

# Nucleation of Huntingtin Aggregation Proceeds via Conformational Conversion of Pre-Formed, Sparsely-Populated Tetramers

Francesco Torricella, Vitali Tugarinov,\* and G. Marius Clore\*

Pathogenic huntingtin exon-1 protein (htt<sup>ex1</sup>), characterized by an expanded polyglutamine tract located between the N-terminal amphiphilic region and a C-terminal polyproline-rich domain, forms fibrils that accumulate in neuronal inclusion bodies, and is associated with a fatal, autosomal dominant neurodegenerative condition known as Huntington's disease. Here a complete kinetic model is described for aggregation/fibril formation of a htt<sup>ex1</sup> construct with a 35-residue polyglutamine repeat, htt<sup>ex1</sup>Q<sub>35</sub>. Using exchange NMR spectroscopy, it is previously shown that the reversible formation of a sparsely-populated tetramer of the N-terminal amphiphilic domain of htt<sup>ex1</sup>Q<sub>35</sub>, comprising a D<sub>2</sub> symmetric four-helix bundle, occurs on the microsecond time-scale and is a prerequisite for subsequent nucleation and fibril formation on a time scale that is many orders of magnitude slower (hours). Here a unified kinetic model of htt<sup>ex1</sup>Q<sub>35</sub> aggregation is developed in which fast, reversible tetramerization is directly linked to slow irreversible fibril formation via conversion of pre-equilibrated tetrameric species to "active", chain elongation-capable nuclei by conformational re-arrangement with a finite, monomer-independent rate. The unified model permits global quantitative analysis of reversible tetramerization and irreversible fibril formation from a time series of <sup>1</sup>H-<sup>15</sup>N correlation spectra recorded during the course of htt<sup>ex1</sup>Q<sub>35</sub> aggregation.

polyproline repeats.<sup>[1]</sup> Extension of the polyQ region beyond ~35 glutamines, by CAG expansion within exon-1 of *HTT*, is responsible for Huntington's disease,<sup>[1-3]</sup> a fatal neurodegenerative condition associated with htt<sup>ex1</sup> aggregation and fibril accumulation within neuronal inclusion bodies.<sup>[4-6]</sup> Solid-state NMR and cryo-electron microscopy studies have shown that htt<sup>ex1</sup> fibrils consist of a central rigid  $\beta$ -hairpin/ $\beta$ -sheet polyQ core stabilized by a network of hydrogen bonds between glutamine side-chains, with NT helices and highly mobile PRD domains located on the outside of the fibril.<sup>[7-13]</sup> Quantitative exchange-based solution NMR studies revealed that a very slowly aggregating htt<sup>ex1</sup> construct with a 7-residue glutamine repeat (htt<sup>ex1</sup>Q<sub>7</sub>), as well as a shorter construct (htt<sup>NT</sup>Q<sub>7</sub>) that lacks the PRD domain, undergo transient tetramerization of the NT region on the microsecond time scale,<sup>[14-16]</sup> thereby increasing the local concentration of the polyQ tracts and providing a template for nucleation.

Reversible, microsecond time-scale oligomerization generating sparsely-populated tetramers of the htt<sup>ex1</sup> NT domain, occurs *prior to* nucleation,<sup>[15]</sup> but is critical for the formation of htt<sup>ex1</sup> fibrils which occurs on a much slower time scale (minutes to hours),<sup>[17]</sup> as evidenced by the fact that reduction of the tetramer population by a variety of mechanisms blocks fibril formation.<sup>[15,17,18]</sup> In our earlier study of pre-nucleation tetramerization and aggregation of a htt<sup>ex1</sup> construct with a 35-residue polyQ repeat, htt<sup>ex1</sup>Q<sub>35</sub>,<sup>[17]</sup> we developed a set of tools for the quantitative analysis of both the kinetics of pre-nucleation tetramerization and fibril formation from a series of 2D <sup>1</sup>H-<sup>15</sup>N band-selective optimized flip-angle short-transient heteronuclear multiple quantum coherence (SOFAST-HMQC)<sup>[19-20]</sup> NMR spectra. The kinetics of reversible tetramerization is derived from the concentration dependence of <sup>15</sup>N/<sup>1</sup>H<sub>N</sub> chemical shifts ( $\delta_{ex}$ ) and <sup>1</sup>H-<sup>15</sup>N cross-peak volume/intensity (V/I) ratios for residues located in the NT region of htt<sup>ex1</sup>Q<sub>35</sub>, while the kinetics of irreversible fibril formation is afforded by the decay of <sup>1</sup>H-<sup>15</sup>N cross-peak intensities within the PRD domain, as the latter are not affected by the fast exchange processes (on the chemical shift time scale) associated with oligomerization.<sup>[17]</sup>

## 1. Introduction

The huntingtin htt<sup>ex1</sup> protein, encoded by exon-1 of the huntingtin gene *HTT*, comprises a 16-residue N-terminal amphiphilic domain (NT), a poly-glutamine stretch (polyQ<sub>n</sub>) of variable length *n*, and a proline-rich domain (PRD) containing two

F. Torricella, V. Tugarinov, G. M. Clore  
Laboratory of Chemical Physics  
National Institute of Diabetes and Digestive and Kidney Diseases  
National Institutes of Health  
Bethesda, MD 20892-0520, USA  
E-mail: vitali.tugarinov@nih.gov; mariusc@mail.nih.gov

The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adv.202309217>

Published 2024. This article is a U.S. Government work and is in the public domain in the USA. Advanced Science published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adv.202309217

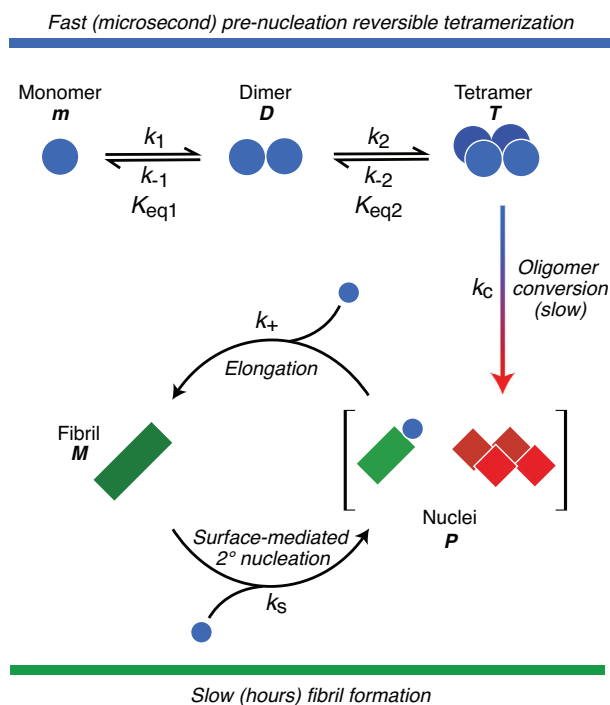
In our earlier study,<sup>[17]</sup> the kinetic analysis of irreversible aggregation of htt<sup>ex1</sup>Q<sub>35</sub> followed a conventional, “classical” approach,<sup>[21–22]</sup> with the rate of primary nucleation represented by an irreversible reaction of oligomerization of order 4. As a result, no direct connection was established in this treatment between fast, reversible tetramerization and the much slower, irreversible process of primary nucleation as, at the onset of aggregation, the concentration of tetramers exceeds that of nascent nuclei by 2-to-3 orders of magnitude.<sup>[17]</sup> Here, we present a unified kinetic model of htt<sup>ex1</sup>Q<sub>35</sub> aggregation that establishes a direct link between fast, reversible tetramerization and slow irreversible fibrillization. The unified model involves the conversion of pre-formed tetramers to “active”, elongation-capable nuclei by conformational re-arrangement with a finite, monomer-independent rate, and quantitatively accounts for the  $\delta_{ex}$  and V/I data and aggregation profiles obtained from a series of SOFAST-HMQC <sup>1</sup>H-<sup>15</sup>N correlation spectra acquired during the course of aggregation. The unified kinetic model presented here establishes a physically meaningful framework for the comprehensive, quantitative interpretation of htt<sup>ex1</sup> aggregation data. Further, the combined analysis of the kinetics of fast, reversible htt<sup>ex1</sup>Q<sub>35</sub> tetramerization and slow, irreversible aggregation provides new insights into the structural properties of htt<sup>ex1</sup>Q<sub>35</sub> nuclei that may potentially be useful for inhibiting the aggregation process.

## 2. Results and Discussion

### 2.1. Unified Kinetic Model of htt<sup>ex1</sup> Pre-Nucleation and Fibril Formation

Figure 1 shows the unified kinetic scheme, in which reversible tetramerization is coupled to irreversible fibril formation. The monomers of htt<sup>ex1</sup>Q<sub>35</sub> (*m*) interconvert on a timescale of ~30  $\mu$ s with sparsely populated dimeric species *D*, which can, in turn, interconvert with sparsely-populated tetramers *T* on a timescale of < 100  $\mu$ s. These two reversible oligomerization reactions are characterized by equilibrium constants  $K_{eq,1}$  and  $K_{eq,2}$ , respectively (see Supporting Information). Note that a minor “off-pathway” branch involving a non-productive dimeric species that cannot further oligomerize,<sup>[14]</sup> is present in all htt<sup>ex1</sup> variants, but can be neglected for the purposes of this work<sup>[16–17]</sup> (and is therefore not shown in the scheme of Figure 1). The tetramer *T* is slowly and irreversibly converted (with a rate constant  $k_c$ ) in a monomer-independent manner into elongation-capable, “active” primary nuclei, followed by monomer-dependent elongation ( $k_+$ ) to produce mature fibrils *M*. Surface-mediated secondary nucleation<sup>[23–25]</sup> ( $k_s$ ) involves interactions of htt<sup>ex1</sup>Q<sub>35</sub> monomers *m* with fibrils *M* to produce secondary nuclei that, although physically distinct from the primary ones (i.e., converted tetramers), add to the total pool of nuclei characterized by the number concentration of extendable ends, *P* (Figure 1).

Analysis of the kinetics of oligomerization and fibril formation depicted in Figure 1, assumes that the equilibrium  $m \rightleftharpoons D \rightleftharpoons T$  is established “instantaneously”, and that the interconversion of oligomeric species at equilibrium occurs fast on the timescale of aggregation; that is the dependence of the concentration of tetramers *T* on the concentration of monomers *m*, follows the same law,  $f(m(t))$ , for each time-point in the course of aggrega-



**Figure 1.** Schematic of the unified kinetic model of fast pre-nucleation tetramerization coupled with slow monomer-independent conversion, chain elongation and fibril surface-mediated secondary nucleation steps of htt<sup>ex1</sup>Q<sub>35</sub> fibrillization.  $k_1$  and  $k_2$  are association rate constants from monomer to dimer and from dimer to tetramer, respectively;  $k_{-1}$  and  $k_{-2}$  are the respective corresponding dissociation rate constants;  $K_{eq,1}$  and  $K_{eq,2}$  are the dimerization and tetramerization equilibrium constants, respectively;  $k_c$  is the first order rate constant for the unimolecular conversion of pre-nucleation transient tetramer *T* to nuclei *P*; and  $k_+$  and  $k_s$  are the rate constants for fibril elongation and surface-mediated secondary nucleation, respectively. The equilibrium  $m \rightleftharpoons D \rightleftharpoons T$  is established “instantaneously”:  $K_{eq,1} = [D]/[m]^2 = k_1/k_{-1}$ ;  $K_{eq,2} = [T]/[D]^2 = k_2/k_{-2}$  and  $[T] = K_{eq,2}[D]^2 = K_{eq,2}(K_{eq,1})^2 = (k_2/k_{-2})(k_1/k_{-1})^2[m]^4$ .

tion. The time course of aggregation can then be described by two coupled ordinary differential equations:<sup>[17,22]</sup>

$$\frac{dP}{dt} = k_c f(m(t)) + k_s m(t)^{n_2} M(t) \quad (1)$$

$$\frac{dM}{dt} = 2k_+ m(t) P(t) \quad (2)$$

where  $f(m(t))$  is the time-dependent concentration of tetramers;  $P(t)$ , is the number concentration of extendable ends of the nascent fibrils (nuclei);  $M(t)$ , is the mass concentration of mature fibrils (in monomer units); and the concentration of monomer is given by  $m(t) = m_{tot} - M(t)$ , where  $m_{tot}$  is the total concentration of htt<sup>ex1</sup>Q<sub>35</sub>.  $k_c$  is the rate constant for the conversion of tetramers to “active” nuclei in units of h<sup>-1</sup>;  $n_2$ , the order of secondary nucleation (in this instance  $n_2 = 1$ );  $k_s$ , the secondary nucleation rate constant in units of M<sup>- $n_2$ h<sup>-1</sup></sup>; and  $k_+$ , the chain elongation rate constant in units of M<sup>-1</sup>h<sup>-1</sup>. Assuming highly skewed fractional populations of the oligomeric states at equilibrium (i.e.,  $p_D$ ,

$p_T \ll p_m$ ),  $f[m(t)]$  can be expressed through the equilibrium/rate constants as:

$$f[m(t)] = (K_{\text{eq},1})^2 K_{\text{eq},2} m(t)^4 = \left(\frac{k_1}{k_{-1}}\right)^2 \left(\frac{k_2}{k_{-2}}\right) m(t)^4 \quad (3)$$

According to Equation (3), the first term on the right-hand side of Equation (1),  $k_c f[m(t)]$ , describing the monomer-independent conversion of all pre-equilibrated tetramers  $T$  to elongation-capable tetramers via conformational re-arrangement, is related to the conventional rate constant of primary nucleation of order 4 (i.e.,  $k_n m(t)^4$  where  $k_n$  is the primary nucleation rate constant) through the simple expression,  $k_n = (k_1/k_{-1})^2 (k_2/k_{-2}) k_c$ . The expressions for the rates of fibril-mediated secondary nucleation (second term in Equation 1), as well as that of chain elongation (Equation 2), are not affected by the current model and remain the same as in our previous treatment.<sup>[17]</sup>

## 2.2. Global Fitting of the Experimental NMR Data

Our experimental NMR strategy follows closely that described in our previous study of htt<sup>ex1</sup>Q<sub>35</sub> aggregation.<sup>[17]</sup> The experimental data together with the global best fits for the unified kinetic model (Figure 1 and Equations 1, 2 and 3) to the concentration-dependent <sup>15</sup>N/<sup>1</sup>H<sub>N</sub>  $\delta_{\text{ex}}$  exchange-induced shifts and  $V/I$  ratios and the decays in monomeric signal intensity (aggregation profiles) at three initial concentrations (160, 250, and 420  $\mu\text{M}$ ) of monomeric htt<sup>ex1</sup>Q<sub>35</sub> obtained from the SOFAST <sup>1</sup>H-<sup>15</sup>N HMQC time series are shown in Figures 2 and 3 (Figure S1, Supporting Information), respectively (see the “Experimental Section” for details of the global fitting procedure and analytical methods used to calculate  $\delta_{\text{ex}}$  and  $V/I$ ). While NMR cross-peak volumes are “immune” to chemical exchange line broadening associated with htt<sup>ex1</sup>Q<sub>35</sub> oligomerization, the cross-peak maximal intensities (heights;  $I$ ) are affected by exchange line broadening. As a result, the volumes decrease monotonically with decreasing monomer concentration during the course of htt<sup>ex1</sup>Q<sub>35</sub> aggregation, whereas the intensities of the cross-peaks involved in exchange initially increase reaching a maximum, before decreasing (Figure 2B, Upper Row). This behavior is a direct consequence of chemical exchange line broadening associated with sub-millisecond tetramerization. As the fractional population of the tetramer  $T$  has a cubic dependence on the concentration of the monomer  $m$ , exchange line broadening is reduced initially, and the concomitant increase of  $I$  due to line narrowing is larger than its decrease from the reduction in monomer concentration. The ratios of  $V$  and  $I$  ( $V/I$ ), however, decrease monotonically with decreasing monomer concentration at the initial stages of aggregation (Figure 2B, Lower Row) following a simple analytical relationship (see “Experimental Section”).

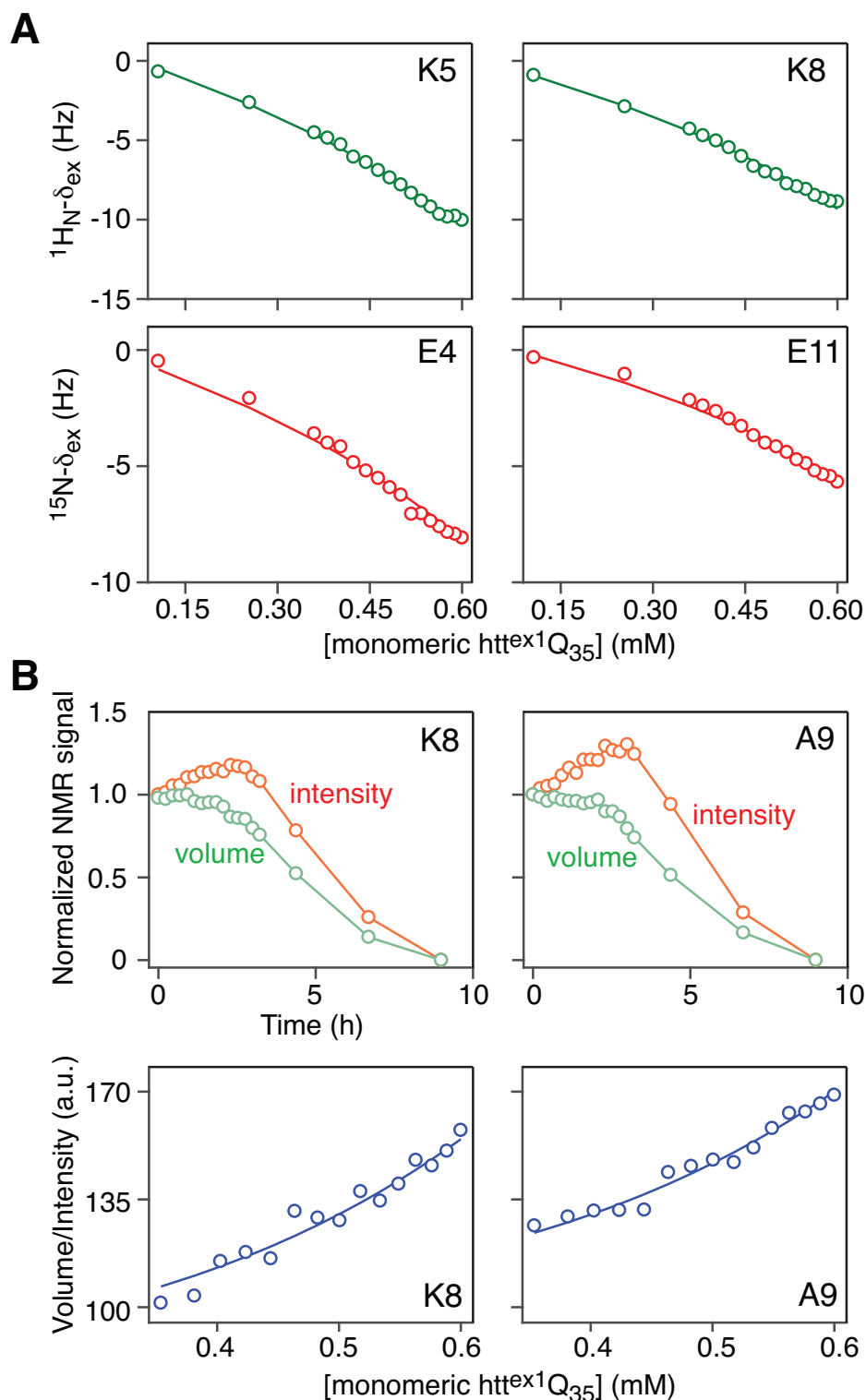
The optimized values for  $K_{\text{eq},1}$  and  $K_{\text{eq},2}$  obtained from the global fit are  $11.4 \pm 0.3$  and  $5.6 (\pm 0.5) \times 10^4 \text{ M}^{-1}$ , corresponding to dimer and tetramer dissociation constants of  $\sim 88$  and  $\sim 18 \mu\text{M}$ , respectively. The dissociation rate constants for the dimer ( $k_{-1}$ ) and tetramer ( $k_{-2}$ ) are  $\geq 3 \times 10^4$  and  $2.4 (\pm 0.5) \times 10^4 \text{ s}^{-1}$ , respectively (Figure S2, Supporting Information). At the highest concentration employed (600  $\mu\text{M}$ ), the fractional populations of dimer and tetramer in monomer units are  $\approx 1.3$  and  $\approx 0.6\%$ , respectively.

The temporal changes in the concentration of tetramers  $T$  in monomer units,  $4[T] = 4f[m(t)]$ , occurring during the course of aggregation for htt<sup>ex1</sup>Q<sub>35</sub> samples with total htt<sup>ex1</sup>Q<sub>35</sub> concentrations  $m_{\text{tot}}$  of 160, 250, and 420  $\mu\text{M}$ , are shown in Figure 3A, while the corresponding htt<sup>ex1</sup>Q<sub>35</sub> aggregation profiles best-fit to the model depicted in Figure 1 and described by Equations (1) and (2), are shown in Figure 3B. Throughout this work, the decrease in the NMR signal arising from NMR visible htt<sup>ex1</sup>Q<sub>35</sub> monomers as a function of time (Figure 3B) is attributed to the formation of any species (e.g., fibrils but also, if present, off-pathway large oligomers)<sup>[26]</sup> that are NMR invisible (unobservable) due to their very high molecular weight (see “Experimental Section”). The amount of tetrameric species available at the initial stages of fibril formation (early times in Figure 3A), defines the rate of primary nucleation (and, hence, the overall initial speed of aggregation) and is inversely proportional to the length of the lag period for the onset of aggregation (Figure 3B). At later times, the fibril-mediated secondary nucleation process ( $k_s$ ) “takes over” and becomes the dominant mechanism for nucleation, thus rendering the amount of tetramers and the rate of their conversion to “active” nuclei largely inconsequential. We note that in the case of htt<sup>ex1</sup>Q<sub>35</sub>, the mechanism of secondary nucleation is likely to involve interactions of the polyQ region of the monomers with the polyQ tracts on the surface of mature fibrils.

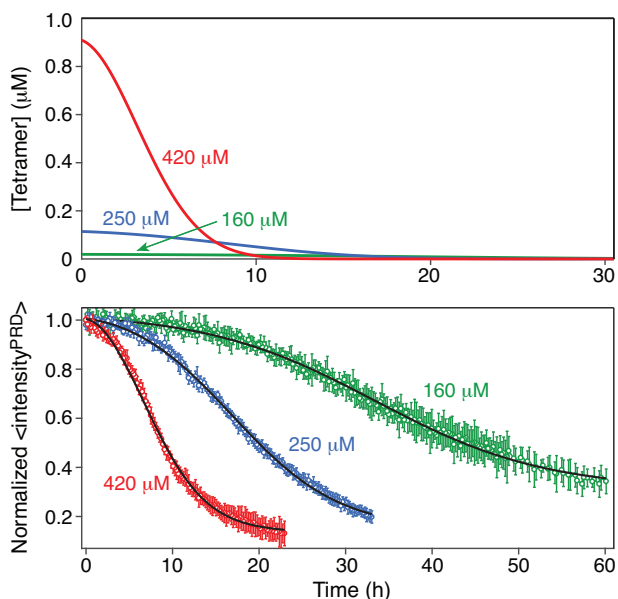
Numerical integration of Equations (1) and (2) in the best-fits of Figure 3B was performed using initial conditions,  $P(0) = 2, 4, \text{ and } 6 \text{ nM}$  for the 160, 250, and 420  $\mu\text{M}$  samples of htt<sup>ex1</sup>Q<sub>35</sub>, respectively (each determined via a separate grid search), while  $M(0)$  was set to 0 for all samples (the fits with “seedless” initial conditions,  $P(0) = 0$ , were of significantly inferior quality), yielding the following optimal values of the rate constants:  $k_c = 0.07 \pm 0.01 \text{ h}^{-1}$ ;  $k_s = 0.3 \pm 0.04 \text{ M}^{-1} \text{ h}^{-1}$ ; and  $k_+ = 6.4 (\pm 0.6) \times 10^5 \text{ M}^{-1} \text{ h}^{-1}$ . The presence of a small amount of nuclei ( $P(0) \neq 0$ ) ensures partial de-correlation<sup>[22,27]</sup> of otherwise highly correlated pairs of rate constants: ( $k_c$ ;  $k_+$ ) and ( $k_s$ ;  $k_+$ ) (see Supporting information). Plots of the time evolution of nuclei  $P$ , mature fibrils  $M$ , and the  $M/P$  ratio, which serves as an approximate measure of fibril length, are shown in Figure 4. The ratio  $M/P$  reaches a maximum of  $\approx 2000$  at  $t \rightarrow \infty$  in all three cases, which corresponds to a fibril limiting length of  $\sim 1 \mu\text{m}$ , assuming a 5 Å separation between polyQ  $\beta$  strands, in agreement with electron microscopy images in our earlier work.<sup>[17]</sup> Of note, the best-fits of htt<sup>ex1</sup>Q<sub>35</sub> aggregation profiles to approximate analytical solutions of Equations (1) and (2) yield very similar values for the rate constants and practically identical quality of fit (see Supporting Information for the closed-form solution of Equations (1) and (2)).

## 3. Discussion and Conclusion

In summary, we have described a kinetic model for aggregation of a pathogenic huntingtin construct with a 35-residue polyglutamine repeat, htt<sup>ex1</sup>Q<sub>35</sub>. The model establishes a clear-cut, quantitative connection between reversible tetramerization of htt<sup>ex1</sup>Q<sub>35</sub>, occurring on the microsecond timescale, and irreversible formation of fibrils that proceeds on a much slower timescale (from minutes to hours). The salient feature of this unified kinetic model involves a slow conversion step, in which the pre-equilibrated pool of htt<sup>ex1</sup>Q<sub>35</sub> tetramers is slowly converted with a finite, monomer-independent rate to “active”, elongation-



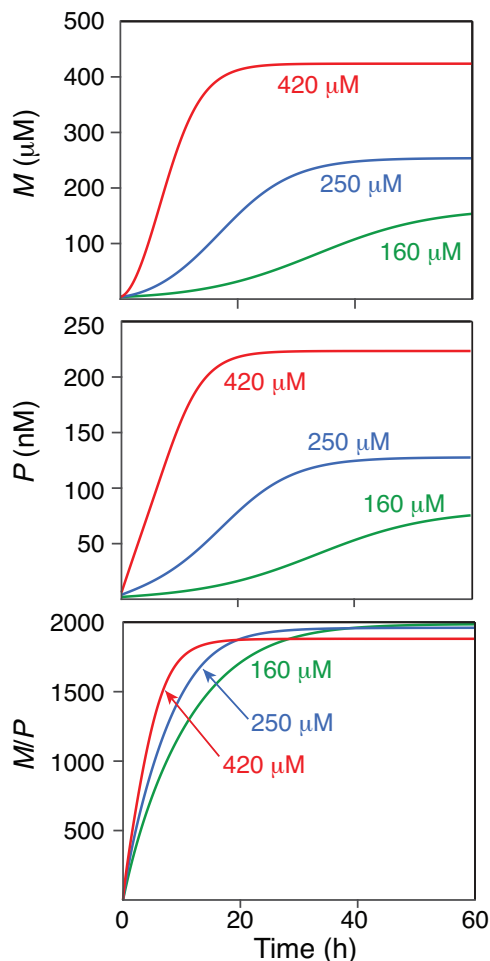
**Figure 2.** Pre-nucleation tetramerization on the microsecond timescale. A) Concentration-dependent  $^1\text{H-}\delta_{\text{ex}}$  (top) and  $^{15}\text{N-}\delta_{\text{ex}}$  (bottom) exchange-induced chemical shifts. B) Time course of  $^1\text{H-}^{15}\text{N}$  cross-peak volumes and intensities during aggregation (top) and  $^1\text{H-}^{15}\text{N}$  cross-peak volume/intensity (V/I) ratios (bottom) as a function of the concentration of monomeric htt<sup>ex1</sup>Q<sub>35</sub>. The total sample concentration was 600  $\mu\text{M}$ . The experimental data are shown as circles and the best fits to the unified kinetic scheme in Figure 1, under the assumption of equal changes in chemical shifts ( $\Delta\omega$ ) between monomers and dimers and monomers and tetramers,<sup>[14,17]</sup> are shown as continuous solid lines. All experiments were recorded at 5 °C and 800 MHz (see the Experimental Section) and the data for other residues used in the analysis are provided in Figure S1 (Supporting Information). The optimized values of  $\Delta\omega$  are given in Table S1 (Supporting Information).



**Figure 3.** Quantitative analysis of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation decay profiles. A) Simulated time-dependence of the concentration of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  tetramers  $T$  in monomer units for  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  samples with total monomer concentrations  $m_{\text{tot}}$  of 160, 250, and 420  $\mu\text{M}$  calculated using the values of  $K_{\text{eq1}}$  and  $K_{\text{eq2}}$  determined from the data in Figure 1. B) Time-dependence of the average  $^1\text{H}$ - $^{15}\text{N}$  cross-peak intensities for residues in the PRD domain of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$ . The experimental data (shown as circles) were recorded at 5 °C and 800 MHz and normalized to the first time point (at  $t \approx 0$  h). The best-fit curves are shown as black continuous lines and were obtained from a global fit to the kinetic scheme in Figure 1 and the model described by Equations (1) and (2) (see Experimental Section for details of the fitting procedure).

capable nuclei via a process of conformational re-arrangement of the polyQ tracts.

While there are some superficial similarities between the unified kinetic model of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation described here and the recently developed kinetic model for aggregation of amyloid  $\beta_{42}$ ,<sup>[28–29]</sup> which involves reversible formation of oligomers that are converted in a monomer-dependent manner to “active” nuclei, there are important differences between the two models that are specific to the system under consideration. First, the equilibrium  $m \rightleftharpoons D \rightleftharpoons T$  in Figure 1 is established “instantaneously” at each time-point sampled during the course of aggregation decay, with transitions between monomers and tetramers occurring on the  $\mu\text{s}$  time-scale – many orders of magnitude faster than the timescale of the nucleation and chain elongation processes. Second, only primary nucleation occurs via the conversion of tetramers in  $\text{htt}^{\text{ex1}}\text{Q}_{35}$ , while the process of secondary nucleation is modeled “classically” and likely involves interactions of the polyQ tracts of monomers with the same polyQ stretches ordered on the surface of mature fibrils. Third, the conversion stage is monomer-independent (the order of the “conversion” reaction with respect to monomer  $m$  is 0) implying a conformational transition within the tetramer to generate an elongation-capable species. The rate of conversion is determined by an unimolecular rate constant  $k_c$  that describes an inherently slow rate of “decay” of the “excited” states in the pool of sparsely-populated



**Figure 4.** Simulation of the time dependence of mature fibrils  $M$ , nuclei  $P$ , and the  $M/P$  ratios during the course of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  fibrillization at 5 °C using the optimized values of the rate constants ( $k_c = 0.07 \pm 0.01 \text{ h}^{-1}$ ;  $k_s = 0.3 \pm 0.04 \text{ M}^{-1}\text{h}^{-1}$ ; and  $k_+ = 6.4 (\pm 0.6) \times 10^3 \text{ M}^{-1}\text{h}^{-1}$ ) obtained from the global fits to the experimental data shown in Figures 2 and 3.

$\text{htt}^{\text{ex1}}\text{Q}_{35}$  tetramers to “aggregation-competent” (i.e., elongation-capable) “ground state” species.

The slow conformational transition of the pool of pre-nucleation tetramers to elongation-active nuclei presumably involves the polyQ domain that eventually comprises the fibril core. In the pre-nucleation tetramer, the polyQ domains are intrinsically disordered. The four-helix bundle formed by the amphiphilic NT domain within the tetramer serves as a template to bring the polyQ domains of a pair of parallel-oriented subunits into the appropriate spatial arrangement conducive to the formation of a four-stranded  $\beta$ -sheet comprising an intra-subunit anti-parallel  $\beta$  sheet/hairpin loop and an inter-subunit parallel  $\beta$ -sheet. This process is expected to be slow since a specific register between the polyQ strands is likely required to form elongation-active nuclei. One may speculate that this process may be in part facilitated by interactions between the hydrophobic moieties of the glutamine sidechains, not dissimilar to the “hydrophobic zipper” model of protein folding.<sup>[30–31]</sup>



We also note that the four-helix NT bundle in the pre-nucleation tetramer cannot be preserved in mature fibrils as these are characterized by a central rigid  $\beta$ -hairpin/ $\beta$ -sheet polyQ core with adjacent NT helices oriented parallel to one another.<sup>[7–13]</sup> Thus the initial four-helix NT bundle likely dissociates upon fibril elongation. However, the four-helix NT bundle motif can be reformed upon intermolecular collision of fibril chains to generate branched fibrils.

As the NMR data described here were obtained at 5 °C, it is of interest to discuss our expectations for changes in  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation kinetics at higher, physiologically relevant temperatures. Although we do not expect the kinetic model as such to undergo any alterations at higher temperatures, the aggregation rate constants, as well as the equilibrium constants for  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  tetramerization are expected to change significantly. As shown by thioflavin T (ThT) fluorescence assays in our earlier study of the interaction of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  with an SH3 domain,<sup>[32]</sup> aggregation is accelerated at higher temperatures: the  $t_{1/2}$  for fibril formation of 100  $\mu\text{M}$   $\text{htt}^{\text{ex1}}\text{Q}_{35}$  is  $\approx 2$  h as monitored by ThT fluorescence, compared to a  $t_{1/2}$  of  $\approx 30$  h for monomer disappearance at 5 °C as monitored by NMR. Thus, NMR-based kinetic studies of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  at 37 °C are not feasible. While the aggregation rate constants in Equations (1) and (2) (including that for tetramer conversion,  $k_c$ ), are expected to increase (possibly between 10- and 20-fold) as the temperature is raised from 5 to 37 °C, it is more difficult to predict the temperature dependence of the equilibrium constants  $K_{\text{eq},1}$  and  $K_{\text{eq},2}$  (Equation 3), as lower intrinsic stabilities of dimers and tetramers (due to entropy losses associated with oligomer formation) may be counteracted by stronger hydrophobic effects at higher temperatures.

To place the unified kinetic model of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation into a wider context, we note that the formation of non-amyloid oligomeric species, so-called “spherical oligomers”, has been reported in previous studies of  $\text{htt}^{\text{ex1}}$  aggregation.<sup>[6,33]</sup> In the framework of our model, such oligomers might be formed via the clustering of  $\alpha$ -helical tetramers into much larger aggregates that are nevertheless distinguishable from amyloid-like fibrils. Although the possibility that “spherical oligomers” of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  are indeed formed cannot be completely ruled out by kinetic modeling, the introduction of additional states/processes into the mechanism of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation was not required to fit the NMR data in the present case or our previous work.<sup>[17]</sup> The high-quality best-fits of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation profiles obtained with the minimalistic kinetic model (Figure 3B and Equations (1) and (2)) fully satisfy the experimental data, and are devoid of any systematic deviations that would serve as an indication of the presence of additional processes not included in the model. In contrast, in our recent NMR-based study of aggregation kinetics of A $\beta$ 42,<sup>[26]</sup> the introduction of an additional step for “off-pathway” oligomer formation to an otherwise similar kinetic model, was critically important to best fit the aggregation profiles quantitatively. We therefore conclude that further complication of the mechanism of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation (i.e., the incorporation of additional processes) is not warranted by our experimental data.

The unified quantitative kinetic model of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation described here bears a superficial similarity to the qualitative (cartoon-like) model of  $\text{htt}^{\text{ex1}}$  aggregation proposed by Wetzel and co-workers<sup>[6,33–34]</sup> (see, for example, Figure 3 of ref. [6]). The common feature of these models is the reversible formation of

$\text{htt}^{\text{ex1}}$  tetrameric species as the first step in the nucleation process. While in our model, the tetramers undergo an irreversible transition to form elongation-capable nuclei, the mechanism of nucleation posited by Wetzel includes an additional reversible step that involves clustering of tetramers into  $\alpha$ -helical higher-order oligomers (see discussion above) that are, in turn, reversibly converted to kinetic nuclei (elongation-capable species) with increased local concentrations of polyQ chains. In this regard, it is important to stress that the qualitative model proposed by Wetzel is not based on detailed global fitting of quantitative kinetic data at multiple  $\text{htt}^{\text{ex1}}$  concentrations, and our current results provide no evidence for the existence of higher-order (>4) oligomers that are subsequently converted into elongation-competent species.

Finally, it is worth noting that the analytical framework described in the paper will allow us to interpret in a physically meaningful, quantitative manner  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation profiles as a function of a variety of experimental conditions (for example, in the presence of crowding agents, macromolecular co-solutes, different salt concentrations, etc.) that we are currently studying in our laboratory. Note that different solution conditions can have an impact on any or all of the stages of the  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  aggregation mechanism. In the absence of a kinetic model that unifies all the involved processes, the monitored changes in aggregation would be largely uninterpretable.

## 4. Experimental Section

**Expression and Purification of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$ :**  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  was expressed as a fusion protein with the immunoglobulin-binding domain of streptococcal protein G, GB1, attached to the N-terminal end and separated by a factor Xa cleavage site as described previously.<sup>[17,32]</sup> Uniform  $^{15}\text{N}$ -labeling and cleavage by factor Xa to remove GB1 were carried out as described previously,<sup>[15,17]</sup> except that dialysis was performed against buffer containing 50 M Tris-HCl (pH 8.0), 100 mM NaCl and 1.5 mM dithiothreitol (DTT) using a 3.5 kDa cut-off dialysis membrane. The presence of significant amounts of DTT prevents the oxidation of the Met<sup>7</sup> side-chain to a sulfoxide. Following the last purification step,<sup>[17]</sup> the lyophilized  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  fractions were dissolved in a 1:1 (v/v) mixture of trifluoroacetic acid (TFA) and hexafluoroisopropanol (HFIP) ensuring complete removal of pre-existing aggregates.<sup>[17,35]</sup> Protein identity and completion of the cleavage reaction were verified by liquid phase chromatography coupled with electrospray ionization mass spectrometry.

**NMR Sample Preparation:** All NMR samples of  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  were prepared by first dissolving an aliquot of purified protein in a 13.8 mM monobasic sodium phosphate buffer, pH 4.6, containing 50 mM NaCl and 10% D<sub>2</sub>O/90% H<sub>2</sub>O (v/v). Upon re-suspension, the protein solution was centrifuged at 20 000 g for 25 min. to remove any pre-formed 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. Protein concentrations were determined by UV absorbance at 205 nm.<sup>[36]</sup> The samples were placed in 5 mm Shigemitsu tubes.

**NMR Spectroscopy:** All NMR experiments were recorded at 5 °C using a Bruker 800 MHz Avance-III spectrometer, equipped with a TCI triple resonance x,y,z-axis gradient cryogenic probe. NMR data were processed using the NMRPipe/NMRDraw software suite<sup>[37]</sup> and analyzed with the program Sparky.<sup>[38]</sup> Aggregation of  $^{15}\text{N}$ -labeled  $\text{htt}^{\text{ex1}}\text{Q}_{35}$  was monitored by following  $^1\text{H}$ - $^{15}\text{N}$  cross-peak intensities (heights) and volumes from a series of 2D  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra.<sup>[19–20]</sup> Each 2D NMR spectrum was acquired using a recycle delay of 0.3 s and a total of 100\* x 2048\* complex data points in the indirect ( $^{15}\text{N}$ ) and direct ( $^1\text{H}$ ) dimensions with respective acquisition times of 44 and 104 ms. The number of scans for each  $t_1$  increment was set to 16 resulting in total acquisition time of  $\approx 11$  min per 2D spectrum.

**Measurement of Concentration Dependent  $^{15}\text{N}/^1\text{H}_\text{N}$  Exchange-Induced Chemical Shifts ( $\delta_\text{ex}$ ):** Concentration dependent changes in  $^{15}\text{N}/^1\text{H}_\text{N}$  chemical shifts were measured on a 600  $\mu\text{M}$   $^{15}\text{N}$ -labeled sample of htt<sup>ex1</sup>Q<sub>35</sub> from a series of 2D  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra. The concentration of monomeric htt<sup>ex1</sup>Q<sub>35</sub> as a function of time,  $m(t)$ , was re-cast from the average normalized cross-peak intensity of 11 residues of the htt<sup>ex1</sup>Q<sub>35</sub> PRD domain ( $\langle I_\text{PRD} \rangle$ ) as,  $m(t) = m_\text{tot}[\langle I_\text{PRD}(t) \rangle - \langle I_\text{PRD}(\infty) \rangle] / [\langle I_\text{PRD}(0) \rangle - \langle I_\text{PRD}(\infty) \rangle]$ . For each time-point  $i$ ,  $^{15}\text{N}/^1\text{H}_\text{N}$ - $\delta_\text{ex}$  values were calculated as,  $\delta_\text{ex}(i) = \delta_\text{obs}(i) - \delta_\text{obs}^\text{ref}$ , where  $\delta_\text{obs}(i)$  is the observed chemical shift at the monomer concentration  $m(i)$ , and  $\delta_\text{obs}^\text{ref}$  is the chemical shift at the concentration close to the end of the aggregation decay ( $\approx 20 \mu\text{M}$ ).

**Global analysis of the concentration dependence of exchange-induced chemical shifts and  $V/I$  ratios together with the aggregation profiles of htt<sup>ex1</sup>Q<sub>35</sub>:** Combined analysis of  $^{15}\text{N}/^1\text{H}_\text{N}$ - $\delta_\text{ex}$ ,  $^1\text{H}^\text{N}$ - $^{15}\text{N}$  cross-peak volume/intensity ( $V/I$ ) ratios and aggregation profiles of htt<sup>ex1</sup>Q<sub>35</sub> followed the procedures described previously.<sup>[17]</sup> The experimental concentration dependence of  $^{15}\text{N}$ - $\delta_\text{ex}$ ,  $^1\text{H}_\text{N}$ - $\delta_\text{ex}$ ,  $^1\text{H}$ - $^{15}\text{N}$  cross peak  $V/I$  ratios, and decay profiles of averaged normalized NMR signal intensities,  $I$ , were fit simultaneously by the minimization of the error function,

$$F = \alpha_1 \sum_i \left( \frac{\delta_\text{N}^{\text{obs},i,k} - \delta_\text{N}^{\text{calc},i,k}}{\sigma_{\delta_\text{N}}^{i,k}} \right)^2 + \alpha_1 \sum_i \left( \frac{\delta_\text{H}^{\text{obs},i,k} - \delta_\text{H}^{\text{calc},i,k}}{\sigma_{\delta_\text{H}}^{i,k}} \right)^2 + \alpha_2 \sum_j \left( \frac{(V/I)^{\text{obs},j,k} - (V/I)^{\text{calc},j,k}}{\sigma_{(V/I)}^{j,k}} \right)^2 + \alpha_3 \sum_l \sum_m \left( \frac{I^{\text{obs},l,m} - I^{\text{calc},l,m}}{\sigma_I^{l,m}} \right)^2 \quad (4)$$

where the first and second terms correspond to the differences between the observed (“obs”) and calculated (“calc”)  $^{15}\text{N}$ - $\delta_\text{ex}$  and  $^1\text{H}_\text{N}$ - $\delta_\text{ex}$  values, respectively, for residue  $i$ , measured at each concentration  $k$  of the monomer; the third term accounts for the differences between observed and calculated  $V/I$  ratios for residue  $j$ , measured at each monomer concentration  $k$ ; and the last term describes the differences in observed and calculated averaged normalized NMR signal intensities,  $I$ , measured over all time-points  $l$  at the total protein concentration  $m$  for the 11 PRD residues of htt<sup>ex1</sup>Q<sub>35</sub>;  $\sigma$  are experimental errors (assumed to be approximately equal to one-half of the standard deviation of the distribution of peak intensities  $I$  in the aggregation profiles); and the coefficients  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  are empirically determined weighting factors.

The following form of the NMR Liouvillian for the three-state exchanging system,  $m \xrightleftharpoons[k_{-1}]{k_1^\text{app}} D \xrightleftharpoons[k_{-2}]{k_2^\text{app}} T$  was used for all calculations of  $^{15}\text{N}/^1\text{H}_\text{N}$ - $\delta_\text{ex}$  values and  $V/I$  ratios:  $\tilde{R} = \tilde{R}^\text{cs} + \tilde{R}^\text{r} + \tilde{R}^\text{ex}$ , where  $\tilde{R}^\text{cs}$ ,  $\tilde{R}^\text{r}$  and  $\tilde{R}^\text{ex}$  are the chemical shift, intrinsic transverse relaxation and chemical exchange contributions, respectively, given by:  $\tilde{R}^\text{cs} = \text{diag}[0; -i\Delta\omega_\text{n}; -i\Delta\omega_\text{n}]$  for single-quantum coherences (see Supporting Information of ref. [17] for the form of  $\tilde{R}^\text{cs}$  used for the zero- and double-quantum magnetization in calculations of  $V/I$  ratios), where  $\Delta\omega_\text{n}$  is the difference between the chemical shifts of the monomer and dimer/tetramer (assumed the same for dimers and tetramers; rad/s) for nucleus  $n$  ( $^{15}\text{N}$  or  $^1\text{H}_\text{N}$ );  $\tilde{R}^\text{r} = -\text{diag}[R_{2,n}; 2R_{2,n}; 4R_{2,n}]$ , where  $R_{2,n}$  is the transverse spin relaxation rate of the monomer for nucleus  $n$  in the absence of exchange ( $\text{s}^{-1}$ ); and the “linearized” form of exchange matrix  $\tilde{R}^\text{ex}$  is given by,

$$\tilde{R}^\text{ex} = - \begin{pmatrix} k_1^\text{app} & -k_{-1} & 0 \\ -k_1^\text{app} & k_2^\text{app} + k_{-1} & -k_{-2} \\ 0 & -k_2^\text{app} & k_{-2} \end{pmatrix} \quad (5)$$

where  $k_1^\text{app}$  and  $k_2^\text{app}$  are apparent, pseudo-first order rate constants given by,  $2k_1[m]$  and  $2k_1k_2[m]^2/k_{-1}$ , respectively.<sup>[17]</sup> The values of  $\delta_\text{ex}$  were calculated from the *imaginary* part of the smallest (by absolute magnitude) eigenvalue of the matrix  $\tilde{R}$ :  $\delta_\text{ex}(\text{Hz}) = \text{Im}(\text{min}[\text{eig}\{\tilde{R}\}]) / 2\pi$ , where  $\text{eig}\{\tilde{R}\}$  is a vector of complex eigenvalues of  $\tilde{R}$ . The ratios  $V/I$  were calculated from the approximate relationship described previously,<sup>[17]</sup> (see Supporting Information of ref. [17] for details) with the “effective” relaxation rates,

$R_{2,\text{eff}} = R_2 + R_\text{ex}$ , estimated from the *real* part of the smallest (by absolute magnitude) eigenvalue of the matrix  $\tilde{R}$ :  $R_{2,\text{eff}} = -\text{Re}(\text{min}[\text{eig}\{\tilde{R}\}])$ . Best-fit concentration-dependent amide backbone  $^{15}\text{N}/^1\text{H}_\text{N}$  exchange-induced shifts and  $V/I$  ratios for the residues of htt<sup>ex1</sup>Q<sub>35</sub> NT domain not included in Figure 2 of the main text are shown in Figure S1 (Supporting Information).

The  $\Delta\omega_\text{N}$  values were set to those previously determined for htt<sup>NT</sup>Q<sub>7</sub><sup>[14]</sup> on the assumption that the structure of the helical coiled-coil tetramer formed by the NT domain is not affected by the length of the polyQ tract, while the  $\Delta\omega_\text{H}$  values were optimized with the starting values reported in the previous study,<sup>[17]</sup> and using constrained minimization with the upper and lower bounds set to 2 times the uncertainties in the reported values (Table S1, Supporting Information). Intrinsic (exchange-free)  $^{15}\text{N}$ - $R_2$  and  $^1\text{H}_\text{N}$ - $R_2$  relaxation rates for htt<sup>ex1</sup>Q<sub>35</sub> monomer were taken from the previous study.<sup>[17]</sup> Since concentration-dependent  $^{15}\text{N}$ - $R_1\rho$  measurements are not feasible on a rapidly aggregating system,<sup>[16]</sup> the approximate range of  $k_{-1}$  values was established using a grid search (Figure S2, Supporting Information), and subsequently fixed at a value of  $30\,000 \text{ s}^{-1}$ .

It is worth noting that if one uses the approximation for  $^{15}\text{N}$ - $\delta_\text{ex}$  valid in the fast-exchange regime on the chemical shift time scale and given by,  $\delta_\text{ex} \approx (p_\text{D} + p_\text{T})\Delta\omega = [2K_\text{eq,1}m + 4(K_\text{eq,1})^2K_\text{eq,2}m^3]\Delta\omega$ , both  $K_\text{eq,1}$  and  $K_\text{eq,2}$  are underestimated by only  $\sim 2.5\%$ ; using the same approximation for both  $^{15}\text{N}$ - $\delta_\text{ex}$  and  $^1\text{H}_\text{N}$ - $\delta_\text{ex}$ , however, results in overestimation of  $K_\text{eq,1}$  and underestimation of  $K_\text{eq,2}$  by  $\sim 2.5\%$  and  $\sim 20\%$ , respectively.

Temporal changes in NMR signal intensities  $I$  monitored in aggregation profiles of htt<sup>ex1</sup>Q<sub>35</sub> were calculated from the expression,

$$I_\text{calc} = A_0 [1 - \alpha' M(t)/m_\text{tot}] \quad (6)$$

where  $\alpha'$  is a scaling factor that takes into account that NMR intensities for the PRD domain do not decay to exactly zero for fully aggregated samples (as the PRD remains flexible and disordered in the fibrils);  $A_0$  is an overall scaling factor that accounts for possible errors in normalization of NMR intensities; and  $M(t)$  is the mass concentration of mature fibrils derived from numerical integration of Equations (1) and (2). Numerical integration of Equations (1) and (2) was performed using an explicit Runge-Kutta (1,4) formula implemented in Matlab “ode45” solver for non-stiff ordinary differential equations,<sup>[39]</sup> with integration steps corresponding to the time-points sampled in the NMR experiments.

The set of global variable parameters in the minimization of the target function thus comprised:  $\{K_\text{eq,1}; K_\text{eq,2}; k_2; k_c; k_s; k_+\}$ . Of note, although a single (combined) target function (Equation 4) was used in the global analysis of the experimental data, the first three terms in Equation (4) drive the determination of  $K_\text{eq,1}$  and  $K_\text{eq,2}$ , since the aggregation profiles are tolerant to the “supplied” values of  $K_\text{eq,1}$  and  $K_\text{eq,2}$  by virtue of the dependence of the rate of production of nuclei,  $dP/dt$  in Equation (1), on the product of the tetramer concentration,  $f[m(t)]$ , and the conversion rate constant  $k_c$  in the model. For  $k_{-1}$  set to  $3 \times 10^4 \text{ s}^{-1}$ ,  $k_2$  optimizes to  $2.4 (\pm 0.5) \times 10^4 \text{ s}^{-1}$ .

The uncertainties in the values of the optimized parameters, corresponding to confidence intervals of  $\pm 1$  standard deviation, 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 error propagation. All calculations were performed using an in-house program written in MATLAB (Math-Works Inc, MA).

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

The authors thank Dr. Atilla Szabo (NIDDK, NIH) for useful suggestions, and 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 National Institutes of Health (DK29023 to G.M.C.).

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

## Keywords

elongation competent nuclei, huntingtin exon-1 protein, NMR, pre-nucleation tetramers, unified kinetic model of aggregation

Received: November 29, 2023

Revised: January 18, 2024

Published online: March 12, 2024

- 
- [1] G. P. Bates, R. Dorsey, J. F. Gusella, M. R. Hayden, C. Kay, B. R. Leavitt, M. Nance, C. A. Ross, R. I. Scahill, R. Wetzel, E. J. Wild, S. J. Tabrizi, *Nat. Rev. Dis. Primers* **2015**, *1*, 15005.
- [2] J. M. Andresen, J. Gayán, L. Djoussé, S. Roberts, D. Brocklebank, S. S. Cherny, *Ann. Hum. Genet.* **2007**, *71*, 295.
- [3] C. A. Ross, S. J. Tabrizi, *Lancet Neurol.* **2011**, *10*, 83.
- [4] C. Zuccato, M. Valenza, E. Cattaneo, *Physiol. Rev.* **2010**, *90*, 905.
- [5] K. Kar, M. Jayaraman, B. Sahoo, R. Kodali, R. Wetzel, *Nat. Struct. Mol. Biol.* **2011**, *18*, 328.
- [6] R. Wetzel, *Acc. Chem. Res.* **2020**, *53*, 2347.
- [7] C. L. Hoop, H. K. Lin, K. Kar, G. Magyarfalvi, J. M. Lamley, J. C. Boatz, A. Mandal, J. R. Lewandowski, R. Wetzel, P. C. van der Wel, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1546.
- [8] J. C. Boatz, T. Piretra, A. Lasorsa, I. Matlahov, J. F. Conway, P. C. A. van der Wel, *J. Mol. Biol.* **2020**, *432*, 4722.
- [9] V. N. Sivanandam, M. Jayaraman, C. L. Hoop, R. Kodali, R. Wetzel, P. C. van der Wel, *J. Am. Chem. Soc.* **2011**, *133*, 4558.
- [10] J. M. Isas, R. Langen, A. B. Siemer, *Biochemistry* **2015**, *54*, 3942.
- [11] H. K. Lin, J. C. Boatz, I. E. Krabbendam, R. Kodali, Z. Hou, R. Wetzel, A. M. Dolga, M. A. Poirier, P. C. A. van der Wel, *Nat. Commun.* **2017**, *8*, 15462.
- [12] F. J. B. Bauerlein, I. Saha, A. Mishra, M. Kalemánov, A. Martínez-Sánchez, R. Klein, I. Dudanova, M. S. Hipp, F. U. Hartl, W. Baumeister, R. Fernández-Busnadiego, *Cell* **2017**, *171*, 179.
- [13] J. G. Galaz-Montoya, S. H. Shahmoradian, K. Shen, J. Frydman, W. Chiu, *Commun. Biol.* **2021**, *4*, 849.
- [14] S. A. Kotler, V. Tugarinov, T. Schmidt, A. Ceccon, D. S. Libich, R. Ghirlando, C. D. Schwieters, G. M. Clore, *Proc. Natl. Acad. Sci.* **2019**, *116*, 3562.
- [15] A. Ceccon, V. Tugarinov, R. Ghirlando, G. M. Clore, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 5844.
- [16] A. Ceccon, V. Tugarinov, G. M. Clore, *J. Phys. Chem. Lett.* **2020**, *11*, 5643.
- [17] A. Ceccon, V. Tugarinov, F. Torricella, G. M. Clore, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2207690119.
- [18] A. Ceccon, V. Tugarinov, G. M. Clore, *J. Am. Chem. Soc.* **2019**, *141*, 94.
- [19] P. Schanda, B. Brutscher, *J. Am. Chem. Soc.* **2005**, *127*, 8014.
- [20] P. Schanda, E. Kupce, B. Brutscher, *J. Biomol. NMR* **2005**, *33*, 199.
- [21] F. Oosawa, S. Asakura, in *Thermodynamics of the Polymerization of Proteins*, Academic Press, Cambridge, **1975**.
- [22] G. Meisl, J. B. Kirkegaard, P. Arosio, T. C. Michaels, M. Vendruscolo, C. M. Dobson, S. Linse, T. P. Knowles, *Nat. Protoc.* **2016**, *11*, 252.
- [23] J. Hofrichter, P. D. Ross, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4864.
- [24] F. A. Ferrone, J. Hofrichter, W. A. Eaton, *J. Mol. Biol.* **1985**, *183*, 611.
- [25] F. A. Ferrone, J. Hofrichter, H. R. Sunshine, W. A. Eaton, *Biophys. J.* **1980**, *32*, 361.
- [26] S. Ghosh, V. Tugarinov, G. M. Clore, *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2305823120.
- [27] S. I. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, T. P. Knowles, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9758.
- [28] T. C. T. Michaels, A. Saric, S. Curk, K. Bernfur, P. Arosio, G. Meisl, A. J. Dear, S. I. A. Cohen, C. M. Dobson, M. Vendruscolo, S. Linse, T. P. J. Knowles, *Nat. Chem.* **2020**, *12*, 445.
- [29] A. J. Dear, T. C. T. Michaels, G. Meisl, D. Klenerman, S. Wu, S. Perrett, S. Linse, C. M. Dobson, T. P. J. Knowles, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 12087.
- [30] K. A. Dill, K. M. Fiebig, H. S. Chan, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1942.
- [31] A. M. Gronenborn, G. M. Clore, *Science* **1994**, *263*, 536.
- [32] A. Ceccon, V. Tugarinov, G. M. Clore, *J. Am. Chem. Soc.* **2021**, *143*, 9672.
- [33] B. Sahoo, D. Singer, R. Kodali, T. Zuchner, R. Wetzel, *Biochemistry* **2014**, *53*, 3897.
- [34] B. Sahoo, I. Arduini, K. W. Drombowski, R. Kodali, L. H. Sanders, J. T. Greenamyre, R. Wetzel, *PLoS One* **2016**, *11*, 0155747.
- [35] S. Chen, R. Wetzel, *Protein Sci.* **2001**, *10*, 887.
- [36] N. J. Anthis, G. M. Clore, *Protein Sci.* **2013**, *22*, 851.
- [37] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277.
- [38] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325.
- [39] L. F. Shampine, M. W. Reichelt, *SIAM J. Sci. Comput.* **1997**, *18*, 1.