation kinetics of the pathogenic htt^{ex1}Q$_{35}$ construct but can also simultaneously provide quantitative information on the interconversion kinetics and equilibria between NMR-visible monomer and NMR-invisible sparsely populated dimers and tetramers that take place on the microsecond timescale (Fig. 1).

While the timescale of the monomer-dimer-tetramer interconversion (50 to 70 μs) and the monomer-dimer equilibrium (K_1^{diss} ~60 mM) for htt^{ex1}Q$_{35}$ are very similar to those previously found in the shorter htt^{ex1}Q$_7$ construct that remains largely monomeric over a period of weeks, the stability of the htt^{ex1}Q$_{35}$ tetramer (K_2^{diss} ~11 μM) is enhanced two- to threefold relative to that of htt^{ex1}Q$_7$ (22, 23). Although the population of tetramer is only ~2.5% at a concentration (600 μM) close to the solubility limit of native htt^{ex1}Q$_{35}$, tetramerization is essential for subsequent primary nucleation, and indeed the Met^{7}O sulfoxide species that does not tetramerize also does not aggregate on its own (Fig. 3).

Nucleation and fibril formation occur cooperatively with residues in the N-terminal amphiphilic region, the polyQ tract, and the PRD domain, all following the same kinetics (Fig. 2 and *SI Appendix*, Fig. S2). Global quantitative analysis of the decay of NMR-visible species, over a range of concentrations spanning 200 to 750 μM, using a modification of the

formalism developed in ref. 31 that takes into account the presence of both native and Met^7O-oxidized species in the sample, shows that fibril formation involves three steps. For the native (reduced) $htt^{ex1}Q_{35}$ species, primary nucleation is a fourth-order process ($nc = 4$) dependent on the population of transient tetramers, while elongation and secondary nucleation are both monomer-dependent, first-order processes ($n2 = 1$). The Met^7O oxidized species, on the other hand, does not undergo either tetramerization or nucleation (both primary and secondary), but it is still incorporated into fibrils with an elongation rate constant k_+ that is only 25% smaller than that for native $htt^{ex1}Q_{35}$. It should be noted that dissection of the contributions of the native and Met^7O species can only be achieved by NMR, as these species give rise to separate signals in the NMR spectra, for example, in contrast to thioflavin T fluorescence assays of aggregation (31), which cannot distinguish between incorporation of chemically different components into fibrils.

How do the data on the kinetics of prenucleation tetramerization and aggregation of $htt^{ex1}Q_{35}$ relate to what is known about the structure and topology of the fibrils deduced from solid-state NMR, negative stain electron microscopy, and cryo-electron tomography (11–14, 16–21)? All these data suggest that the filamentous fibril comprises a rigid β-hairpin/β-sheet polyQ core extending along the long axis of the fibril. Both the partially ordered NT helices and the intrinsically disordered PRD domains lie on the outside of the fibril core, and the NT helices can be observed in subtomogram-averaged cryo-electron tomography images as a pseudoperiodic pattern of repeating lumps (21). Irrespective of whether neighboring β-hairpin/β-sheet subunits are oriented parallel or anti-parallel to one another (and both arrangements can presumably occur), the NT helices protruding on each side of the fibril must all be parallel to one another. The NT tetramer facilitates fibril nucleation by bringing the polyQ tracts into close proximity, thereby providing a template for polyQ self-association. The NT domain tetramer is a dimer of dimers: the dimer is an anti-parallel helical coiled-coil that assembles into a tetramer with the two dimers oriented approximately orthogonal to one another (24). The shortest distance between C-terminal ends of the NT domains, marking the beginning of the polyQ tract, occurs between dimer units in the tetramer (24). Thus, it seems likely that once fibril nucleation and elongation are initiated, dissociation of the NT domain tetramer takes place, leaving behind two interacting parallel coiled-coil helices that stabilize the stacking of successive layers of polyQ β-sheets on top of one another. The NT domain tetramer could potentially persist to generate wider fibril polymorphs by bridging two polyQ cores. In addition, serial addition of $htt^{ex1}Q_{35}$ monomers upon binding in an anti-parallel orientation to a preexisting, parallel NT helical dimer on the outside of a monofilament fibril will lead to secondary nucleation and subsequent branching.

Finally, we note that the approach used here to quantitatively analyze fast prenucleation oligomerization events simultaneously with the relatively slow kinetics of fibrillization should be widely applicable to many aggregating systems comprising intrinsically disordered or misfolded regions of proteins that form amyloid-like fibrils.

Experimental Procedures

Expression and Purification of $htt^{ex1}Q_{35}$. $htt^{ex1}Q_{35}$ was expressed as fusion protein with the immunoglobulin-binding domain of streptococcal protein G, GB1, attached to the N-terminal end separated by a factor Xa cleavage site (27). Uniform ^{15}N-labeling and cleavage of factor Xa to remove GB1 were carried out as described previously (22). Following the last step of purification, consisting of reverse-phase high performance liquid chromatography, the lyophilized $htt^{ex1}Q_{35}$ fractions were dissolved in a 1:1 (vol/vol) mixture of trifluoracetic acid (TFA) and hexafluoroisopropanol. As previously described by Chen and Wetzel (45), the latter procedure ensures removal of the vast majority of preexisting aggregates that can accelerate the aggregation of $htt^{ex1}Q_{35}$. The solvent mixture was then removed under a stream of N_2 gas, and the resulting peptide film was further lyophilized for 16 h to ensure complete removal of both solvents. Prior to final lyophilization, the protein film was dissolved in 0.1 mM TFA. Protein identity and completion of the cleavage reaction were checked by liquid phase chromatography coupled with electrospray mass spectrometry.

Oxidation of Met^7 to a sulfoxide was carried out by incubating the GB1–$htt^{ex1}Q_{35}$ fusion protein with 4% (vol/vol) H_2O_2 for ~2 h (35). Complete removal of H_2O_2 was achieved by exhaustive dialysis against a 50 mM Tris·HCl buffer, pH 8.0, and 100 mM NaCl (overnight at 4 °C), prior to factor Xa cleavage to remove GB1. Oxidation of Met^7 to a sulfoxide results in a +16 Da addition to the molecular mass of $htt^{ex1}Q35$ as verified by mass spectrometry.

NMR Sample Preparation. All NMR samples of $htt^{ex1}Q_{35}$ were prepared by first dissolving an aliquot of purified protein in 13.8 mM monobasic sodium phosphate buffer, pH 4.6, containing 50 mM NaCl and 10% D_20/90% H_2O (vol/vol). Upon resuspension, the protein solution was centrifuged at 20,000 × g for 25 min to remove any preformed insoluble aggregates. The pH of the buffer was subsequently adjusted to 6.5 by adding dibasic sodium phosphate for a final sodium phosphate concentration of 20 mM. Peptide concentrations were determined by ultraviolet absorbance at 205 nm (46) since $htt^{ex1}Q_{35}$ does not contain any tryptophan or tyrosine residues (24). Although the procedure of Chen and Wetzel (45) was implemented during purification to remove the majority of preexisting aggregates (see above), we found that some seeding was inevitable at the start of the NMR experiments. For this reason, following the formalism described in refs. 31 and 37, the concentration of extendable ends of the fibril chain, $P(0)$ present at $t = 0$, was initially optimized (together with the rate constants) in a first pass of minimization, as described in the main text.

NMR Spectroscopy. All NMR experiments were recorded at 5 °C using a Bruker 800 MHz Avance-III spectrometer, equipped with a TCI triple resonance z axis gradient cryogenic probe. All NMR data were processed using the NMRPipe/NMRDraw software suite (47) and analyzed with the program Sparky (48).

Fibrillization/aggregation of ^{15}N-labeled $htt^{ex1}Q_{35}$ was monitored by following ^1H-^{15}N cross-peak intensities (I, peak height) and peak volumes (V) from serially acquired two-dimensional (2D) ^1H-^{15}N SOFAST-HMQC spectra (see Figs. 1 and 4) using the pulse sequence of Schanda and Brutscher (29). Each 2D NMR spectrum was acquired using eight dummy scans, a recycle delay of 0.3 s, and a total of 100* × 2,048* complex data points in the indirect (^{15}N) and direct (^1H) dimensions with respective acquisition times of 44 and 104 ms. The number of scans (ns) used in all 2D ^1H-^{15}N SOFAST-HMQC experiments was 16, with the exception of the data acquired for the lowest concentration sample (200 μM) where ns = 32.

The total measurement time for each 2D ^{1}H-^{15}N SOFAST-HMQC spectrum was either ~6 or 12 min for ns = 16 or 32, respectively.

Measurement of Concentration-Dependent ^{15}N Exchange-Induced Chemical Shifts (δ_{ex}). Concentration-dependent changes in ^{15}N chemical shifts were measured on the 750-μM ^{15}N- htt^{ex1}Q$_{35}$ sample from a series of 2D ^{1}H-^{15}N SOFAST-HMQC spectra acquired over 15 h. At $t = 0$ h, the concentration of htt^{ex1}Q$_{35}$, m_{tot}, was 600 μM of native reduced htt^{ex1}Q$_{35}$ ($f = 0.8$) and 150 μM of the Met^{7}O oxidized form. The concentration of monomeric native htt^{ex1}Q$_{35}$ as a function of time, $m_{red}(t)$, was directly recast from the average normalized cross-peak intensity ($<I_{PRD}>$) of the PRD residues as $m_{red}(t) = fm_{tot}[< I_{PRD}(t) > - < I_{PRD}(\infty) >]/[< I_{PRD}(0) > - < I_{PRD}(\infty) >]$. For a given time-point t, ^{15}N-δ_{ex} is calculated as $\delta_{ex}(t) = \delta_{obs}(t) - \delta_{obs}^{ref}$, where $\delta_{obs}(t)$ is the observed chemical shift at a native monomer concentration $m_{red}(t)$, and δ_{obs}^{ref} is the chemical shift at a concentration of 10 μM htt^{ex1}Q$_{35}$. The errors in peak positions were determined as described previously (24, 49).

^{15}N-R_2 and ^{1}H$_N$-R_2 Relaxation Rates. Intrinsic (exchange-free) ^{15}N-$R_{1\rho}$ (with a 3-kHz spin lock) and ^{15}N-R_1 rates for the htt^{ex1}Q$_{35}$ monomer were measured on a 10-μM sample using the pulse sequences and procedures described previously (22, 50). A two–time point measurement was employed for $R_{1\rho}$, $R_{1\rho} = (1/\Delta t)\ln(I_2/I_1)$, where Δt is the difference in the two relaxation delays (2 and 60 ms) and I_n is the peak intensity at the corresponding relaxation delay ($n = 1$ or 2). The maximum relaxation delay was adjusted such that the NMR signal intensity was attenuated by ~50% of the initial intensity. The intrinsic ^{15}N-R_2 values were calculated as $(R_{1\rho} - R_1\cos^2\theta)/\sin^2\theta$, where θ is the angle between the effective spin-lock field and the external magnetic (B_0) field. The two relaxation delays used for the ^{15}N-R_1 measurements were 40 and 320 ms. Free precession ^{1}H$_N$-R_2 relaxation rates for the htt^{ex1}Q$_{35}$ monomer were also obtained using a two–time point measurement (relaxation delays of 4 and 24 ms) on the 10 μM sample using standard procedures (51). The intrinsic values of ^{15}N-R_2 and ^{1}H$_N$-R_2 for the monomer are provided in the *SI Appendix*, Table S2.

Global Fitting of Concentration Dependence of ^{15}N-δ_{ex} and ^{1}H^{N}-^{15}N Cross-Peak *V/I* Ratios. The experimental concentration dependence of ^{15}N-δ_{ex} and ^{1}H-^{15}N cross peak *V/I* ratios were fit simultaneously by minimization of the error function

$$F = \alpha_1 \sum_i \left(\frac{\delta^{obs,i,k} - \delta^{calc,i,k}}{\sigma_\delta^{i,k}} \right)^2 + \alpha_2 \sum_j \left(\frac{(V/I)^{obs,j,k} - (V/I)^{calc,j,k}}{\sigma_{(V/I)}^{j,k}} \right)^2, \quad [5]$$

where the first and second terms correspond to the differences between the observed (obs) and calculated (calc) ^{15}N-δ_{ex} and *V/I* values, respectively, for residues i or j measured at each concentration k of the monomer; σ values are the experimental errors; and the coefficients α_1 and α_2 are empirically determined weighting factors set to 1 and 0.05, respectively. The set of global variable parameters in the minimization of the target function comprised $\{K_1^{diss}; K_2^{diss}; k_{-1}\}$ (see Fig. 3*C* and the main text for definitions). The space of local (residue-specific) parameters included $\{\Delta\omega^H; \Delta\omega^N(F10), \alpha'; \varphi^N\}$, where $\Delta\omega^H$ values are the differences between the chemical shifts of the dimer/tetramer and the monomeric species for ^{1}H$_N$; $\Delta\omega^N(F10)$ is the difference in chemical shift of the dimer/tetramer and monomeric species for ^{15}N of Phe10 (which was not previously determined for htt^{ex1}Q$_7$ (22); α' is a scaling factor (see Eq. **1**); and φ^N is an offset that accounts for small errors in the reference value of the chemical shift, δ_{ref} (measured for 10 μM htt^{ex1}Q$_{35}$). The other $\Delta\omega^N$ values were set to those previously determined from htt^{NT}Q$_7$ (22) on the reasonable assumption that the structure of the helical coiled-coil tetramer formed by the NT domain is not affected by the length of the polyQ domain. Since concentration-dependent ^{15}N-$R_{1\rho}$ measurements (23) are not feasible on a rapidly aggregating system, a lower limit on k_{-2} was established using a grid search (*SI Appendix*, Fig. S3). The values of $\Delta\omega^N$ and $\Delta\omega^H$ are listed in *SI Appendix*, Table S3.

The 3 × 3 matrix $\tilde{R}$ for monomer ↔ dimer ↔ tetramer interconversion that enters into the calculation of ^{15}N-δ_{ex} and the ^{1}H-^{15}N cross-peak *V/I* ratios (Eq. **1**) is given by $\tilde{R} = \tilde{R}^{CS} + \tilde{R}^{rel} + \tilde{R}^{ex}$, where $\tilde{R}^{CS}$, $\tilde{R}^{rel}$ and $\tilde{R}^{ex}$ are the chemical shift, intrinsic relaxation, and chemical exchange contributions, respectively, given by $\tilde{R}^{CS} = (0\ \ 0\ \ 0;\ \ 0\ \ \Delta\omega^{E2}\ \ 0;\ \ 0\ \ 0\ \ \Delta\omega^{E4})$, $\tilde{R}^{rel} = (R_2^E\ \ 0\ \ 0;\ \ 0\ \ R_2^{E2}\ \ 0;\ \ 0\ \ 0\ \ R_2^{E4})$, and $\tilde{R}^{ex} = (-k_1^{app}\ \ k_{-1}\ \ 0;\ k_1^{app}\ -k_{-1} - k_2^{app}\ \ k_{-2};\ \ 0\ \ k_2^{app} - k_{-2})$. $\Delta\omega$ values are expressed in rad/s; the superscripts E, E2, and E4 stand for monomer, dimer, and tetramer, respectively; R_2^{E2} and R_2^{E4} are assumed to be equal to $2R_2^E$ and $4R_2^E$, respectively; and k_1^{app} and k_2^{app} are apparent pseudofirst-order rate constants given by $2k_1[E]$ and $2k_1k_2[E]2/k_{-1}$, respectively. δ_{ex} values were calculated from the imaginary part of the smallest (by absolute magnitude) eigenvalue of the relaxation matrix $\tilde{R}$: δ_{ex} (Hz) = $Im(\min[eig\{\tilde{R}\}])/2\pi$, where $eig\{\tilde{R}\}$ is a vector of complex eigenvalues of $\tilde{R}$.

With regard to the calculation of *V/I* ratios from Eq. **1**, ^{15}N-R_2 and ^{1}H$_N$-R_2 are used for the ^{15}N (indirect) and ^{1}H (direct) dimensions, respectively, of the ^{1}H-^{15}N SOFAST-HMQC experiment. Further, the intrinsic ^{15}N-^{1}H double quantum (^{15}N-$R_{2,DQ}$) and zero-quantum (^{15}N-$R_{2,ZQ}$) transverse relaxation rates are assumed to be equal to the single-quantum ^{15}N-R_2 values measured experimentally, as described in the previous section (also see *SI Appendix*).

The uncertainties in the values of the optimized parameters corresponding to confidence intervals of ±1 SD were determined from the variance-covariance matrix of the nonlinear fit. Uncertainties in the rate constants recalculated from the optimized parameters (k_1 and k_2) were determined by standard procedures for error propagation. All calculations were performed using an in-house program written in MATLAB (MathWorks Inc) (53).

As discussed previously (22, 24), prenucleation oligomerization involves a very fast on-pathway (τ_{ex} ~20 to 50 μs) leading to formation of a tetramer via a productive helical coiled-coil dimer and an order of magnitude slower off-pathway (τ_{ex} ~750 μs) leading to a nonproductive dimer (E2*) that does not oligomerize further. To ascertain that omission of the slower off-pathway does not affect the above analysis or the values of the on-pathway rate and equilibrium constants, we also fit a branched four-state scheme to the experimental δ_{ex} and *V/I* data in Fig. 3 *A* and *B* with the values of the rate (k_{-3} = 1350 s^{-1} and equilibrium (K_3^{diss} = 0.23 M) constants and ^{15}N-$\Delta\omega^{E2*}$ held fixed to those determined previously (22) and with ^{1}H-$\Delta\omega^{E2*}$ = ^{1}H-$\Delta\omega^{E2}/2$. The values of k_{-1}, K_1^{diss}, and K_2^{diss} obtained for the branched four-state model are 4.0 (±1.1) × 10^4 s^{-1}, 67 ± 15 mM, and

10 ± 5 μM, respectively, which are identical within error to the values of 3.8 (± 1.1) $\times 10^4$ s^{-1}, 62 ± 12 mM, and 11 ± 4 μM, respectively, obtained for the three-state monomer ↔ dimer ↔ tetramer scheme with the nonproductive pathway omitted. Further, the lower limit for k_{-2} for both three- and four-state models is 2×10^4 s^{-1}. Thus, the concentration dependence of *VI I*, like that of $R_{1\rho}$ (23), is not affected by the presence or absence of slower off-pathway dimerization.

Global Fitting of Fibrillization/Aggregation Kinetics. The decay profiles of averaged normalized NMR signal intensities, I, measured at multiple total protein concentrations, m_{tot}, were analyzed by the minimization of the sum of squares target function, $F = \sum_i \sum_j [(I_{obs} - I_{calc})/\sigma]^2$, where I_{obs} and I_{calc} are the experimental and calculated normalized intensities, respectively; σ is the SE (assumed to be equal to one-half of the SD of the distributions of peak intensities for the 11 PRD residues used); the index j in the inner sum of F runs over all time points, and the index i runs over all sample concentrations analyzed. The differential equations describing the kinetics of aggregation (Eq. **4**) were integrated numerically using an explicit Runge-Kutta [1,4] formula implemented in the Matlab "ode45" solver for nonstiff ordinary differential equations (52), with integration steps corresponding to the time points in the NMR experiments.

Visualization of htt^{ex1}Q$_{35}$ Fibrils by Negative Stain Electron Microscopy. Complete aggregation of 4- or 40-μM htt^{ex1}Q$_{35}$ was achieved in an Eppendorf tube at 5 °C in ∼70 h. Following fibrillization, ∼10 μL htt^{ex1}Q$_{35}$ solution was blotted onto carbon grids (carbon-coated copper grids; Ultrathin Carbon Film/Holey Carbon; Ted Pella) and left for 1 min. The grids were quickly rinsed with deionized H_2O and stained with 2% uranyl acetate for 20 s. Electron microscopy images were recorded using a FEI Tecnai T12 electron microscope at 120 kV equipped with a Gatan US 1000 CCD camera.

Data Availability. Experimental NMR data (concentration dependence of ^{15}N-exchange–induced shifts and ^{1}H-^{15}N cross-peak *V/I* ratios and aggregation kinetics data from the time dependence of ^{1}H-^{15}N cross-peak intensities and volumes) in digital format, together with the MATLAB scripts used in global fitting, have been deposited on figshare (DOI: 10.6084/m9.figshare.19694983) (53). All experimental data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Drs. James Baber, Dan Garrett, and Jinfa Ying for technical support. This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases at the NIH (Grant DK29023 to G.M.C.).

Author affiliations: [a]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, MD 20892-0520

Author contributions: A.C., V.T., and G.M.C. designed research; A.C. and F.T. performed research; A.C., V.T., F.T., and G.M.C. analyzed data; and A.C., V.T., and G.M.C. wrote the paper.

1. J. M. Andresen *et al.*; US-Venezuela Collaborative Research Group; HD MAPS Collaborative Research Group, The relationship between CAG repeat length and age of onset differs for Huntington's disease patients with juvenile onset or adult onset. *Ann. Hum. Genet.* **71**, 295–301 (2007).
2. C. A. Ross, S. J. Tabrizi, Huntington's disease: From molecular pathogenesis to clinical treatment. *Lancet Neurol.* **10**, 83–98 (2011).
3. G. P. Bates *et al.*, Huntington disease. *Nat. Rev. Dis. Primers* **1**, 15005 (2015).
4. C. Zuccato, M. Valenza, E. Cattaneo, Molecular mechanisms and potential therapeutical targets in Huntington's disease. *Physiol. Rev.* **90**, 905–981 (2010).
5. K. Kar, M. Jayaraman, B. Sahoo, R. Kodali, R. Wetzel, Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* **18**, 328–336 (2011).
6. R. Wetzel, Exploding the repeat length paradigm while exploring amyloid toxicity in Huntington's disease. *Acc. Chem. Res.* **53**, 2347–2357 (2020).
7. C. Landles *et al.*, Proteolysis of mutant huntingtin produces an exon 1 fragment that accumulates as an aggregated protein in neuronal nuclei in Huntington disease. *J. Biol. Chem.* **285**, 8808–8823 (2010).
8. K. Sathasivam *et al.*, Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2366–2370 (2013).
9. E. Scherzinger *et al.*, Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4604–4609 (1999).
10. W. C. Duim, Y. Jiang, K. Shen, J. Frydman, W. E. Moerner, Super-resolution fluorescence of huntingtin reveals growth of globular species into short fibers and coexistence of distinct aggregates. *ACS Chem. Biol.* **9**, 2767–2778 (2014).
11. C. L. Hoop *et al.*, Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1546–1551 (2016).
12. J. C. Boatz *et al.*, Protofilament structure and supramolecular polymorphism of aggregated mutant Huntingtin exon 1. *J. Mol. Biol.* **432**, 4722–4744 (2020).
13. V. N. Sivanandam *et al.*, The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils. *J. Am. Chem. Soc.* **133**, 4558–4566 (2011).
14. R. Schneider *et al.*, Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol.* **412**, 121–136 (2011).
15. C. L. Hoop *et al.*, Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance. *Biochemistry* **53**, 6653–6666 (2014).
16. J. M. Isas, R. Langen, A. B. Siemer, Solid-state nuclear magnetic resonance on the static and dynamic domains of Huntingtin exon-1 fibrils. *Biochemistry* **54**, 3942–3949 (2015).
17. H. K. Lin *et al.*, Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core. *Nat. Commun.* **8**, 15462 (2017).
18. B. G. Caulkins, S. A. Cervantes, J. M. Isas, A. B. Siemer, Dynamics of the proline-rich C-terminus of huntingtin exon-1 fibrils. *J. Phys. Chem. B* **122**, 9507–9515 (2018).
19. I. Matlahov, P. C. van der Wel, Conformational studies of pathogenic expanded polyglutamine protein deposits from Huntington's disease. *Exp. Biol. Med. (Maywood)* **244**, 1584–1595 (2019).
20. F. J. B. Bäuerlein *et al.*, In situ architecture and cellular interactions of polyQ inclusions. *Cell* **171**, 179–187.e10 (2017).
21. J. G. Galaz-Montoya, S. H. Shahmoradian, K. Shen, J. Frydman, W. Chiu, Cryo-electron tomography provides topological insights into mutant huntingtin exon 1 and polyQ aggregates. *Commun. Biol.* **4**, 849 (2021).
22. A. Ceccon, V. Tugarinov, R. Ghirlando, G. M. Clore, Abrogation of prenucleation, transient oligomerization of the Huntingtin exon 1 protein by human profilin I. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5844–5852 (2020).
23. A. Ceccon, V. Tugarinov, G. M. Clore, Kinetics of fast tetramerization of the Huntingtin Exon 1 protein probed by concentration-dependent on-resonance $R_{1\rho}$ measurements. *J. Phys. Chem. Lett.* **11**, 5643–5648 (2020).
24. S. A. Kotler *et al.*, Probing initial transient oligomerization events facilitating Huntingtin fibril nucleation at atomic resolution by relaxation-based NMR. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3562–3571 (2019).
25. A. Ceccon, V. Tugarinov, G. M. Clore, TiO_2 nanoparticles catalyze oxidation of Huntingtin exon 1-derived peptides impeding aggregation: A quantitative NMR study of binding and kinetics. *J. Am. Chem. Soc.* **141**, 94–97 (2019).
26. M. A. Wälti, S. A. Kotler, G. M. Clore, Probing the interaction of Huntingtin exon-1 polypeptides with the chaperonin nanomachine GroEL. *ChemBioChem* **22**, 1985–1991 (2021).
27. A. Ceccon, V. Tugarinov, G. M. Clore, Quantitative exchange NMR-based analysis of Huntingtin-SH3 interactions suggests an allosteric mechanism of inhibition of huntingtin aggregation. *J. Am. Chem. Soc.* **143**, 9672–9681 (2021).
28. S. Vieweg *et al.*, The Nt17 domain and its helical conformation regulate the aggregation, cellular properties and neurotoxicity of mutant Huntingtin exon 1. *J. Mol. Biol.* **433**, 167222 (2021).
29. P. Schanda, B. Brutscher, Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127**, 8014–8015 (2005).
30. P. Schanda, E. Kupce, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **33**, 199–211 (2005).
31. S. I. Cohen *et al.*, Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9758–9763 (2013).
32. R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon Press, Oxford, NY, 1987).
33. K. Pervushin, B. Vögeli, A. Eletsky, Longitudinal ^{1}H relaxation optimization in TROSY NMR spectroscopy. *J. Am. Chem. Soc.* **124**, 12898–12902 (2002).
34. H. S. Atreya, T. Szyperski, G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9642–9647 (2004).
35. A. Ceccon *et al.*, Interaction of Huntingtin exon-1 peptides with lipid-based micellar nanoparticles probed by solution NMR and Q-band pulsed EPR. *J. Am. Chem. Soc.* **140**, 6199–6202 (2018).
36. H. M. McConnell, Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430–431 (1958).
37. G. Meisl *et al.*, Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc.* **11**, 252–272 (2016).
38. S. I. Cohen *et al.*, Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *J. Chem. Phys.* **135**, 065105 (2011).
39. S. I. Cohen, M. Vendruscolo, C. M. Dobson, T. P. Knowles, Nucleated polymerization with secondary pathways. II. Determination of self-consistent solutions to growth processes described by non-linear master equations. *J. Chem. Phys.* **135**, 065106 (2011).
40. S. I. Cohen, M. Vendruscolo, C. M. Dobson, T. P. Knowles, Nucleated polymerization with secondary pathways. III. Equilibrium behavior and oligomer populations. *J. Chem. Phys.* **135**, 065107 (2011).
41. P. Arosio, T. P. Knowles, S. Linse, On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys.* **17**, 7606–7618 (2015).
42. M. Törnquist *et al.*, Secondary nucleation in amyloid formation. *Chem. Commun. (Camb.)* **54**, 8667–8684 (2018).

43. F. Oosawa, S. Asakura, *Thermodynamics of the Polymerization of Proteins* (Molecular Biology, Academic Press, London, UK, 1975).
44. G. Meisl *et al.*, Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9384–9389 (2014).
45. S. Chen, R. Wetzel, Solubilization and disaggregation of polyglutamine peptides. *Protein Sci.* **10**, 887–891 (2001).
46. N. J. Anthis, G. M. Clore, Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci.* **22**, 851–858 (2013).
47. F. Delaglio *et al.*, NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
48. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
49. G. Kontaxis, G. M. Clore, A. Bax, Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J. Magn. Reson.* **143**, 184–196 (2000).
50. N. A. Lakomek, J. Ying, A. Bax, Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).
51. L. Deshmukh, V. Tugarinov, J. M. Louis, G. M. Clore, Binding kinetics and substrate selectivity in HIV-1 protease-Gag interactions probed at atomic resolution by chemical exchange NMR. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9855–E9862 (2017).
52. L. F. Shampine, M. W. Reichelt, The Matlab ODE suite. *SIAM J. Sci. Comput.* **18**, 1–22 (1997).
53. A. Ceccon, V. Tugarinov, F. Torricella, G. M. Clore, Digital data deposition and Matlab fitting scripts for "Quantitative NMR analysis of the kinetics of pre-nucleation oligomerization and aggregation of a pathogenic exon 1-derived huntingtin polypeptide. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2207690119 (2022)". figshare. https://figshare.com/articles/dataset/Digital_data_Deposition_and_Matlab_fitting_scripts_for_the_paper_entitled_Quantitative_NMR_analysis_of_the_kinetics_of_pre-nucleation_oligomerization_and_aggregation_of_a_pathogenic_exon_1-derived_huntingtin_polypeptide_Proc_Natal_Acad_Sci_/19694983. Deposited 2 May 2022.