2018

Profilin reduces aggregation and phase separation of huntingtin N-terminal fragments by preferentially binding to soluble monomers and oligomers

Ammon E. Posey
Washington University School of Medicine in St. Louis

Kiersten M. Ruff
Washington University School of Medicine in St. Louis

Tyler S. Harmon
Washington University School of Medicine in St. Louis

Scott L. Crick
Washington University School of Medicine in St. Louis

Aimin Li
Washington University School of Medicine in St. Louis

See next page for additional authors

Recommended Citation

Posey, Ammon E.; Ruff, Kiersten M.; Harmon, Tyler S.; Crick, Scott L.; Li, Aimin; and Diamond, Marc I., "Profilin reduces aggregation and phase separation of huntingtin N-terminal fragments by preferentially binding to soluble monomers and oligomers." Journal of Biological Chemistry. 293, 10. 3734-3746. (2018).
https://digitalcommons.wustl.edu/open_access_pubs/6663

Profilin reduces aggregation and phase separation of huntingtin N-terminal fragments by preferentially binding to soluble monomers and oligomers

Received for publication, October 12, 2017, and in revised form, January 16, 2018 Published, Papers in Press, January 22, 2018, DOI 10.1074/jbc.RA117.000357

Ammon E. Posey[‡], Kiersten M. Ruff[‡], Tyler S. Harmon[‡], Scott L. Crick[‡], Aimin Li[§], Marc I. Diamond[¶], and Rohit V. Pappu[‡1]

From the [‡]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University, St. Louis, Missouri 63130, the [¶]Center for Alzheimer's and Neurodegenerative Diseases, University of Texas Southwestern Medical Center, Dallas, Texas 75390, and the [§]Department of Neurology, Washington University School of Medicine, St. Louis, Missouri 63108

Edited by Karen G. Fleming

Huntingtin N-terminal fragments (Htt-NTFs) with expanded polyglutamine tracts form a range of neurotoxic aggregates that are associated with Huntington's disease. Here, we show that aggregation of Htt-NTFs, irrespective of polyglutamine length, yields at least three phases (designated M, S, and F) that are delineated by sharp concentration thresholds and distinct aggregate sizes and morphologies. We found that monomers and oligomers make up the soluble M phase, ~25-nm spheres dominate in the soluble S phase, and long, linear fibrils make up the insoluble F phase. Previous studies showed that profilin, an abundant cellular protein, reduces Htt-NTF aggregation and toxicity in cells. We confirm that profilin achieves its cellular effects through direct binding to the C-terminal proline-rich region of Htt-NTFs. We show that profilin preferentially binds to Htt-NTF M-phase species and destabilizes aggregation and phase separation by shifting the concentration boundaries for phase separation to higher values through a process known as polyphasic linkage. Our experiments, aided by coarse-grained computer simulations and theoretical analysis, suggest that preferential binding of profilin to the M-phase species of Htt-NTFs is enhanced through a combination of specific interactions between profilin and polyproline segments and auxiliary interactions between profilin and polyglutamine tracts. Polyphasic linkage may be a general strategy that cells utilize to regulate phase behavior of aggregation-prone proteins. Accordingly, detailed knowledge of phase behavior and an understanding of how ligands modulate phase boundaries may pave the way for developing new therapeutics against a variety of aggregation-prone proteins.

Many diseases are associated with protein misfolding and aggregation (1, 2). The aggregation process is often characterized by the presence of one or more threshold concentrations at which a sharp, discontinuous change to some aspect of the assembly state (*e.g.* size, conformational characteristics, material properties) occurs (3–6). Such a change can be described using the concepts of phase transitions. Phase separation, a subcategory of phase transitions, has recently received considerable attention due to increasing recognition of its importance in cell biology (7–13). Phase separation refers to aggregation-related changes in molecular density that give rise to the coexistence of dilute macromolecule-deficient phases and dense macromolecule-rich phases (3, 14, 15). Examples of multiple coexisting phases have been observed in biological contexts (15–18), and these phases can be liquid, solid, or semisolid (*e.g.* a gel) (10, 19–25). Driving forces for phase separation are quantified in terms of saturation concentrations (14, 19, 26). For a given two-phase system, the saturation concentration is the bulk concentration of the protein, beyond which the solution separates into two coexisting phases. The lower the saturation concentration, the stronger the driving force for aggregation and phase separation (14, 19, 26–29).

Ligand binding can alter saturation concentrations and shift phase boundaries to either stabilize or destabilize specific phases (30). Wyman and Gill (31) introduced this concept, known as polyphasic linkage, to interpret the impact of ligands on saturation concentrations. In the current work, we report results from our studies of ligand binding to N-terminal fragments of the huntingtin protein (Htt-NTFs).[2]

Huntington's disease (HD) is a devastating neurological disorder that affects medium spiny and cerebellar neurons (32). The age of onset as well as severity of HD are inversely correlated with the length of the polyglutamine (polyQ) tract within the N-terminal region of huntingtin (Htt) (32). The disease is characterized by deposition of insoluble intranuclear deposits of Htt-NTFs with expanded polyQ tracts in afflicted neurons (33, 34). Htt-NTFs consist of a 17-residue N-terminal stretch (MATLEKLMKAFESLKSF) designated as N17, a polyQ tract of length n (Q_n), and a 38-residue C-terminal stretch (C38) that includes two polyproline (polyP) modules, P_{11} and P_{10}, connected by a 17-residue linker denoted as L17 (QLPQPPPQAQ-

This work was supported by National Institutes of Health Grants R01NS056114 (to R. V. P.) and R01NS089932 (to M. I. D. and R. V. P.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains Table S1, Figs. S1–S7, and Methods.

[1] To whom correspondence should be addressed. E-mail: pappu@wustl.edu.

[2] The abbreviations used are: Htt-NTF, huntingtin N-terminal fragment; HD, Huntington's disease; polyQ, polyglutamine; polyP, polyproline; Q_n, polyQ tract of length n; CFP, cyan fluorescent protein; YFP, yellow fluorescent protein; ThT, thioflavin T; TEM, transmission electron microscopy; ANOVA, analysis of variance.

PLLPQPQ) (Fig. 1*a*). In transgenic mouse models, Htt-NTFs produce robust degeneration of the relevant neurons (35).

Although HD is characterized by Htt-NTF inclusions, the role of these inclusions as agents of neurotoxicity has been the topic of considerable debate (34–44). In a recent study, Ramdzan *et al.* (45) showed that both soluble species and insoluble deposits engender neurotoxicity, albeit through very different mechanisms. Soluble forms of Htt-NTFs promote cell death via apoptosis due to mitochondrial dysfunction. In contrast, insoluble inclusions weaken apoptosis and induce slow cell death via necrosis by impacting cellular metabolism and promoting co-aggregation and phase separation of other proteins with Htt-NTFs (45). In light of this recent study, it is reasonable to postulate that the entire aggregation process and the full spectrum of Htt-NTF aggregates might have deleterious roles to play in neurons. This postulate argues against the notion that a single species of aggregate is toxic and suggests that attempts to target a single toxic species are likely to be unfruitful. It may be more advantageous to instead consider toxicity in terms of supersaturation with respect to a phase boundary (or boundaries) and the alleviation of toxicity in terms of ligands that effectively reduce supersaturation by shifting phase boundaries via thermodynamic linkage.

A ligand of interest that has been shown to reduce toxicity of Htt-NTFs in multiple HD models is the ubiquitously expressed protein profilin-1, hereafter referred to as profilin for brevity. Profilin is a 15-kDa protein that binds to G-actin, polyP, and phosphatidylinositol lipids through three distinct binding sites (46–49) and is essential for cell division and survival during embryogenesis (50). Shao *et al.* (51) showed that profilin suppresses Htt-NTF aggregation through direct interactions with Htt-NTFs. Increased expression of profilin reduces intracellular aggregation of Htt-NTFs and diminishes the toxicity of Htt-NTFs in cell lines, primary neurons, model organisms, and transgenic mice (52). Therefore, a direct therapeutic approach to HD could involve the design of ligands that mimic the effects of profilin on Htt-NTF aggregation and alleviation of toxicity. To realize such a therapeutic, we need a mechanistic understanding of how profilin suppresses Htt-NTF aggregation. This is the focus of the current work.

Here, we show that Htt-NTFs form at least three distinct phases that are preferred across distinct concentration regimes, and we identify the concentration thresholds that demarcate these concentration regimes. By quantifying the driving forces for aggregation and phase separation in the presence and absence of profilin, we show that the binding of profilin to the proline-rich C38 region of Htt-NTFs decreases the driving forces for aggregation and phase separation through polyphasic linkage. Specifically, systematic measurements of profilin binding aided by coarse-grained simulations reveal that profilin preferentially binds low-molecular weight species (monomers/oligomers). This leads to the stabilization of monomers and oligomers and a suppression of large-scale aggregation and phase separation. In addition to direct interactions between profilin and polyP, our analysis points to hitherto unknown auxiliary interactions between profilin and the polyQ domain that engender an enhancement in profilin binding to Htt-NTF monomers/oligomers.

Results

C38 is necessary for profilin to bind to Htt-NTFs

Shao *et al.* (51) previously identified unphosphorylated profilin as an intracellular ligand that reduces Htt-NTF aggregation and toxicity. In that study, a direct interaction was inferred from pulldowns between non-mutant Htt-NTF (Q_{25}) and either wildtype profilin or profilin variants with a mutated polyP binding pocket. The results indicated that the polyP binding pocket of profilin is essential for binding Htt-NTFs. We carried out follow-up experiments to assess, directly, whether the Htt-NTF C38 stretch is essential to elicit profilin-dependent reduction in aggregation. We used an intracellular aggregation assay (51, 53–55) to quantify the impact of profilin overexpression on the aggregation of constructs with and without the C38 stretch. Htt-NTFs with expanded polyQ tracts (Q_{72}) were fused to cyan (CFP) or yellow fluorescent proteins (YFP). CFP and YFP are used as FRET pairs, and when fused to Htt-NTFs they serve as intracellular reporters of aggregation. We expressed N17-Q_{72}-C38-CFP/YFP or N17-Q_{72}-CFP/YFP in HEK293 cells cultured in a 96-well format (51), with or without profilin. We incubated the cells for 48 h before fixing the plates in 2% paraformaldehyde and measuring intracellular, intermolecular FRET in a fluorescence plate reader. As described previously (51), wildtype profilin reduced N17-Q_{72}-C38-CFP/YFP aggregation by 30%. However, when the polyP-containing C38 stretch was removed, the aggregation of N17-Q_{72}-CFP/YFP was unaffected by profilin overexpression (see *first* and *second green bars* in Fig. S1). The ability to suppress intracellular aggregation by profilin was blocked by the phosphomimic mutation S137D and was preserved by the S137A mutation (*second* and *third gray bars*, respectively, in Fig. S1), whereas these mutations had no effects on constructs lacking C38 (*green bars* in Fig. S1). These data confirm that direct interactions between profilin and polyP modules in C38 play a central role in the modulation of Htt-NTF aggregation via profilin binding.

Constructs for in vitro experiments

Recent *in vivo* studies from Yang and co-workers (56, 57) showed that a ΔN17 variant of huntingtin, which lacks the N17 module, leads to faithful reproduction of HD phenotypes in transgenic mice. In light of these observations, and given the relative ease of synthesis and purification of Htt-NTFs that lack the N17 module (19), we used constructs of the form Q_n-C38 for a majority of our *in vitro* experiments. Here, n refers to the length of the polyQ tract. For most of the experiments, a polyQ length of $n = 40$ was used for the feasibility of synthesis and handling while still being long enough to be of a mutant or pathological length (58). In the interest of completeness, we also used a set of distinct Htt-NTF sequences for some of the experiments. Fig. 1*a* shows a detailed inventory of all of the sequence constructs used in one or more of our studies.

Htt-NTFs form at least three distinct phases

Previous studies (19) showed that Htt-NTFs separate into soluble and insoluble species at bulk concentrations that are above construct-specific saturation concentrations. Accordingly, one set of saturation concentrations can be defined in

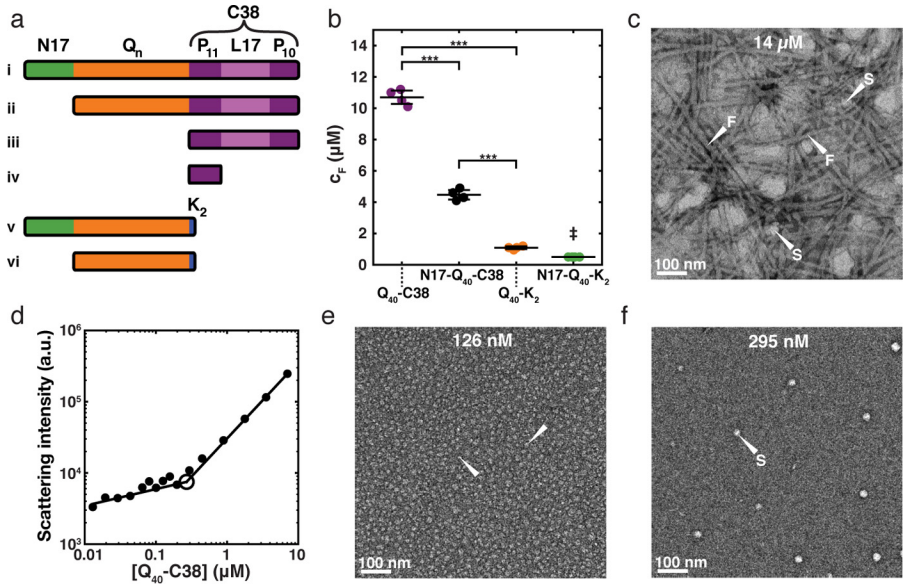

Figure 1. Experimental evidence for Htt-NTF phase behavior with multiple phase boundaries. *a*, Htt-NTF sequence architectures used in this study. *i*, full Htt exon 1, N17-Q_n-C38; *ii*, Q_n-C38; *iii*, C38; *iv*, P_{11}; *v*, N17-Q_n-K_2; *vi*, Q_n-K_2. Here, $n = 40$ in all experiments except for the in-cell FRET aggregation assay (see Fig. S1), where $n = 72$, and some experiments where $n = 30$ was compared with $n = 40$ (see Fig. 5*e*). Additionally, K_2 corresponds to two lysines. *b*, previously reported (19) saturation concentrations (denoted here as c_F) determined by measuring the concentration of soluble protein remaining in the supernatant (colorimetric micro-BCA assay) following quiescent incubation at 30 °C and centrifugation of the indicated Htt-NTF constructs. Consistent with the presence of a saturation concentration, the same concentration was arrived at for a given construct regardless of the starting concentration, so long as the starting concentration was above the saturation concentration indicated. The saturation concentration is modulated by the N17 and C38 sequence modules that flank polyQ in Htt-NTFs. *Black bars*, mean ± S.D. (*error bars*) of four independent experiments. The N17-Q_{40}-K_2 c_F was below the detection limit of the assay for all trials, so the quantity plotted is the detection limit and represents an upper bound for the N17-Q_{40}-K_2 c_F. Statistical significance was determined by one-way ANOVA with Tukey's range test for post hoc analysis. ***, $p < 0.001$. ‡, the N17-Q_{40}-K_2 c_F data were not included in the statistical significance analysis for the reason cited above. *c*, TEM image of Q_{40}-C38 fibrils observed at a concentration supersaturated with respect to c_F (14 μM). Two of the fibrils in the image are indicated with *white arrows labeled F*, and two examples of spherical aggregates are indicated with *white arrows labeled S*. See Fig. 4*c* for an additional example of fibrils and spheres. *d*, right-angle static light scattering of Q_{40}-C38 in 20 mM Tris, 5 mM EDTA, 1 mM DTT, pH 7.4, as a function of Q_{40}-C38 concentration. The discontinuity at ~290 nM is indicative of a phase boundary. *e*, TEM image of Q_{40}-C38 monomers and small oligomers observed at a concentration subsaturated with respect to c_S (126 nM). This particular field of view contains monomers/small oligomers, two of which are indicated with *white arrows*. *f*, TEM image of Q_{40}-C38 spherical aggregates observed at a concentration supersaturated with respect to c_S but subsaturated with respect to c_F (295 nM). One of the spherical aggregates is indicated with a *white arrow labeled S*. All TEM images (*c*, *e*, and *f*) are at the same scale. *Scale bar*, 100 nm.

terms of the solubility limit, which we shall designate as c_F for fibrillar-phase saturation concentration. It is important to clarify that c_F was designated as c_S in the work of Crick *et al.* (19). For bulk concentrations that are greater than c_F, the concentration of soluble protein in the supernatant will be equal to c_F, provided that equilibrium has been established between the soluble and insoluble phases. The value of c_F is quantifiable using previously published colorimetric assays (19). In this approach, solutions with varying concentrations of fully disaggregated Htt-NTFs (see "Experimental procedures" and supporting Methods), were equilibrated via incubation in a temperature-controlled water bath without shaking for up to 2 weeks. The solutions were separated, via centrifugation, into soluble supernatants that coexist with insoluble pellet fractions. For bulk concentrations above a construct-specific threshold value of c_F, the concentration of protein remaining in solution was found to be constant and independent of starting concentration. This is a prerequisite for the existence of a saturation concentration (59). Values of c_F for different Htt-NTF constructs with polyQ lengths of $n = 40$ at 37 °C are shown in Fig. 1*b*. The values of c_F varied with Htt-NTF constructs, where each construct either has a different sequence module that flanks the polyQ segment or, for a given construct, the polyQ length is different (19).

At concentrations above c_F, the insoluble material in the pellet stains positively with amyloid-sensitive dyes, such as thioflavin T (ThT) (19). Transmission electron microscopy (TEM) images confirmed that the insoluble phase is predominantly composed of fibrillar aggregates (Fig. 1*c*). We also observed coexisting small spherical aggregates in TEM images. These observations suggest that spherical aggregates might represent either a separate stable phase that is in equilibrium with the fibrils and the unaggregated protein in the bulk solution, or they are metastable phases that form as precursors to fibrils. Given their destabilization with increased supersaturation with respect to c_F, we investigated the possibility that spherical aggregates form as a distinct phase with saturation concentrations that are below c_F. On average, the spherical species appeared to be too small for sedimentation by centrifugation. Therefore, we turned to right-angle static light scattering to investigate the presence of a second phase boundary at concentrations that were subsaturated with respect to c_F (*i.e.* for concentrations less than c_F).

The average size and the concentration of scattering species determine the intensity of scattered light in static light-scattering experiments. The appearance of a sharp discontinuity in scattering intensity *versus* concentration is indicative of the presence of a saturation concentration in a phase-separating system (60). To test for the possible presence of a distinct saturation concentration that lies below c_F, we measured right-angle static light scattering of Htt-NTF solutions as a function of peptide concentration (Fig. 1*d*). We observed a discontinuity in

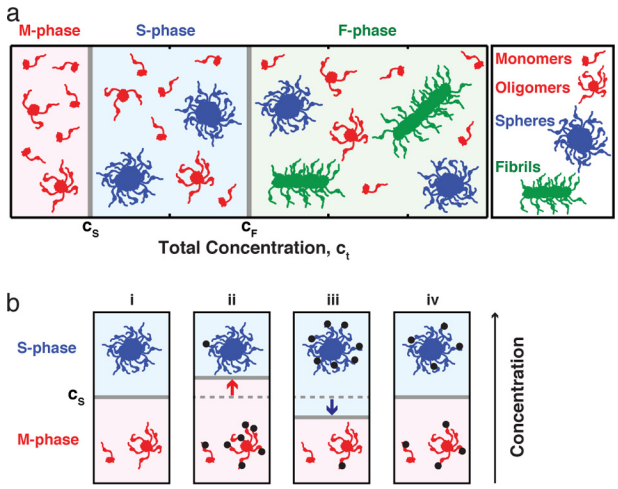

Figure 2. Schematic representation of Htt-NTF phase behavior and polyphasic linkage. *a*, Htt-NTF phase behavior is described by three phases (M, S, and F phase), *color-coded* as *red*, *blue*, and *green*, respectively, and two phase boundaries (c_S and c_F) indicated by *gray lines* between the phases. The total monomer-equivalent Htt-NTF concentration (*horizontal axis*) with respect to these boundaries will determine whether monomers and oligomers, spherical aggregates, or fibrils form. *b*, the influence of a binding partner (profilin) on the Htt-NTF c_S phase boundary. The phase boundary is depicted as a *solid gray line* in the absence of profilin (*i*) and for three different profilin binding scenarios (*ii–iv*). The relative number of profilin (*black circles*) associated with each phase (M or S) indicates which phase is preferentially bound. The *dotted gray line* in *ii* and *iii* indicates the c_S in the absence of profilin, and the *red* and *blue arrows* indicate the direction of the shift in the c_S if the M or S phase is preferentially bound, respectively. In scenario *iv*, profilin binds both phases equally well, and thus no shift in c_S is observed. The orientation of *b* is rotated 90° with respect to *a* for formatting purposes. See "Results" for further description of polyphasic linkage.

the concentration dependence of the light-scattering intensity at ~290 nM for Q_{40}-C38. This is consistent with the presence of a phase boundary or saturation concentration, which we term c_S for spherical-phase saturation concentration. TEM images indicate that species considerably smaller than the 10–20-nm size, which presumably include monomers and oligomers, are the predominant morphological species at concentrations below c_S (Fig. 1*e*). Conversely, spherical aggregates with a diameter of ~25 nm are the preferred morphologies for concentrations between c_S and c_F (Fig. 1*f*). We analyzed multiple independent TEM images to quantify the distribution of aggregate sizes. This analysis confirmed that a single size distribution was present at concentrations below c_S, whereas two distinct size distributions were present at concentrations that lie between c_S and c_F (Fig. S2).

Our results, derived from a joint analysis of solubility, light scattering, and morphological measurements, lead to a phase diagram that delineates the presence of at least three distinct phases separated by two phase boundaries (Fig. 2*a*). For total concentrations c_t that are lower than c_S, Htt-NTFs populate monomers and oligomers (M phase). For concentrations between c_S and c_F, the concentration of monomers/oligomers saturates, and these species are in equilibrium with large spherical aggregates (S phase). Soluble, spherical aggregates that are ~25 nm in diameter are the predominant morphologies in the S phase. For $c_t > c_F$, the concentration of soluble species saturates, and equilibrium is established between soluble species and insoluble, large, β-sheet–rich fibrils with a bottlebrush architecture that constitute the F phase (61). We use the term "aggregates" for molecules engaged in intermolecular interactions that can range from dimers to fibrils, whereas "oligomers" refers to aggregates smaller than the spherical aggregates that characterize the S phase.

Profilin preferentially binds M-phase species

The quantitative phase diagrams for Htt-NTFs become useful when considering the impact of binding partners on phase behavior. Polyphasic linkage provides a framework for understanding how ligand binding modulates the phase equilibria of molecules that undergo aggregation and phase separation (30, 31) (Fig. 2*b*). Polyphasic linkage can be explained by considering the simple example of a macromolecular solution that separates into two distinct phases labeled as A and B, respectively. Here, A could be the dilute phase and B the dense phase. In the absence of ligand, we denote c_A as the saturation concentration of the macromolecule in the A phase. The establishment of phase equilibrium in the presence of ligand, which refers to the equalization of chemical potentials of the macromolecule across the two phases, leads to the following scenarios. Assuming the ligand binds to the macromolecule in both phases, then the saturation concentration for phase separation of the macromolecule in the presence of the ligand becomes $c_A^{(L)} = c_A(p^A/p^B)$ (30, 31). Here, p^A and p^B are the binding polynomials for the binding of ligand to species in the A and B phases, respectively. For reversible macromolecular associations, a binding polynomial quantifies the sum of the activities (concentrations) of all bound and unbound species involving the macromolecule relative to the activity (concentration) of the unbound macromolecule (62). The negative logarithm of the binding polynomial is proportional to the free energy of binding. Accordingly, if p^A is greater than p^B, then species in the A phase will have higher affinity for the ligand when compared with species in the B phase. In this scenario, ligand binding will help retain a higher concentration of the macromolecule in the A phase, and hence $c_A^{(L)}$ will be greater than c_A. The converse is true if p^A is less than p^B, which would engender increased partitioning of the ligated macromolecule into the dense phase, B. Finally, ligand binding will not impact the phase equilibrium if $p^A = p^B$.

To understand the basic tenets of polyphasic linkage, we consider the impact that profilin binding has on the phase boundary between the M and S phases. In accordance with polyphasic linkage, if profilin were to bind preferentially to the M phase of Htt-NTFs, then the value of c_S would shift to higher concentrations (Fig. 2*b*, *ii*). In contrast, preferential binding to the S phase would shift c_S to lower concentrations (Fig. 2*b*, *iii*). Equivalent binding to species in the two phases would have no effect on the phase boundary (Fig. 2*b*, *iv*). These scenarios are depicted qualitatively in Fig. 2*b*.

Given knowledge of the phase diagram for Htt-NTFs, we hypothesized that profilin must reduce aggregation/phase separation through a polyphasic linkage mechanism, whereby it binds preferentially to M-phase species to destabilize the S and F phases. This hypothesis is directly testable by measuring saturation concentrations in the absence (*e.g.* Fig. 1*d*) and presence of profilin. A lack of change in the saturation concentration in the presence of profilin may indicate no binding or

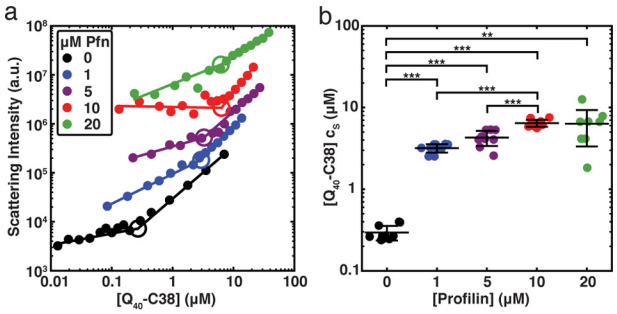

Figure 3. Profilin shifts the c_S of Q_{40}-C38 through polyphasic linkage. *a*, right-angle static light scattering of Q_{40}-C38 solution as a function of Q_{40}-C38 concentration, in the absence of profilin (*black*, reproduced from Fig. 1*d*) and in the presence of 1, 5, 10, or 20 μM profilin (*blue*, *purple*, *red*, and *green*, respectively) in 20 mM Tris, 5 mM EDTA, 1 mM DTT, pH 7.4. One representative trial at each profilin concentration is shown, and the plots are arbitrarily offset in the vertical direction for the sake of clarity. The values of c_S for Q_{40}-C38 in the presence of each of these concentrations of profilin are indicated by the discontinuity in slope, and each discontinuity is marked using an *open circle*. We note that the slope of the low-concentration arm of the 10 μM profilin data set differs from the other data sets and suspect that this is the result of an experimental artifact (see "Experimental procedures"). Regardless, there was good agreement between trials concerning the location of the discontinuity, as is evident in *b*. *b*, Q_{40}-C38 c_S measured in *a* plotted as a function of the profilin concentration. The intersections of "best-fit" lines from jackknife sampling for at least three trials at each profilin concentration are plotted (see supporting Methods). *Black bars*, mean ± S.D. (*error bars*). Statistical significance was determined by Welch's ANOVA with Games–Howell post hoc analysis. **, $p < 0.01$; ***, $p < 0.001$. *Colors* are the same as in *a*. Even the lowest concentration of profilin tested (1 μM) resulted in more than an order of magnitude increase in the c_S of Q_{40}-C38.

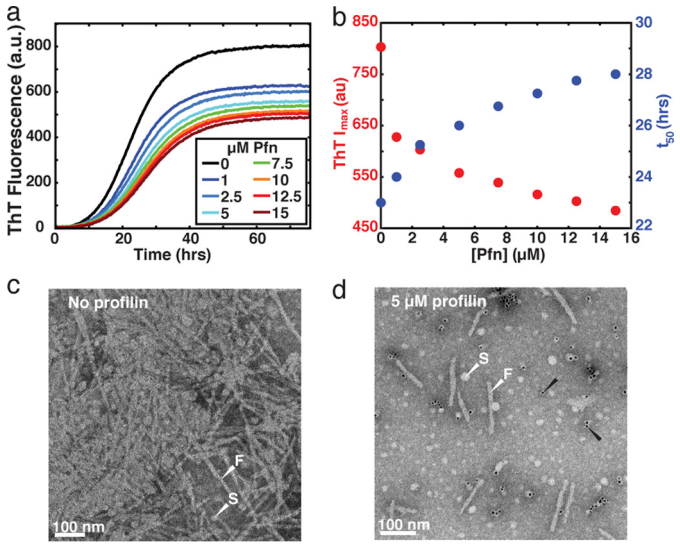

Figure 4. The influence of profilin on Q_{40}-C38 fibril formation. *a*, ThT assay of Q_{40}-C38 aggregation kinetics at a fixed concentration of Q_{40}-C38 in the presence of various concentrations of profilin. ThT fluoresces upon binding to amyloid-like fibrils, thus providing a readout of fibril formation in the form of fluorescence intensity. *b*, the plateau ThT fluorescence intensity (ThT I_{max}) and the time to reach 50% I_{max} (t_{50}) (*red* and *blue*, respectively), are plotted as a function of profilin concentration. *c*, representative TEM image of a sample taken from an aggregation assay in the absence of profilin showing an abundance of long fibrils coexisting with spherical aggregates. *d*, representative TEM image of a sample taken from an aggregation assay in the presence of 5 μM profilin shows short fibrils few in number as well as spherical aggregates and oligomers or possibly monomers. His-tagged profilin was marked on the carbon-coated copper grid with 5-nm nickel-nitrilotriacetic acid-gold nanoparticles, which appear as *black dots*. Examples of gold nanoparticles are indicated with *black arrows*. Examples of fibrils and spheres are indicated with *white arrows labeled F* and *S*, respectively, in *c* and *d*.

equivalent binding to the two phases. Alternatively, if the saturation concentration changes in the presence of profilin, then the direction of the change will point to the phase that is stabilized or destabilized by ligand binding.

We performed static right-angle light scattering measurements as a function of Q_{40}-C38 concentration in the presence of fixed concentrations of profilin (Fig. 3*a*). The discontinuity in the concentration-dependent scattering profile, which quantifies the c_S of Q_{40}-C38, increases by more than an order of magnitude in the presence of profilin. The inferred value of c_S increases from ~290 nM in the absence of profilin to ~3 μM in the presence of 1 μM profilin (Fig. 3*b* and Table S1). The value of c_S increases with increasing profilin concentration and appears to plateau to a value of ~6 μM in the presence of 20 μM profilin. The increase of c_S in the presence of profilin suggests that binding to profilin destabilizes the S phase. Our reasoning, based on thermodynamic linkage relations, suggests that profilin must bind preferentially to either the M or F phases, whereby ligand-induced changes to one phase boundary in a closed system will necessarily lead to changes in all phase boundaries.

To infer the impact of profilin binding on the stability of the F phase, we quantified the impact of the presence of profilin on the kinetics of fibril formation and the morphologies of fibrils that are formed in the presence of profilin. We monitored the kinetics of fibril formation by following the rate of change of ThT fluorescence. In these measurements, we fixed the total concentration (c_t) of Q_{40}-C38 to be above its intrinsic c_F such that the degree of supersaturation with respect to c_F, quantified as $\ln(c_t/c_F)$, was 0.75. A series of independent measurements were performed, each in the presence of a different concentration of profilin (Fig. 4*a*). As the profilin concentration increased, fewer fibrillar aggregates formed, as indicated by the maximum ThT intensity reached, and the rate of fibril formation was slowed, as indicated by the time required to reach half-maximum intensity (Fig. 4*b*). These results are consistent with ligand binding shifting the intrinsic c_F of Q_{40}-C38 to higher values. Negative stain TEM images of samples from the end point of aggregation assays confirms that in the presence of profilin, fibrils appear to be reduced in both length and number (Fig. 4, *c* and *d*).

To determine whether profilin was bound to fibrils or other aggregates, we tagged profilin on TEM sample grids with 5-nm gold nanoparticles functionalized with a nickel moiety that binds histidine tags. Imidazole buffer washes were used to prevent nonspecific binding. The gold nanoparticles were found predominantly near oligomers and less frequently near larger spherical aggregates or fibrils (Fig. 4*d*). In total, 257 gold nanoparticles in three TEM images were counted and classified according to their association with M-, S-, or F-phase species. Of the 257 gold nanoparticles, 133 (51.7%) were associated with M-phase species, whereas only 49 (19.1%) were associated with S-phase species, and 75 (29.2%) were associated with F-phase species (see supporting Methods for details). Taken together, the results from Figs. 3 and 4 suggest that profilin binding destabilizes the S and F phases in favor of the M phase via polyphasic linkage. This is attributable to the preferential binding of profilin to species of the M phase. In light of this suggestion, we investigated the mode of binding of profilin to species of the M

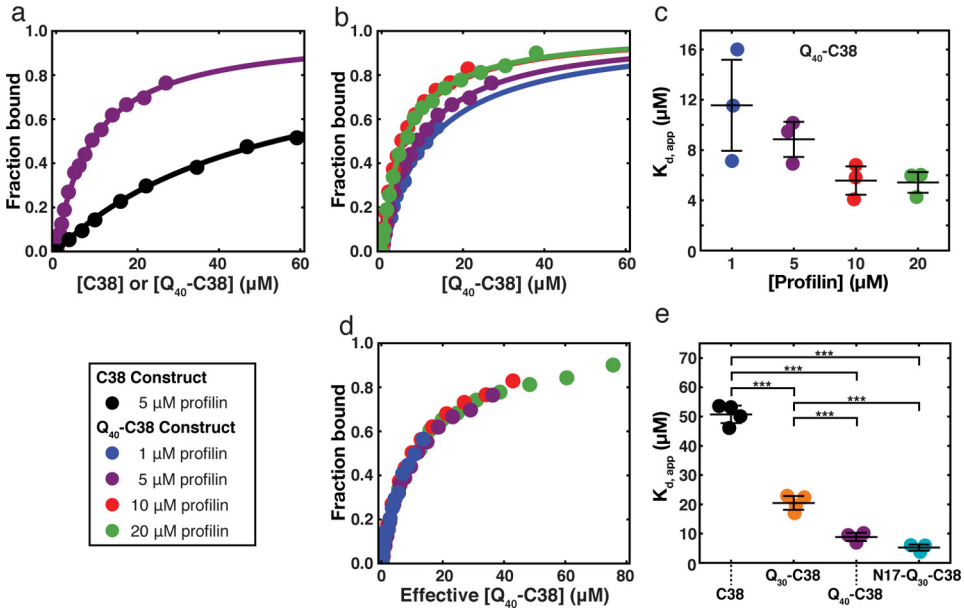

Figure 5. The effect of polyQ on profilin binding to Htt-NTFs. *a*, binding isotherms of 5 μM profilin with C38 (*black*) *versus* 5 μM profilin with Q_{40}-C38 (*purple*). The presence of the polyQ tract enhances binding at a given profilin concentration. *b*, binding isotherms of 1, 5, 10, and 20 μM profilin with Q_{40}-C38. *c*, apparent K_d values extracted from fits to the isotherms in *a*. A slight dependence of the $K_{d,app}$ on profilin concentration is apparent, although a one-way ANOVA test of the data revealed that the difference is at the threshold of significance, with a *p* value of 0.0507. *d*, when the concentration of each binding isotherm is adjusted by the magnitude of the profilin-dependent shift in c_S ($c_{S,Pfn}/c_{S,intrinsic}$), then the binding isotherms collapse to a single curve, confirming that the observed dependence of $K_{d,app}$ on the profilin concentration is due to profilin-dependent changes in partitioning of Q_{40}-C38 molecules between the M- and S-phases (see "Results" for details). *e*, comparison of apparent K_d values of C38 (*black*) and Q_{40}-C38 (*purple*) with Q_{30}-C38 (*orange*) and N17-Q_{30}-C38 (*teal*), all measured at 5 μM profilin. Statistical significance was determined by one-way ANOVA with Tukey's range test for post hoc analysis. ***, $p < 0.001$. Binding of profilin appears to correlate with aggregation propensity, which is dictated by polyQ length (Q_{30} *versus* Q_{40}) and flanking sequence context (absence or presence of N17). *Black bars* in *c* and *e* indicate the mean ± S.D. (*error bars*) of three or four independent experiments.

phase to uncover a mechanistic understanding of how the M-phase species are stabilized by profilin binding.

Profilin binds to polyP and binding to Htt-NTFs is aided by the presence of the polyQ region

Profilin binds to polyP through a distinct binding site (47–49, 63, 64). The C38 region of Htt-NTFs encompasses two polyP modules (P_{11} and P_{10}) connected by a 17-residue linker (L17) that is rich in Pro, Gln, and Leu residues. As described above, pulldowns performed by Shao *et al.* (51) indicated that the polyP binding pocket of profilin is responsible for binding Htt-NTFs, and our follow-up experiments confirmed that the Htt-NTF C38 stretch is essential to elicit profilin-dependent reduction in aggregation (Fig. S1). Thus, the main mode of interaction between profilin and Htt-NTFs is through direct binding to polyP modules in the C38 region.

To place the interactions of profilin with Htt-NTFs on a quantitative footing, we first measured the binding affinity of profilin to polyP in isolation and in the context of C38 (Fig. S3*a*). We used a fluorescence-based binding assay developed by the Pollard laboratory (49). The binding of polyP ligands to profilin leads to increased fluorescence of Trp-3 and Trp-31 from profilin and a blue shift in the maximal wavelength for Trp fluorescence. We used the fluorescence intensity at 320 nm to quantify the binding of profilin to C38 and compared this with the binding of profilin to an 11-mer of polyP (P_{11}). The binding data for P_{11} and C38 were fit to a model with a single binding site (Fig. S3*a*). Here, profilin is the macromolecule, and P_{11} and C38 are the ligands. The dissociation constant $K_{d,P11}$ that we obtain for the binding of P_{11} to profilin (106.5 μM) matches published estimates (49, 65).

In contrast to P_{11}, C38 is a bivalent ligand with two polyP modules. The dissociation constant $K_{d,C38}$ for the binding of profilin to C38 is 50.7 μM, which is essentially half the value of $K_{d,P11}$. When we used the number of polyP tracts to normalize the binding isotherms, the isotherms overlay, indicating that the difference in dissociation constants results from C38 having two binding sites for profilin (Fig. S3*b*). The fact that $K_{d,P11} \approx 2K_{d,C38}$ suggests that profilin binds to the two sites on C38 independently. This was confirmed using Hill and Scatchard analyses as shown in Fig. S4. We obtained an average value of 1.0 as our estimate for the Hill coefficient, which quantifies the size of the cooperative unit for binding. Therefore, we conclude that profilin molecules bind independently to the individual polyP modules in C38. Accordingly, the lowering of $K_{d,C38}$ derives exclusively from doubling the concentration of polyP by doubling the polyP valency in C38. Additionally, in three separate sets of fluorescence experiments, we fixed the concentration of profilin and titrated the concentrations of N17, polyQ (Q_{30}), and L17. We were unable to detect any measurable changes in Trp fluorescence of profilin for a range of concentrations of each of N17, polyQ, and L17 as ligands. We conclude that these modules have undetectably weak interactions with profilin at or near the polyP-binding site, although interactions with other distal sites cannot be ruled out.

Next, we measured the binding of profilin to Q_{40}-C38, a construct possessing both a polyQ and C38 tract. The binding data were fit to a 1:1 binding model, and the inferred apparent dis-

sociation constant ($K_{d,app}$) in the presence of 5 μM profilin was found to be 8.8 μM for Q_{40}-C38 (Fig. 5*a*). This value for $K_{d,app}$ is ~6-fold lower than the intrinsic value of $K_{d,C38}$ = 50.7 μM for C38. We also measured the binding of Q_{40}-C38 to profilin at several different fixed concentrations of profilin. These measurements yield a modest decrease in the inferred $K_{d,app}$ as profilin concentration increases (Fig. 5 (*b* and *c*) and Table S1) and suggest one of two possibilities; either profilin undergoes a concentration-dependent change, such as dimerization, that influences its binding to Q_{40}-C38, or alternatively, the effect can be explained in terms of a shift in the concentration of Q_{40}-C38 within the M phase that is caused by profilin binding leading to a modulation of the phase boundary. Rennella *et al.* (66) have shown that profilin only forms sparsely populated oligomers at millimolar concentrations. In our studies, we never exceeded 20 μM profilin. Therefore, it appears safe to conclude that profilin oligomerization is not relevant.

We know now that profilin binding increases the value of c_S. Because the c_S dictates the relative amounts of molecules in the M and S phases, which in turn influences $K_{d,app}$, the observed dependence of $K_{d,app}$ on the profilin concentration could be due to changes in the partitioning of Q_{40}-C38 molecules between the M and S phases resulting from profilin-dependent changes in the c_S. In support of this hypothesis, we find that adjusting the concentration of each binding isotherm by the magnitude of the profilin-dependent shift in c_S (defined here as $c_{S,Pfn}/c_{S,intrinsic}$, where $c_{S,Pfn}$ is the c_S measured in the presence of a given concentration of profilin, and $c_{S,intrinsic}$ is the c_S measured in the absence of profilin (see Fig. 3)), leads to binding isotherms that collapse onto a single master curve (Fig. 5*d*). Therefore, lowering of $K_{d,app}$ at higher profilin concentrations is thermodynamically linked to the profilin-dependent increase in c_S, which increases the M-phase concentration of Q_{40}-C38 molecules.

Our measurements also show that the binding of profilin to Htt-NTF constructs depends on the length of the polyQ segment and the presence of the N17 module (Fig. 5*e* and Table S1). Specifically, the value of $K_{d,app}$, measured in the presence of a fixed profilin concentration, is lowered upon increasing the polyQ length in Q_n-C38 constructs. The value of $K_{d,app}$ is further lowered upon the addition of the N17 module, which is shown by comparing the values for $K_{d,app}$ obtained for N17-Q_{30}-C38 with those obtained for Q_{30}-C38 and Q_{40}-C38. The decrease in $K_{d,app}$ in going from Q_{30}-C38 to N17-Q_{30}-C38 is equivalent to the decrease we observe in going from Q_{30}-C38 to Q_{40}-C38. Given that increasing the polyQ length and the presence of N17 both promote Htt-NTF oligomerization/aggregation (19, 68, 69), these results are consistent with the idea that the binding of profilin to Htt-NTFs is enhanced due to polyQ- and N17-dependent assembly. These results further suggest that oligomerization can engender positive cooperativity in binding, which results in a reduced $K_{d,app}$. Consistent with this interpretation, Scatchard analysis of profilin binding to Q_{40}-C38 indicated positive cooperativity, which was absent in the binding of C38 to profilin (Fig. S4). Additionally, binding of profilin to Q_{30}-C38 and N17-Q_{30}-C38 also exhibited positive cooperativity (Fig. S5), offering a clear correlation between the observed cooperativity and the presence of polyQ or N17-polyQ. However, the evidence for cooperativity in binding does not illuminate the specific mechanism through which profilin binding is enhanced in the presence of polyQ and N17. To explore possible origins for this behavior, we used coarse-grained simulations to explicitly test distinct models.

Modeling suggests that profilin modulates Htt-NTF phase behavior through a combination of specific interactions with polyP and auxiliary interactions with polyQ

To understand the source of the enhanced profilin binding in constructs with flanking N-terminal polyQ regions, we turned to coarse-grained simulations to identify the most plausible mechanistic models for our observations. Our modeling probes how the apparent binding affinity is influenced by increased local concentration and auxiliary interactions that involve regions on profilin that do not overlap with its polyP-binding site. Here, increased local concentration refers to the clustering of polyP segments due to Q_{40}-C38 oligomerization. Auxiliary interactions can also contribute to the apparent increase in binding affinity through multivalency of both Q_{40}-C38 oligomers and profilin (70–72).

Using a phenomenological coarse-grained model, in which we represent profilin and Q_{40}-C38 as multiple spheres or "beads" connected by "springs" that mimic their atomistic architectures, we performed simulations in which binding was quantified between profilin and C38 or Q_{40}-C38 as a function of Q_{40}-C38 oligomer size. The general architecture of the coarse-grained model is shown in Fig. 6*a*. Each profilin molecule has a polyP-binding site that can interact with a polyP segment in either C38 or Q_{40}-C38. We refer to this interaction as the *primary interaction* between profilin and C38 or Q_{40}-C38, given that this is the interaction that is measured in the fluorescence experiments. If an increased local concentration of polyP segments leads to the observed experimental results, then the inclusion of this primary interaction between profilin and C38 or Q_{40}-C38 should be sufficient to reproduce trends in binding that are consistent with the experimental results. In this scenario, Q_{40}-C38 oligomerization should increase the local concentration of C38 modules. To be consistent with experimentally measured data, the simulation results would need to reproduce two key findings: 1) binding to C38 alone should be weaker than binding to Q_{40}-C38 (Fig. 5*a*), and 2) binding to M-phase oligomers should be stronger than binding to S-phase aggregates (Fig. 3). Our results show that a model restricted to inclusion of primary interactions alone will not reproduce either of the experimental observations listed above. Instead, in a model that only accounts for primary interactions, we found that binding to C38 alone was stronger than binding to Q_{40}-C38, and the fraction of profilin molecules bound decreased from that of the Q_{40}-C38 monomer with increasing Q_{40}-C38 oligomer size (Fig. 6*c*). This decrease originates from steric-effects, whereby bound profilin molecules, as well as other C38 stretches, get in the way of unbound profilin, occluding the C38-binding sites and reducing the effective concentration of free C38. This result suggests that an increased local concentration of polyP segments alone cannot account for the experimentally observed trends.

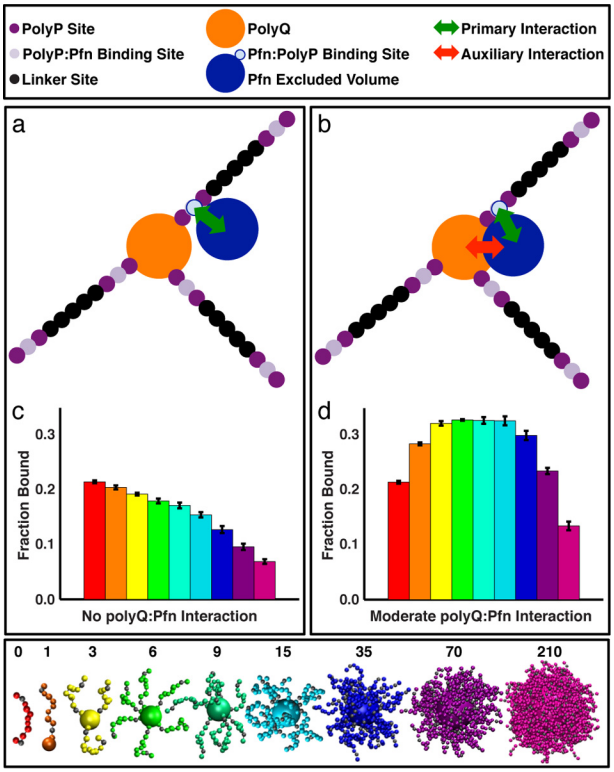

Figure 6. Coarse-grained model identifies an auxiliary interaction necessary for the experimentally observed profilin effect. *a*, visual representation of the architecture used for the increased local concentration model. Q_{40}-C38 and profilin are represented by a set of spheres or "beads" as defined in the *top legend* (see supporting Methods for model details). The primary interaction between profilin (Pfn) and the polyP segment of C38 is denoted by a *green arrow*. *b*, visual representation of the architecture used for the polyQ-Pfn auxiliary interaction model. This model is identical to the increased local concentration model, except an auxiliary interaction is added between the polyQ bead and a nonspecific region of the profilin molecule that does not overlap with the polyP interaction site (*red arrow*). *c* and *d*, fraction of profilin molecules bound to polyP segments observed in coarse-grained simulations of the increased local concentration model and the polyQ-Pfn auxiliary interaction model, respectively. Each simulation contained 210 Pfn molecules and 630 C38 stretches. Additionally, each simulation had a homogeneous distribution of cluster sizes, *X* (*X* = 1, 3, 6, 9, 15, 35, 70, 210), such that the number of clusters multiplied by *X* equaled 630. Each *bar* denotes results from simulations performed with a distinct cluster size as indicated by the *bottom legend*. A cluster size of zero denotes C38 without the polyQ domain. *Gray beads* in the *legend* denote beads that interact with profilin via the primary Pfn-PolyP interaction. *Error bars*, S.E. from five independent simulations.

Next, we tested three distinct models for *auxiliary interactions* that engender multivalency and might explain the observed experimental results. The distinct auxiliary interactions are denoted as polyQ-Pfn interactions, linker-Pfn interactions, and Pfn-Pfn interactions. Here, Pfn refers to profilin. In each of these models, the coarse-grained profilin molecule can engage in an auxiliary interaction, in addition to the primary polyP interaction. In the polyQ-Pfn interaction model, the auxiliary interaction is between the polyQ bead of a Q_{40}-C38 molecule and a nonspecific region on the profilin molecule that is distinct from the polyP-binding site (Fig. 6*b*). This model accounts for experimental results, which suggest that long polyQ tracts and polyQ aggregates can interact nonspecifically with other molecules, given the uniformly "sticky" surface of the polyQ domain (73–76). This model allows for the possibility of a heterotypic multivalent complex in which a single profilin molecule is engaged in two different types of interactions with a Q_{40}-C38 oligomer. In the linker-Pfn interaction model, the auxiliary interaction is between the non-proline residues of L17 and the face opposing the polyP-binding site on profilin, which is enriched in polar residues (Fig. S6*b*) (64). This model queries the possibility of a heterotypic complex in which a single profilin molecule interacts with the glutamine and leucine residues of L17 from one Q_{40}-C38 molecule and a polyP region from another Q_{40}-C38 molecule. Finally, in the Pfn-Pfn interaction model, the auxiliary interaction is a stable dimerization of two profilin molecules (Fig. S6*c*). This model is intended to address published conjectures suggesting that profilin can undergo dimerization (66, 77) and allows for the possibility of a homotypic complex in which a profilin dimer engages in interactions with two distinct polyP sites.

Fig. 6*d* and Fig. S6 (*d–f*) show the results obtained from simulations based on each of the three auxiliary interaction models. Only the polyQ-Pfn interaction model (Fig. S7) was able to reproduce both experimental findings. Specifically, our calculations based on this model show that binding to C38 alone was weaker than binding to Q_{40}-C38, and binding to small oligomers was stronger than binding to large aggregates (Fig. 6*d* and Fig. S6*d*). Neither the linker-Pfn interaction model nor the Pfn-Pfn interaction model was able to reproduce both experimental results (Fig. S6, *e* and *f*). Taken together, the results from coarse-grained modeling suggest that an auxiliary interaction between profilin and the polyQ domain within Htt-NTFs probably plays an important role in modulating Q_{40}-C38 phase behavior.

Discussion

We showed that aggregation and phase separation of Htt-NTF constructs is characterized by at least three distinct phases. We also showed that the values of c_S shift to higher concentrations, and fibril formation is destabilized in the presence of profilin. This implies that profilin binds preferentially to M-phase species of Htt-NTFs. In most situations that involve binding between a protein and a heterogeneous mixture of oligomers or aggregates, it would be impossible to determine which species in the heterogeneous mixture is/are preferentially bound. However, with a quantitative description of the phase boundaries that govern the distribution of species, we can apply the concepts of polyphasic linkage to identify the phases that are preferentially bound. This is important because if one or more types of aggregate species are toxic to cells, it is essential that one knows which phase is preferentially bound by a ligand or small molecule to understand its effect on the equilibrium and therefore toxicity. This is essential, given that preferential binding to any one species will influence the equilibrium populations of all aggregate species.

We complemented experiments that probe the impact of profilin binding on phase boundaries with measurements of binding isotherms. Our results from experiments and modeling suggest that increasing the local concentration of C38 domains through polyQ-dependent oligomerization cannot explain the results of the binding experiments. Instead, the experimental results are consistent with a model that invokes a combination of primary interactions between polyP and profilin and weak, auxiliary interactions between the polyQ domain and profilin.

Previous studies have shown that increases in apparent affinity can result from a combination of specific and nonspecific interactions and that the magnitude of the increase depends on the size of the nonspecific interaction surface (78). This result is consistent with the decreasing $K_{d,app}$ between Q_n-C38 and profilin observed as the length of the polyQ domain was increased. Based on our coarse-grained simulations, we propose that these auxiliary interactions will involve sites on the surface of profilin that are non-overlapping with the polyP-binding site.

Given that overexpression of profilin in cellular models reduces Htt-NTF dependent toxicity (51) and that multivalency appears to be important for the relevant interactions needed between profilin and Htt-NTFs, it appears that exploiting multivalent interactions might be an important strategy for designing therapeutics targeting Htt-NTFs. Such strategies have been developed for the design of antimicrobial peptides as a means to treat bacterial infections (79) and for selective targeting of different amyloid-β species as a means to alleviate neurotoxicity (80). Our results suggest that the design of multivalent peptides can increase binding affinity, thus allowing these peptides to compete more effectively with other ligands in the cell.

Understanding the effects of Htt-NTF binding partners in terms of polyphasic linkage

Htt-NTFs engage in a wide variety of protein-protein interactions (75, 81–90). These "interactomes" are tissue-specific, and within a tissue type, the nodes and edges in an interaction network are known to be different for Htt-NTFs with wildtype *versus* mutant polyQ lengths (90). The polyphasic linkage formalism of Wyman and Gill (30, 31) provides a route for modeling the phase behavior of Htt-NTFs and quantifying the impact of protein-protein interaction nodes on this phase behavior.

From a phenotypic standpoint, gain-of-function toxicity resulting from the accumulation of S- and F-phase species appears to be weakened by interactions with profilin (91). It is also conceivable that other heterotypic interactions with Htt-NTFs modulate phase boundaries in very different ways. For example, preferential binding to S-phase species will, via linkage effects, lead to a lowering of c_S and a widening of the gap between c_S and c_F. Conversely, preferential binding to F-phase fibrils will lower c_F, destabilize the M and S phases, and promote the formation of long, stable fibrils. Indeed, there appear to be proteins in interactomes that preferentially interact with high-molecular weight S- and F-phase species. For example, SH3GL3, a protein involved in endocytosis, binds the linker region of C38 and has been shown to bind preferentially to mutant Htt-NTF and accelerate mutant Htt-NTF aggregation (92, 93). Recent studies showed that SH3GL3 binds monomeric wildtype and mutant Htt-NTF similarly (92). However, upon introduction of aggregated mutant Htt-NTF, retention of mutant Htt-NTF on SH3GL3 pull-downs was increased in a manner that correlated with increased aggregation of mutant Htt-NTF. Together, these results suggest that the binding of SH3GL3 increases the stability of the F phase of mutant as opposed to wildtype Htt-NTFs due to the increased aggregation propensity of mutant Htt-NTFs. This points to preferential binding of SH3GL3 to larger fibrillar aggregates. This should shift c_F to lower values. Our suggestion is consistent with findings of increased fibril formation when SH3GL3 is co-expressed with Htt-NTFs in cellular models (93).

Profilin is one of the most abundant cellular proteins, with a copy number of around 10^7/cell (94). We speculate that the high natural abundance coupled to its intrinsic ability to bind polyP tracts leads to the serendipitous consequence of being able to influence cellular aggregation and phase separation of Htt-NTFs. This raises the possibility that there may be other naturally occurring proteins that show even greater potency but are not readily identified in cellular screens due to their low cellular concentrations. Indeed, given the wide prevalence of polyP-binding modules, such as WW, SH3, and EVH1 domains (95), a plausible approach would be to identify other, potentially more powerful modifiers of intracellular aggregation and phase separation. This could be achieved by driving overexpression of putative modifiers based on the combined criteria of polyP-binding domains and low molecular weight. Similarly, profilin itself may represent a useful scaffold upon which protein engineering (96) could be used to tune its specificity and affinity for Htt-NTFs. Regardless of the "physiological relevance" of interactions between profilin and Htt-NTFs, it is clear that profilin is capable of reducing Htt-NTF aggregation and toxicity, and therefore it serves as a useful model for the design of therapeutics for Huntington's disease. The same logic applies to any protein or molecule shown to bind a pathologically aggregating protein and reduce its toxicity, irrespective of the "physiological relevance" of that interaction.

Experimental procedures

Synthesis of peptide constructs

Peptides were purchased from Watsonbio Sciences (Houston, TX) in crude form and stored at −20 °C until use.

Peptide purification

Crude peptides were disaggregated in undiluted formic acid, diluted to 30% formic acid with deionized water, and purified using reverse-phase HPLC on a semipreparative 300SB-C3 reverse-phase column (Agilent, Santa Clara, CA). Elution fractions containing pure peptide were pooled, lyophilized, and stored at −20 °C until use.

Peptide disaggregation

Peptides were disaggregated using the method reported previously (19). Further details are given in the supporting Methods.

Recombinant expression and purification of profilin

The genetic sequence of human profilin-1 (hereafter referred to as profilin) in the pRK174 vector, which includes a T7 promoter, a His_6 tag, and an ampicillin resistance gene, was verified by DNA sequencing (GENEWIZ, South Plainfield, NJ). The pRK174-His_6-profilin plasmid was transformed into *Escherichia coli* BL21 (DE3) cells. Cells were grown at 37 °C in Luria broth with 100 μg/ml ampicillin to an A_{600} of ~0.8 with shaking at 220 rpm. Protein expression was induced by the addition of 1.0 mM isopropyl 1-thio-β-D-galactopyranoside and harvested

by centrifugation at 4,500 rpm for 30 min at 4 °C, ~4 h after induction. The resulting cell pellets were resuspended in nickel column buffer (20 mM sodium phosphate, 500 mM NaCl, 30 mM imidazole, pH 7.5). Protease inhibitors (Roche Applied Science) were added to the suspension, and cells were lysed on ice by sonication using four 30-s pulses with 2-min rests between each pulse. Lysates were clarified by centrifugation at 20,000 rpm for 30 min at 4 °C in a Sorvall SS34 rotor. Protein was isolated from the resulting supernatant by affinity chromatography using a HisTrap HP column on an AKTA pure FPLC system (GE Healthcare). Fractions containing His-tagged proteins were pooled and dialyzed against 20 mM Tris, 100 mM NaCl, 5 mM EDTA, 0.5 mM tris(2-carboxyethyl)phosphine, pH 7.7, using Spectra/Por dialysis membrane (Spectrum Laboratories, Rancho Dominguez, CA) at 4 °C for 48 h with multiple changes to fresh buffer. Sample purity was typically greater than 95%, as determined by Coomassie G250-stained SDS-PAGE. Protein concentrations were determined by UV spectroscopy using an extinction coefficient of 18,450 M^{-1} cm^{-1}. The protein stock was aliquoted, flash-frozen in liquid nitrogen, and stored at −80 °C until use. Protein yield was typically >160 mg/liter of cell culture.

Steady-state tryptophan fluorescence

Steady-state fluorescence emission spectra of profilin were measured using a PTI QuantaMaster fluorometer (Horiba Scientific, Edison, NJ). Disaggregated peptide was reconstituted in Tris buffer as described, and profilin was added from a concentrated stock solution. Profilin was maintained at a constant concentration while the concentration of the Htt-NTF construct was titrated by dilution using a profilin solution that was identical to the Htt-NTF + profilin solution, except that it contained no Htt-NTF peptide. This solution was then used for stepwise dilution of the Htt-NTF + profilin solution. All experiments were repeated at least three times. The reverse titration was also performed (by adding concentrated Htt-NTF stock to increase the total sample concentration) with C38 and profilin, and the same result was obtained regardless of the direction of the titration. Binding experiments were carried out in a 3 × 3-mm fluorometer cell (Starna, Atascadero, CA) at 22 °C. Emission spectra were recorded from 310 to 380 nm, with excitation at 295 nm (5-nm slit width). All spectra were corrected for the contribution of buffer. Fluorescence emission intensities at 320 nm were recorded as a function of Htt-NTF concentration. Binding isotherms were fit using nonlinear least-squares analysis as described in the supporting Methods.

Right-angle static light scattering

The intensity of static right-angle light scattering at 350 nm (I) was measured at each titration point immediately before each fluorescence measurement using the same instrument. The intersection of linear fits to I *versus* [Q_{40}-C38] (log-log plot) identified discontinuity points in the concentration dependence of the scattering intensity, which is indicative of a phase transition. Due to the sensitive nature of the measurement, it was somewhat prone to experimental artifact (such as persistent aggregates resulting from incomplete peptide disaggregation or the introduction of dust or other foreign material to the sample). Even minute contaminants can introduce noise or otherwise obscure the signal of interest, and it was not always possible to completely exclude minor contaminants from the sample. Because of this, none of the trials carried out were completely devoid of artifact. However, despite these challenges, our method yielded good agreement across trials with regard to the location of the discontinuity. In contrast, a physical interpretation of the slope on either side of a discontinuity was not pursued further because of sensitivity to the potential artifacts discussed above. The method of determining optimal linear fits is given in the supporting Methods.

Negative-stain TEM

Samples were either recovered from fluorescence and light scattering experiments or freshly prepared in the same manner (*i.e.* samples were disaggregated using the TFA/HFIP method, reconstituted in buffer, and then diluted to various concentrations). Sample solutions were briefly vortexed, and 10 μl of each sample was applied to a parafilm strip. Carbon-coated copper grids (Ted Pella, catalog no. 01843, 300-mesh) were inverted on the sample droplets for 15–20 s. The grids were then washed by inverting them on 50-μl droplets of deionized water for about 5 s, after which the water was wicked away with Whatman filter paper. Finally, the grids were inverted on droplets of 0.5% uranyl acetate for 1 min, wicked dry with Whatman filter paper, and air-dried overnight. For samples in the presence of (His-tagged) profilin, the location of profilin was marked on the EM grid with 5-nm nickel-nitrilotriacetic acid-gold nanoparticles. Unbound profilin was rinsed away with buffer before applying the gold nanoparticle solution, and unbound or nonspecifically bound gold nanoparticles were washed away with imidazole buffer before imaging. TEM images were collected on an FEI Tecnai G2 Spirit transmission electron microscope operating at 120 kV and equipped with a Gatan digital camera. All images are representative of multiple distinct locations on the EM grid.

Sedimentation assay

Peptides were disaggregated using the TFA/HFIP method described in the supporting Methods and aliquoted (1.5 ml) into 1.6-ml Eppendorf tubes, which were then capped and sealed with Parafilm. The samples were incubated in a water bath at the specified temperature for 2 weeks. After 2 weeks of incubation, the samples were centrifuged for 100 min at 25,000 × *g*. Two aliquots of the supernatant were removed, and the peptide concentrations were determined using the micro-BCA (bicinchoninic acid) assay (97) (Pierce).

Thioflavin T fluorescence measurements

ThT assays were performed as described by Crick *et al.* (19). Further details are given in the supporting Methods.

Statistical significance tests

Data sets were first subjected to either Bartlett's or Levene's test for equal variance. Data sets exhibiting equal variance were subjected to a one-way ANOVA with Tukey's range test for post hoc analysis. Data sets that did not pass the equal variance test were subjected to Welch's ANOVA with Games–Howell post hoc analysis. Significance levels are indicated with one,

two, or three asterisks corresponding to the following p value cutoffs: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Analyses were implemented in Python.

Molecular simulations

All atom simulations used to parameterize our coarse-grained model were performed using the CAMPARI simulation package utilizing the ABSINTH implicit solvation model and force field paradigm (98, 99). Coarse-grained Langevin dynamics simulations were performed using the LAMMPS simulation package (67). Details regarding the set-up and execution of the simulations are given in the supporting Methods.

Author contributions—A.E.P., K.M.R., T.S.H., and R.V.P. conceptualization; A.E.P. K.M.R., S. L. C., and A. L. data curation; A.E.P., K.M.R., and S.L.C., A.L. formal analysis; A.E.P., K.M.R., S.L.C., and A.L. investigation; A.E.P. and K.M.R. visualization; A.E.P., K.M.R., T.S.H., S.L.C., A.L., M.I.D., and R.V.P. methodology; A.E.P., K.M.R., and R.V.P. writing-original draft; A.E.P., K.M.R., and R.V.P. writing-review and editing; K.M.R. and T.S.H. software; M.I.D. and R.V.P. funding acquisition; R.V.P. resources; R.V.P. supervision; R.V.P. project administration.

Acknowledgments—We thank Carlos Castañeda, Jeong-Mo Choi, Alex Holehouse, Ralf Langen, and J. Paul Taylor for helpful discussions.

References

1. Ross, C. A., and Poirier, M. A. (2004) Protein aggregation and neurodegenerative disease. *Nat. Med.* **10,** S10–S17 CrossRef Medline
2. Aguzzi, A., and O'Connor, T. (2010) Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nat. Rev. Drug Discov.* **9,** 237–248 CrossRef Medline
3. Pappu, R. V., Wang, X., Vitalis, A., and Crick, S. L. (2008) A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **469,** 132–141 CrossRef Medline
4. Greer, S. C. (2002) Reversible polymerizations and aggregations. *Annu. Rev. Phys. Chem.* **53,** 173–200 CrossRef Medline
5. Sfatos, C. D., Gutin, A. M., and Shakhnovich, E. I. (1994) Phase transitions in a "many-letter" random heteropolymer. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **50,** 2898–2905 CrossRef Medline
6. Shakhnovich, E. I., and Finkelstein, A. V. (1989) Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers* **28,** 1667–1680 CrossRef Medline
7. Li, P., Banjade, S., Cheng, H. C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V., King, D. S., Banani, S. F., Russo, P. S., Jiang, Q. X., Nixon, B. T., and Rosen, M. K. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483,** 336–340 CrossRef Medline
8. Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., and Hyman, A. A. (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324,** 1729–1732 CrossRef Medline
9. Zhang, H., Elbaum-Garfinkle, S., Langdon, E. M., Taylor, N., Occhipinti, P., Bridges, A. A., Brangwynne, C. P., and Gladfelter, A. S. (2015) RNA controls PolyQ protein phase transitions. *Mol. Cell* **60,** 220–230 CrossRef Medline
10. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., and Taylor, J. P. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163,** 123–133 CrossRef Medline
11. Su, X., Ditlev, J. A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D. S., Taunton, J., Rosen, M. K., and Vale, R. D. (2016) Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352,** 595–599 CrossRef Medline
12. Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18,** 285–298 CrossRef Medline
13. Hyman, A. A., Weber, C. A., and Jülicher, F. (2014) Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30,** 39–58 CrossRef Medline
14. Brangwynne, C. P., Tompa, P., and Pappu, R. V. (2015) Polymer physics of intracellular phase transitions. *Nat. Phys.* **11,** 899–904 CrossRef
15. Harmon, T. S., Holehouse, A. S., Rosen, M. K., and Pappu, R. V. (2017) Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6,** e30294 Medline
16. Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., Kriwacki, R. W., Pappu, R. V., and Brangwynne, C. P. (2016) Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165,** 1686–1697 CrossRef Medline
17. Jacobs, W. M., and Frenkel, D. (2017) Phase transitions in biological systems with many components. *Biophys. J.* **112,** 683–691 CrossRef Medline
18. Jain, S., Wheeler, J. R., Walters, R. W., Agrawal, A., Barsic, A., and Parker, R. (2016) ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell* **164,** 487–498 CrossRef Medline
19. Crick, S. L., Ruff, K. M., Garai, K., Frieden, C., and Pappu, R. V. (2013) Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 20075–20080 CrossRef Medline
20. Riback, J. A., Katanski, C. D., Kear-Scott, J. L., Pilipenko, E. V., Rojek, A. E., Sosnick, T. R., and Drummond, D. A. (2017) Stress-triggered phase separation is an adaptive, Evolutionarily tuned response. *Cell* **168,** 1028–1040.e19 CrossRef Medline
21. Murakami, T., Qamar, S., Lin, J. Q., Schierle, G. S., Rees, E., Miyashita, A., Costa, A. R., Dodd, R. B., Chan, F. T., Michel, C. H., Kronenberg-Versteeg, D., Li, Y., Yang, S. P., Wakutani, Y., Meadows, W., *et al.* (2015) ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron* **88,** 678–690 CrossRef Medline
22. Audas, T. E., Audas, D. E., Jacob, M. D., Ho, J. J., Khacho, M., Wang, M., Perera, J. K., Gardiner, C., Bennett, C. A., Head, T., Kryvenko, O. N., Jorda, M., Daunert, S., Malhotra, A., Trinkle-Mulcahy, L., *et al.* (2016) Adaptation to stressors by systemic protein amyloidogenesis. *Dev. Cell* **39,** 155–168 CrossRef Medline
23. Weber, S. C., and Brangwynne, C. P. (2012) Getting RNA and protein in phase. *Cell* **149,** 1188–1191 CrossRef Medline
24. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., *et al.* (2015) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162,** 1066–1077 CrossRef Medline
25. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., *et al.* (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149,** 753–767 CrossRef Medline
26. Ciryam, P., Kundra, R., Morimoto, R. I., Dobson, C. M., and Vendruscolo, M. (2015) Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol. Sci.* **36,** 72–77 CrossRef Medline
27. Ciryam, P., Tartaglia, G. G., Morimoto, R. I., Dobson, C. M., and Vendruscolo, M. (2013) Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Rep.* **5,** 781–790 CrossRef Medline
28. Ciryam, P., Lambert-Smith, I. A., Bean, D. M., Freer, R., Cid, F., Tartaglia, G. G., Saunders, D. N., Wilson, M. R., Oliver, S. G., Morimoto, R. I., Dobson, C. M., Vendruscolo, M., Favrin, G., and Yerbury, J. J. (2017) Spinal motor neuron protein supersaturation patterns are associated with inclusion body formation in ALS. *Proc. Natl. Acad. Sci. U.S.A.* **114,** E3935–E3943 CrossRef Medline

29. So, M., Hall, D., and Goto, Y. (2016) Revisiting supersaturation as a factor determining amyloid fibrillation. *Curr. Opin. Struct. Biol.* **36,** 32–39 CrossRef Medline
30. Wyman, J., and Gill, S. J. (1980) Ligand-linked phase changes in a biological system: applications to sickle cell hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **77,** 5239–5242 CrossRef Medline
31. Wyman, J., and Gill, S. J. (1990) *Binding and Linkage: Functional Chemistry of Biological Macromolecules*, University Science Books, Mill Valley, CA
32. Walker, F. O. (2007) Huntington's disease. *Lancet* **369,** 218–228 CrossRef Medline
33. Crook, Z. R., and Housman, D. E. (2012) Dysregulation of dopamine receptor D2 as a sensitive measure for Huntington disease pathology in model mice. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 7487–7492 CrossRef Medline
34. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431,** 805–810 CrossRef Medline
35. Bates, G. (2003) Huntingtin aggregation and toxicity in Huntington's disease. *Lancet* **361,** 1642–1644 CrossRef Medline
36. Arrasate, M., and Finkbeiner, S. (2012) Protein aggregates in Huntington's disease. *Exp. Neurol.* **238,** 1–11 CrossRef Medline
37. Nucifora, L. G., Burke, K. A., Feng, X., Arbez, N., Zhu, S., Miller, J., Yang, G., Ratovitski, T., Delannoy, M., Muchowski, P. J., Finkbeiner, S., Legleiter, J., Ross, C. A., and Poirier, M. A. (2012) Identification of novel potentially toxic oligomers formed *in vitro* from mammalian-derived expanded huntingtin exon-1 protein. *J. Biol. Chem.* **287,** 16017–16028 CrossRef Medline
38. Preisinger, E., Jordan, B. M., Kazantsev, A., and Housman, D. (1999) Evidence for a recruitment and sequestration mechanism in Huntington's disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **354,** 1029–1034 CrossRef Medline
39. Donaldson, K. M., Li, W., Ching, K. A., Batalov, S., Tsai, C. C., and Joazeiro, C. A. (2003) Ubiquitin-mediated sequestration of normal cellular proteins into polyglutamine aggregates. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 8892–8897 CrossRef Medline
40. Kim, W. Y., Fayazi, Z., Bao, X., Higgins, D., and Kazemi-Esfarjani, P. (2005) Evidence for sequestration of polyglutamine inclusions by *Drosophila* myeloid leukemia factor. *Mol. Cell. Neurosci.* **29,** 536–544 CrossRef Medline
41. Schipper-Krom, S., Juenemann, K., Jansen, A. H., Wiemhoefer, A., van den Nieuwendijk, R., Smith, D. L., Hink, M. A., Bates, G. P., Overkleeft, H., Ovaa, H., and Reits, E. (2014) Dynamic recruitment of active proteasomes into polyglutamine initiated inclusion bodies. *FEBS Lett.* **588,** 151–159 CrossRef Medline
42. Zeng, L., Wang, B., Merillat, S. A., N Minakawa, E., Perkins, M. D., Ramani, B., Tallaksen-Greene, S. J., Costa, M. D. C., Albin, R. L., and Paulson, H. L. (2015) Differential recruitment of UBQLN2 to nuclear inclusions in the polyglutamine diseases HD and SCA3. *Neurobiol. Dis.* **82,** 281–288 CrossRef Medline
43. O'Nuallain, B., and Wetzel, R. (2002) Conformational Abs recognizing a generic amyloid fibril epitope. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 1485–1490 CrossRef Medline
44. Yang, W., Dunlap, J. R., Andrews, R. B., and Wetzel, R. (2002) Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. *Hum. Mol. Genet.* **11,** 2905–2917 CrossRef Medline
45. Ramdzan, Y. M., Trubetskov, M. M., Ormsby, A. R., Newcombe, E. A., Sui, X., Tobin, M. J., Bongiovanni, M. N., Gras, S. L., Dewson, G., Miller, J. M. L., Finkbeiner, S., Moily, N. S., Niclis, J., Parish, C. L., Purcell, A. W., *et al.* (2017) Huntingtin inclusions trigger cellular quiescence, deactivate apoptosis, and lead to delayed necrosis. *Cell Rep.* **19,** 919–927 CrossRef Medline
46. Jockusch, B. M., Murk, K., and Rothkegel, M. (2007) The profile of profilins. *Rev. Physiol. Biochem. Pharmacol.* **159,** 131–149 Medline
47. Goldschmidt-Clermont, P. J., Machesky, L. M., Baldassare, J. J., and Pollard, T. D. (1990) The actin-binding protein profilin binds to PIP2 and inhibits its hydrolysis by phospholipase C. *Science* **247,** 1575–1578 CrossRef Medline
48. Goldschmidt-Clermont, P. J., Machesky, L. M., Doberstein, S. K., and Pollard, T. D. (1991) Mechanism of the interaction of human platelet profilin with actin. *J. Cell Biol.* **113,** 1081–1089 CrossRef Medline
49. Petrella, E. C., Machesky, L. M., Kaiser, D. A., and Pollard, T. D. (1996) Structural requirements and thermodynamics of the interaction of proline peptides with profilin. *Biochemistry* **35,** 16535–16543 CrossRef Medline
50. Witke, W., Sutherland, J. D., Sharpe, A., Arai, M., and Kwiatkowski, D. J. (2001) Profilin I is essential for cell survival and cell division in early mouse development. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 3832–3836 CrossRef Medline
51. Shao, J., Welch, W. J., Diprospero, N. A., and Diamond, M. I. (2008) Phosphorylation of profilin by ROCK1 regulates polyglutamine aggregation. *Mol. Cell. Biol.* **28,** 5196–5208 CrossRef Medline
52. Li, M., Huang, Y., Ma, A. A. K., Lin, E., and Diamond, M. I. (2009) Y-27632 improves rotarod performance and reduces huntingtin levels in R6/2 mice. *Neurobiol. Dis.* **36,** 413–420 CrossRef Medline
53. Fuentealba, R. A., Marasa, J., Diamond, M. I., Piwnica-Worms, D., and Weihl, C. C. (2012) An aggregation sensing reporter identifies leflunomide and teriflunomide as polyglutamine aggregate inhibitors. *Hum. Mol. Genet.* **21,** 664–680 CrossRef Medline
54. Fuentealba, R. A., Udan, M., Bell, S., Wegorzewska, I., Shao, J., Diamond, M. I., Weihl, C. C., and Baloh, R. H. (2010) Interaction with polyglutamine aggregates reveals a Q/N-rich domain in TDP-43. *J. Biol. Chem.* **285,** 26304–26314 CrossRef Medline
55. Pollitt, S. K., Pallos, J., Shao, J., Desai, U. A., Ma, A. A. K., Thompson, L. M., Marsh, J. L., and Diamond, M. I. (2003) A rapid cellular FRET assay of polyglutamine aggregation identifies a novel inhibitor. *Neuron* **40,** 685–694 CrossRef Medline
56. Veldman, M. B., Rios-Galdamez, Y., Lu, X. H., Gu, X., Qin, W., Li, S., Yang, X. W., and Lin, S. (2015) The N17 domain mitigates nuclear toxicity in a novel zebrafish Huntington's disease model. *Mol. Neurodegener.* **10,** 67 CrossRef Medline
57. Gu, X., Cantle, J. P., Greiner, E. R., Lee, C. Y., Barth, A. M., Gao, F., Park, C. S., Zhang, Z., Sandoval-Miller, S., Zhang, R. L., Diamond, M., Mody, I., Coppola, G., and Yang, X. W. (2015) N17 Modifies mutant Huntingtin nuclear pathogenesis and severity of disease in HD BAC transgenic mice. *Neuron* **85,** 726–741 CrossRef Medline
58. Andrew, S. E., Goldberg, Y. P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., and Kalchman, M. A. (1993) The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* **4,** 398–403 CrossRef Medline
59. Garai, K., Sahoo, B., Sengupta, P., and Maiti, S. (2008) Quasihomogeneous nucleation of amyloid β yields numerical bounds for the critical radius, the surface tension, and the free energy barrier for nucleus formation. *J. Chem. Phys.* **128,** 045102 CrossRef Medline
60. Sedlák, M. N., and Amis, E. J. (1992) Concentration and molecular weight regime diagram of salt-free polyelectrolyte solutions as studied by light scattering. *J. Chem. Phys.* **96,** 826 CrossRef
61. Bugg, C. W., Isas, J. M., Fischer, T., Patterson, P. H., and Langen, R. (2012) Structural features and domain organization of huntingtin fibrils. *J. Biol. Chem.* **287,** 31739–31746 CrossRef Medline
62. Freire, E., Schön, A., and Velazquez-Campoy, A. (2009) Isothermal titration calorimetry: general formalism using binding polynomials. *Methods Enzymol.* **455,** 127–155 CrossRef Medline
63. Mahoney, N. M., Janmey, P. A., and Almo, S. C. (1997) Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. *Nat. Struct. Biol.* **4,** 953–960 CrossRef Medline
64. Mahoney, N. M., Rozwarski, D. A., Fedorov, E., Fedorov, A. A., and Almo, S. C. (1999) Profilin binds proline-rich ligands in two distinct amide backbone orientations. *Nat. Struct. Biol.* **6,** 666–671 CrossRef Medline
65. Kaiser, D. A., and Pollard, T. D. (1996) Characterization of actin and poly-L-proline binding sites of *Acanthamoeba* profilin with monoclonal antibodies and by mutagenesis. *J. Mol. Biol.* **256,** 89–107 CrossRef Medline
66. Rennella, E., Sekhar, A., and Kay, L. E. (2017) Self-assembly of human profilin-1 detected by Carr-Purcell-Meiboom-Gill nuclear magnetic resonance (CPMG NMR) spectroscopy. *Biochemistry* **56,** 692–703 CrossRef Medline

67. Plimpton, S. (1995) Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117,** 1–19 CrossRef
68. Shen, K., Calamini, B., Fauerbach, J. A., Ma, B., Shahmoradian, S. H., Serrano Lachapel, I. L., Chiu, W., Lo, D. C., and Frydman, J. (2016) Control of the structural landscape and neuronal proteotoxicity of mutant Huntingtin by domains flanking the polyQ tract. *Elife* **5,** e18065 Medline
69. Sahoo, B., Singer, D., Kodali, R., Zuchner, T., and Wetzel, R. (2014) Aggregation behavior of chemically synthesized, full-length huntingtin exon1. *Biochemistry* **53,** 3897–3907 CrossRef Medline
70. Krishnamurthy, V. M., Estroff, L. A., and Whitesides, G. M. (2006) Multivalency in ligand design. In *Fragment-based Approaches in Drug Discovery* (Jahnke, W., and Erlanson, D. A., eds), pp. 11–53, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany
71. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed. Engl.* **37,** 2754–2794 CrossRef
72. Jencks, W. P. (1981) On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U.S.A.* **78,** 4046–4050 CrossRef Medline
73. Kayatekin, C., Matlack, K. E., Hesse, W. R., Guan, Y., Chakrabortee, S., Russ, J., Wanker, E. E., Shah, J. V., and Lindquist, S. (2014) Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 12085–12090 CrossRef Medline
74. Wear, M. P., Kryndushkin, D., O'Meally, R., Sonnenberg, J. L., Cole, R. N., and Shewmaker, F. P. (2015) Proteins with intrinsically disordered domains are preferentially recruited to polyglutamine aggregates. *PLoS One* **10,** e0136362 CrossRef Medline
75. Ripaud, L., Chumakova, V., Antonin, M., Hastie, A. R., Pinkert, S., Körner, R., Ruff, K. M., Pappu, R. V., Hornburg, D., Mann, M., Hartl, F. U., and Hipp, M. S. (2014) Overexpression of Q-rich prion-like proteins suppresses polyQ cytotoxicity and alters the polyQ interactome. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 18219–18224 CrossRef Medline
76. Warner, J. B., 4th, Ruff, K. M., Tan, P. S., Lemke, E. A., Pappu, R. V., and Lashuel, H. A. (2017) Monomeric huntingtin exon 1 has similar overall structural features for wild type and pathological polyglutamine lengths. *J. Am. Chem. Soc.* **139,** 14456–14469 CrossRef Medline
77. Korupolu, R. V., Achary, M. S., Aneesa, F., Sathish, K., Wasia, R., Sairam, M., Nagarajaram, H. A., and Singh, S. S. (2009) Profilin oligomerization and its effect on poly (l-proline) binding and phosphorylation, *Int. J. Biol. Macromol.* **45,** 265–273 CrossRef Medline
78. Lundgren, A., Agnarsson, B., Zirbs, R., Zhdanov, V. P., Reimhult, E., and Höök, F. (2016) Nonspecific colloidal-type interaction explains size-dependent specific binding of membrane-targeted nanoparticles. *ACS Nano* **10,** 9974–9982 CrossRef Medline
79. Liu, S. P., Zhou, L., Lakshminarayanan, R., and Beuerman, R. W. (2010) Multivalent antimicrobial peptides as therapeutics: design principles and structural diversities. *Int. J. Pept. Res. Ther.* **16,** 199–213 CrossRef Medline
80. Kim, Y., Lee, J. H., Ryu, J., and Kim, D. J. (2009) Multivalent and multifunctional ligands to β-amyloid. *Curr. Pharm. Des.* **15,** 637–658 CrossRef Medline
81. Riechers, S. P., Butland, S., Deng, Y., Skotte, N., Ehrnhoefer, D. E., Russ, J., Laine, J., Laroche, M., Pouladi, M. A., Wanker, E. E., Hayden, M. R., and Graham, R. K. (2016) Interactome network analysis identifies multiple caspase-6 interactors involved in the pathogenesis of HD. *Hum. Mol. Genet.* **25,** 1600–1618 CrossRef Medline
82. Kim, Y. E., Hosp, F., Frottin, F., Ge, H., Mann, M., Hayer-Hartl, M., and Hartl, F. U. (2016) Soluble oligomers of polyQ-expanded huntingtin target a multiplicity of key cellular factors. *Mol. Cell* **63,** 951–964 CrossRef Medline
83. O'Brien, R., DeGiacomo, F., Holcomb, J., Bonner, A., Ring, K. L., Zhang, N., Zafar, K., Weiss, A., Lager, B., Schilling, B., Gibson, B. W., Chen, S., Kwak, S., and Ellerby, L. M. (2015) Integration-independent transgenic Huntington disease fragment mouse models reveal distinct phenotypes and life span *in vivo*. *J. Biol. Chem.* **290,** 19287–19306 CrossRef Medline
84. Kalathur, R. K., Giner-Lamia, J., Machado, S., Barata, T., Ayasolla, K. R., and Futschik, M. E. (2015) The unfolded protein response and its potential role in Huntington's disease elucidated by a systems biology approach. *F1000Research* **4,** 103 CrossRef Medline
85. Hosp, F., Vossfeldt, H., Heinig, M., Vasiljevic, D., Arumughan, A., Wyler, E., Genetic and Environmental Risk for Alzheimer's Disease GERAD1 Consortium, Landthaler, M., Hubner, N., Wanker, E. E., Lannfelt, L., Ingelsson, M., Lalowski, M., Voigt, A., and Selbach, M. (2015) Quantitative interaction proteomics of neurodegenerative disease proteins. *Cell Rep.* **11,** 1134–1146 CrossRef Medline
86. Brehme, M., Voisine, C., Rolland, T., Wachi, S., Soper, J. H., Zhu, Y., Orton, K., Villella, A., Garza, D., Vidal, M., Ge, H., and Morimoto, R. I. (2014) A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell Rep.* **9,** 1135–1150 CrossRef Medline
87. Shirasaki, D. I., Greiner, E. R., Al-Ramahi, I., Gray, M., Boontheung, P., Geschwind, D. H., Botas, J., Coppola, G., Horvath, S., Loo, J. A., and Yang, X. W. (2012) Network organization of the huntingtin proteomic interactome in mammalian brain. *Neuron* **75,** 41–57 CrossRef Medline
88. Ratovitski, T., Chighladze, E., Arbez, N., Boronina, T., Herbrich, S., Cole, R. N., and Ross, C. A. (2012) Huntingtin protein interactions altered by polyglutamine expansion as determined by quantitative proteomic analysis. *Cell Cycle* **11,** 2006–2021 CrossRef Medline
89. Berchtold, L. A., Størling, Z. M., Ortis, F., Lage, K., Bang-Berthelsen, C., Bergholdt, R., Hald, J., Brorsson, C. A., Eizirik, D. L., Pociot, F., Brunak, S., and Størling, J. (2011) Huntingtin-interacting protein 14 is a type 1 diabetes candidate protein regulating insulin secretion and β-cell apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* **108,** E681–E688 CrossRef Medline
90. Langfelder, P., Cantle, J. P., Chatzopoulou, D., Wang, N., Gao, F., Al-Ramahi, I., Lu, X. H., Ramos, E. M., El-Zein, K., Zhao, Y., Deverasetty, S., Tebbe, A., Schaab, C., Lavery, D. J., Howland, D., *et al.* (2016) Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice. *Nat. Neurosci.* **19,** 623–633 CrossRef Medline
91. Li, M., Yasumura, D., Ma, A. A. K., Matthes, M. T., Yang, H., Nielson, G., Huang, Y., Szoka, F. C., Lavail, M. M., and Diamond, M. I. (2013) Intravitreal administration of HA-1077, a ROCK inhibitor, improves retinal function in a mouse model of Huntington disease. *PLoS One* **8,** e56026 CrossRef Medline
92. Davranche, A., Aviolat, H., Zeder-Lutz, G., Busso, D., Altschuh, D., Trottier, Y., and Klein, F. A. (2011) Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. *Hum. Mol. Genet.* **20,** 2795–2806 CrossRef Medline
93. Sittler, A., Wälter, S., Wedemeyer, N., Hasenbank, R., Scherzinger, E., Eickhoff, H., Bates, G. P., Lehrach, H., and Wanker, E. E. (1998) SH3GL3 associates with the Huntingtin exon 1 protein and promotes the formation of polygln-containing protein aggregates. *Mol. Cell* **2,** 427–436 CrossRef Medline
94. Beck, M., Schmidt, A., Malmstroem, J., Claassen, M., Ori, A., Szymborska, A., Herzog, F., Rinner, O., Ellenberg, J., and Aebersold, R. (2011) The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7,** 549 Medline
95. Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) The structure and function of proline recognition domains. *Sci. STKE* **2003,** RE8 Medline
96. Zhang, W., Ben-David, M., and Sidhu, S. S. (2017) Engineering cell signaling modulators from native protein-protein interactions. *Curr. Opin. Struct. Biol.* **45,** 25–35 Medline
97. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150,** 76–85 CrossRef Medline
98. Radhakrishnan, A., Vitalis, A., Mao, A. H., Steffen, A. T., and Pappu, R. V. (2012) Improved atomistic Monte Carlo simulations demonstrate that poly-L-proline adopts heterogeneous ensembles of conformations of semi-rigid segments interrupted by kinks. *J. Phys. Chem. B* **116,** 6862–6871 CrossRef Medline
99. Vitalis, A., and Pappu, R. V. (2009) ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **30,** 673–699 CrossRef Medline

Profilin reduces aggregation and phase separation of huntingtin N-terminal fragments by preferentially binding to soluble monomers and oligomers

Ammon E. Posey, Kiersten M. Ruff, Tyler S. Harmon, Scott L. Crick, Aimin Li, Marc I. Diamond and Rohit V. Pappu

J. Biol. Chem. 2018, 293:3734-3746.
doi: 10.1074/jbc.RA117.000357 originally published online January 22, 2018

Profilin reduces aggregation and phase separation of huntingtin N-terminal fragments by preferentially binding to soluble monomers and oligomers

Ammon E. Posey[1], Kiersten M. Ruff[1], Tyler S. Harmon[1], Scott L. Crick[1], Aimin Li[3], Marc I. Diamond[2], Rohit V. Pappu[1]*

[1]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA
[2]Center for Alzheimer's and Neurodegenerative Diseases, University of Texas, Southwestern Medical Center, Dallas, TX 75390, USA
[3]Department of Neurology, Washington University School of Medicine, St. Louis, MO 63108, USA

*Corresponding author: R.V.P (pappu@wustl.edu)

SUPPORTING INFORMATION APPENDIX

Contents:

SUPPORTING METHODS

Peptide disaggregation: An appropriate amount of peptide was weighed in a Wheaton glass vial and dissolved in a 1:1 mixture of trifluoroacetic acid (TFA):hexafluoroisopropanol (HFIP), such that the final concentration was roughly 1 mg/mL peptide. The TFA-HFIP protocol was previously shown to be reasonable for polyQ containing peptides that lack the N17 module (1,2). This solution was sonicated for ~45 seconds and left to incubate at room temperature for 1 hour. The solution was then thoroughly evaporated under a gentle nitrogen stream until a clear peptide film remained on the walls of the glass vial. The peptides were resuspended from this film in a small volume of buffer (20 mM Tris, 5 mM EDTA, 1 mM DTT) at room temperature with sonication. The solution was then diluted to the desired concentration by the addition of buffer and the pH was re-adjusted to 7.4. DTT was included in the buffer to reduce cysteine bonds in profilin, but to be consistent DTT was also used in experiments that did not include profilin. For experiments performed in the presence of profilin, an appropriate amount of a concentrated profilin stock solution was added to the peptide after the peptide was resuspended by sonication as described above. All chemicals were purchased from Sigma unless otherwise noted.

Steady-state tryptophan fluorescence data analysis: Using nonlinear least-squares analysis, the data were fit to the equation $F = \frac{F_{max}[L]}{(K_d + [L])}$ where F_{max} is the fluorescence maximum recovered from fitting the raw data and [*L*] is the concentration of Q_{40}-C38. The data were then converted to fraction bound, where fraction bound is defined as: F/F_{max}. The equation for estimating F is a simplified form that is based on the assumption that the macromolecule (profilin) concentration is not considerably larger than K_d, which is true in this case: even the highest profilin concentration used (20 μM), is not significantly larger than the apparent K_d which is ~10 μM. To verify that changing the profilin concentration did not have an effect on fitting with the simplified equation, each isotherm was also fit with the full quadratic equation

$F = F_{max} \frac{\left([M]+[L]+K_d\right) - \sqrt{\left([M]+[L]+K_d\right)^2 - 4[M][L]}}{2[M]}$, which includes the macromolecule concentration ([M]) in the fit. This analysis yields values for the apparent dissociation constant that are essentially equivalent to those obtained using the simplified fit in each instance.

Right-angle static light scattering data analysis: Optimal linear fits were determined using a modified jackknife approach in the following manner: each dataset was subjected to two independent series of linear fits, with one series of fits for the low concentration arm (hereafter referred to as LCA) and the second series of fits for the high concentration arm (hereafter referred to as HCA). For each arm (the LCA and the HCA), a series of fits was initiated with a linear fit to the four lowest or highest concentration data points, respectively, and the root mean square error (RMSE) of each fit was recorded. Following these initial fits, the next highest or lowest concentration data point was added to the respective LCA or HCA data set, the expanded data sets were re-fit, and the new RMSEs were recorded. This process was continued, expanding the fitted dataset by one data point at a time, until all the points in the full dataset were included in both the LCA and HCA linear fits. Fits that included data points from the opposite side of the putative discontinuity point caused the RMSE to deviate significantly and could be observed on a plot of RMSE vs. number of points fit. In our case, either the mean or the median RMSE of the fits for a given arm of the data consistently correlated with the point at which the RMSE began to deviate significantly. Consequently, the mean or median RMSE was used as a cutoff; all fits with an RMSE above the cutoff were discarded, and all fits with an RMSE below the cutoff were deemed “best fits”. The intersections of each pair of LCA and HCA best fits were recorded and the average intersection point for all best fits, for all trials at a given concentration of profilin, were determined. This is the value reported as c_S herein.

Thioflavin T fluorescence measurements: Peptides were dissolved and disaggregated in formic acid and the solution was adjusted to 20 mM Tris buffer at pH 7.4, with 5 mM EDTA and 1 mM DTT. Note that identical results were also obtained in 50 mM phosphate buffer. Profilin from a concentrated stock solution and / or buffer was added in the appropriate amounts to achieve the specified final profilin concentrations while maintaining equivalent dilution of peptide across all samples. The final concentration of ThT was 20 μM. Samples were prepared in triplicate in 96-well plates that were subsequently covered with a transparent film to prevent sample evaporation. Measurements were carried out at 23°C in a Tecan plate reader with 5 seconds of orbital shaking once per minute. Samples were excited at 436 nm (20 nm bandwidth) and emission was collected at 482 nm (28 nm bandwidth).

TEM image analysis: For size analysis of M- and S-phase particles, TEM images of Q40C38 samples at two concentrations (126 nM and 295 nM) were analyzed using a custom algorithm implemented using ImageJ and Python. 22,384 M-phase particles and 342 S-phase particles were sampled and histogrammed. Mean particle sizes (diameters) were determined by fitting each histogram to a Gaussian (126 nM sample) or sum of two Gaussian curves (295 nM sample). For the analysis of the binding of profilin to the various phases (with profilin identified using 5 nm Ni-gold nanoparticles bound to the profilin His-tag), the custom algorithm described above was modified to create false-colored images wherein M-, S- and F-phase aggregates were color-coded. A 19 nm diameter cutoff was used to distinguish between M- and S-phases and an aspect ratio cutoff of 2.0 was used to distinguish fibrils from non-fibrillar aggregates. Five TEM images of aggregation assay samples with 16.4 μM Q40C38 and 5 μM profilin were color-coded in this way, then the number of gold nanoparticles associated with M-, S- and F-phases were manually counted/classified.

All Atom Simulations: All atom simulations used to extract coarse-grained parameters were performed using the CAMPARI simulation package (http://campari.sourceforge.net) with the ABSINTH implicit solvation model and forcefield paradigm (3,4). The details of the ABSINTH implicit solvation model and move sets used in the CAMPARI modeling package have been published previously (3,5). Simulations were based on the abs3.2_opls.prm parameter set. Temperature replica exchange was used to enhance sampling using a temperature schedule of *T*=[288 K, 293 K, 298 K, 305 K, 310 K, 315 K, 320 K, 325 K, 335 K, 345 K, 360 K, 375 K, 390 K, 405 K]. Each simulation consisted of 6.15×10^7 steps. Steps can be either Metropolis Monte Carlo moves or temperature swaps. The first 10^7 steps are taken as equilibration and all analyses are performed over the last 5.15×10^7 steps. Temperature swaps were proposed every 5×10^4 steps. Simulations were run in spherical droplets of radius 125 Å. This radius sized was picked in order to ensure against confinement artifacts. The specific sequences used for coarse-grained parameterization were Ace-Q_{22}-Nme, Ace-Q_{40}-Nme, and Ace-P_{11}-QLPQPPPQAQPLLPQPQ-P_{10}-Nme. Here, the subscripts denote how many times that amino acid was repeated, Ace is the N-terminal acetyl unit, and Nme is the C-terminal N-methyl amide.

Coarse-Grained Langevin Dynamics Simulations: Langevin dynamics (LD) simulations were used to study Q_{40}-C38:Pfn interactions. All LD simulations utilized the LAMMPS simulation package (http://lammps.sandia.gov) (6). The force on each bead *i* is given by

$$\mathbf{F}_i = -\nabla W_{\text{eff},i} - \frac{m_i}{\gamma_i}\mathbf{v}_i + \mathbf{R}_i. \quad (1)$$

Here, W_{eff} is the effective energy function described below (Equation (2)). The second term denotes the frictional force which is proportional to the velocity of bead *i*, $\mathbf{v}_i$. The damping term, γ_i, is calculated using $\gamma_i = C\frac{m_i}{6\pi\eta R_i}$. Here, m_i is the mass of bead *i*, $\eta=6.29\times10^{-4}$ kg m^{-1} s^{-1} is the viscosity of water at 315 K, R_i is the radius of bead *i*, and *C*=20 is a scaling factor. The last

term in Equation (1) denotes the random force, derived from the fluctuation dissipation theorem, exerted on bead *i* from collisions with the bath. The equation of motion is integrated using a velocity Verlet algorithm with an integration time step of 2 fs.

Coarse-grained simulations of C38:Pfn and Q_{40}-C38:Pfn interactions were performed in the canonical ensemble with and without the presence of an auxiliary Pfn interaction. Simulations were conducted in cubic boxes of length 703.9 Å with periodic boundary conditions. Each simulation included 210 Pfn molecules and 630 C38 stretches. Each Q_{40}-C38:Pfn simulation had M Q_{40}-C38 clusters of size X such that $M \times X = 630$. Here, X=1, 3, 6, 9, 15, 35, 70, 210. By restricting each simulation to only have clusters of a particular size, we bypass problems associated with the kinetics of aggregation in coarse-grained simulations. Simulations were initiated by replicating Pfn molecules and Q_{40}-C38 clusters on three-dimensional lattices. Following initiation, energy minimization was performed utilizing the Polak-Ribiere conjugate gradient algorithm (6). Then, 10^8 steps of LD simulations are performed using a time step of 2 fs. For each combination of Q_{40}-C38 cluster size and auxiliary interaction 5 independent simulations were performed.

Coarse-Grained Model Architecture: Q_{40}-C38 clusters of size X were constructed using a single scaled bead for the aggregated polyQ domains with X C38 stretches protruding from the surface of the polyQ bead. Specifically, the radius of the polyQ bead was determined using the equation $R_g = R_0 N^v$. Here, R_g is the radius of gyration, N is the number of glutamine residues in the cluster, v=0.33 for globules, and R_0=2.85 is the pre-factor. The pre-factor R_0 was determined by fitting the $\langle R_g \rangle$ from all atom simulations of Q_{22} and Q_{40} to the equation $\langle R_g \rangle = R_0 N^{0.33}$.

The coarse-grained architecture of the C38 stretches is p-p_{Pfn}-p-l-l-l-l-l-p-p_{Pfn}-p. Here, p denotes an excluded volume polyP bead, p_{Pfn} denotes a polyP bead that interacts with Pfn through its primary polyP binding site, and l denotes a linker bead. Given this architecture, each C38 bead corresponds to 3-4 amino acids. The specific sequence-to-bead mapping is given by

(PPPP)-(PPP)-(PPPP)-(QLPQ)-(PPP)-(QAQ)-(PLL)-(PQPQ)-(PPP)-(PPP)-(PPPP), where each set of amino acids within parentheses corresponds to a bead. The radius of each C38 bead was determined by calculating the average R_g over the residues that make up the bead. The mass of each bead was taken to be the sum of the mass of the amino acids that make up the bead.

The base coarse-grained architecture for Pfn is an excluded volume bead (Pfn_{EV}) and a polyP interacting bead (Pfn_p). The radius of the Pfn_{EV} bead was taken to be the R_g extracted from the crystal structure of human Pfn 1 (PDB: 1pfl). The radius of the Pfn_p bead was taken to be that of the C38 p_{Pfn} bead. The mass of the Pfn_p bead was taken to be that of the C38 p_{Pfn} bead (291.39 g/mol) and the rest of the Pfn mass was distributed to the Pfn_{EV} bead. In order to model the linker:Pfn and Pfn:Pfn auxiliary interactions an additional coarse-grained Pfn bead was added to the Pfn model. The details of these additions are discussed below.

Coarse-Grained Model Energy Function: The coarse-grained energy function takes the form:

$$W_{\text{eff}} = W_b + W_\theta + W_{LJ} + W_{\text{Primary}} + W_{\text{Auxiliary}}. \quad (2)$$

Here, W_b, W_θ, W_{LJ}, W_{Primary}, and $W_{\text{Auxiliary}}$ correspond to the bond length, bond angle, Lennard-Jones, primary, and auxiliary interaction potentials. The primary interaction potential refers to the interaction between Pfn and polyP. The auxiliary interaction potential refers to the interaction between Pfn and either polyQ, linker, or Pfn.

The bonded potentials in Equation (2) are given by

$$W_b = \sum_{i=1}^{N_b} \frac{K_i\left(b_i - b_{0i}\right)^2}{2},$$

$$W_\theta = \sum_{i=1}^{N_\theta} \frac{L_i\left(\theta_i - \theta_{0i}\right)^2}{2}.$$

Here, N_b and N_θ are the total number of bonds and angles in the system and K_i and L_i are the force constants associated with the equilibrium bond and angle values given by b_{0i} and θ_{0i}, respectively. The Boltzmann inversion procedure (7) was used to extract K_i, L_i, b_{0i}, and θ_{0i} for

C38 beads from all atom simulations of C38 at 315 K. Given that the bonded potentials are described by harmonic potentials, inversion of the Boltzmann relationship yields the following analytical relationships for K_i, L_i, b_{0i}, and θ_{0i}

$$b_{0i} = \langle b_i \rangle,\ K_i = \frac{k_B T}{\langle b_i^2 \rangle - \langle b_i \rangle^2},$$
$$\theta_{0i} = \langle \theta_i \rangle,\ \mathrm{L}_i = \frac{k_B T}{\langle \theta_i^2 \rangle - \langle \theta_i \rangle^2}.$$

Here, b_i and θ_i are the sets of bond lengths and bond angles extracted from the all atom simulations. For the bond between the polyQ bead and the first C38 bead, b_{0i} and K_i were set to the radius of the polyQ bead and 10 kcal/mol/Å^2, respectively. For the bond between the two Pfn beads, b_{0i} and K_i were set to the sum of the radii of the two Pfn beads and 10 kcal/mol/Å^2, respectively. Finally, θ_{0i} and L_i were set to 180 degrees and 10 kcal/mol/radians^2, respectively, for the angle between the polyQ bead and the first two C38 beads.

The W_{LJ} term, denotes the Lennard-Jones potential calculated over non-bonded pairs of beads with the form

$$\begin{aligned} W_{LJ} &= \sum_{i=1}^{N_{nb}} \sum_{j<i} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right],\ r_{ij} < r_c \\ &= 0,\ r_{ij} \geq r_c. \end{aligned}$$

Here, N_{nb} is the number of beads, r_{ij} is the distance between beads i and j, σ_{ij} is the distance at which the inter-bead potential is zero, and ε_{ij} is the strength of the interaction. Except for the primary and auxiliary interactions between Q_{40}-C38 clusters and Pfn molecules described in detail below, all non-bonded inter-bead interactions were defined by $\sigma_{ij} = \left(2R_g^i + 2R_g^j\right)/2$ Å,

where R_g^i is the R_g extracted from all atom simulations or the Pfn crystal structure for bead *i* as explained above, ε_{ij}=0.01 kcal/mol, and $r_c = 2.5\sigma_{ij}$ Å.

The primary interaction potential models the interaction between Pfn and polyP. In order to model this interaction, we combined inverted Gaussian (w_G) and Lennard-Jones (w_{LJ}) potentials. Explicitly,

$$\begin{aligned} W_{\text{Primary}} &= w_G + w_{LJ}, \\ w_G &= \sum_{i=1}^{N_{p_{\text{Pfn}}}} \sum_{j=1}^{N_{\text{Pfn}_p}} -\varepsilon_G \exp\left[-\frac{(r_{ij}-r_G)^2}{2\sigma_G^2}\right], \; r_{ij} < r_G^{\text{cutoff}} \\ &= 0, \; r_{ij} \geq r_G^{\text{cutoff}}, \\ w_{LJ} &= \sum_{i=1}^{N_{p,p_{\text{Pfn}}}} \sum_{j=1}^{N_{\text{Pfn}_{\text{EV}}}} 4\varepsilon_{LJ}\left[\left(\frac{\sigma_{LJ}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{LJ}}{r_{ij}}\right)^{6}\right], \; r_{ij} < r_{LJ}^{\text{cutoff}} \\ &= 0, \; r_{ij} \geq r_{LJ}^{\text{cutoff}}. \end{aligned}$$

Here, $N_{p_{\text{Pfn}}}$, N_{Pfn_p}, $N_{p,p_{\text{Pfn}}}$, and $N_{\text{Pfn}_{\text{EV}}}$ correspond to the number of p_{Pfn} beads, the number of Pfn_p beads, the number of *p* and p_{Pfn} beads, and the number of Pfn_{EV} beads, respectively. For w_G, r_{ij} is the distance between beads *i* and *j*, ε_G = 6 kcal/mol is the interaction strength, r_G = 0 Å is the position of the well, σ_G = 2.27 Å controls the width of the well, and $r_G^{\text{cutoff}} = 2.5R^{p_{\text{Pfn}}}$, where $R^{p_{\text{Pfn}}}$ is the radius of the C38 p_{Pfn} bead. The width of the well was chosen such that the energy for two p_{Pfn} beads to bind the Pfn_p bead on a single Pfn molecule was half that of the energy for a single p_{Pfn} bead to bind the Pfn_p bead. This choice was made in order to disfavor a single Pfn_p from binding multiple p_{Pfn} beads. Explicitly, in the dimer case $r_{p_{\text{Pfn}},\text{Pfn}_p} = R_g$ since $\sigma_{p_{\text{Pfn}},p_{\text{Pfn}}} = 2R_g$, where R_g is the radius of the p_{Pfn} bead. This yields

$$\frac{1}{2}U_M = U_D,$$
$$-\frac{1}{2}\varepsilon_G \exp\left(-\frac{0^2}{2\sigma_G^2}\right) = -2\varepsilon_G \exp\left(-\frac{R_g^{\,2}}{2\sigma_G^2}\right). \quad \textbf{(3)}$$

Rearranging Equation (3), yields

$$\sigma_G = \sqrt{\frac{-R_g^2}{2\ln\left(\frac{1}{4}\right)}} = 2.27\text{Å}$$

For w_{LJ}, r_{ij} is the distance between beads *i* and *j*, ε_{LJ} = 0.01 kcal/mol is the interaction strength, σ_{LJ} is the distance at which the inter-bead potential is zero, and $r_{LJ}^{\text{cutoff}} = 2.5\sigma_{LJ}$. Here, σ_{LJ} was set to the radius of the Pfn_{EV} bead. This choice prevented excluded volume effects between Pfn_{EV} beads and *p* and p_{Pfn} beads from disfavoring p_{Pfn}:Pfn_p interactions.

The auxiliary interaction potential takes on three different forms for the three different models tested to account for the experimental observations. The three models are a polyQ:Pfn interaction, a linker:Pfn interaction, and a Pfn:Pfn interaction. The details of each model will be discussed below.

The auxiliary interaction potential for the polyQ:Pfn interaction is modeled as a Lennard-Jones potential between the polyQ bead and the Pfn_{EV} bead. Long polyQ stretches and polyQ aggregates have been shown to interact non-specifically with other molecules (8-10). Thus, we assumed that the polyQ cluster should be uniformly and non-specifically interactive with Pfn. Explicitly, the model takes the form

$$\begin{aligned} W_{\text{Auxiliary}}^{\text{polyQ:Pfn}} &= \sum_{i=1}^{N_{\text{polyQ}}}\sum_{j=1}^{N_{\text{Pfn}_{\text{EV}}}} 4\varepsilon_{\text{polyQ:Pfn}}\left[\left(\frac{\sigma_{\text{polyQ:Pfn}}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{\text{polyQ:Pfn}}}{r_{ij}}\right)^{6}\right],\ r_{ij} < r_{\text{polyQ:Pfn}}^{\text{cutoff}} \\ &= 0,\ r_{ij} \geq r_{\text{polyQ:Pfn}}^{\text{cutoff}}. \end{aligned} \quad (4)$$

In Equation (4), N_{polyQ} and $N_{\text{Pfn}_{\text{EV}}}$ denote the number of polyQ and Pfn_{EV} beads, respectively, r_{ij} is the distance between beads *i* and *j*, $\varepsilon_{\text{polyQ:Pfn}}$ is the strength of the interaction, which was set to either 1 or 2 kcal/mol, and $\sigma_{\text{polyQ:Pfn}} = R^{\text{polyQ}}$ Å is the distance at which the inter-bead potential is zero. Here, R^{polyQ} is the radius of the polyQ bead. The cutoff, $r_{\text{polyQ:Pfn}}^{\text{cutoff}}$, was set to $2.5\sigma_{\text{polyQ:Pfn}}$.

For the linker:Pfn auxiliary interaction potential an additional bead was added to the Pfn architecture. This bead, termed $\text{Pfn}_{\text{linker}}$, was set to have the same mass and radius as the Pfn_p bead. The rest of the mass was distributed to the Pfn_{EV} bead. The linker:Pfn interaction potential is modeled as a Lennard-Jones potential between the $\text{Pfn}_{\text{linker}}$ bead and all non-polyP linker beads, with the addition of bond length and angle terms to account for the extra Pfn bead. Specifically,

$$
\begin{aligned}
W_{\text{Auxiliary}}^{\text{linker:Pfn}} &= w_b^{\text{linker:Pfn}} + w_\theta^{\text{linker:Pfn}} + w_{LJ}^{\text{linker:Pfn}}, \\
w_b^{\text{linker:Pfn}} &= \sum_{i=1}^{N_{\text{Pfn}}} \frac{K_i\left(b_i - b_{0i}\right)^2}{2}, \\
w_\theta^{\text{linker:Pfn}} &= \sum_{i=1}^{N_{\text{Pfn}}} \frac{L_i\left(\theta_i - \theta_{0i}\right)^2}{2}, \\
w_{LJ}^{\text{linker:Pfn}} &= \sum_{i=1}^{N_{\text{linker}}} \sum_{j=1}^{N_{\text{Pfn}_{\text{linker}}}} 4\varepsilon_{\text{linker:Pfn}} \left[\left(\frac{\sigma_{\text{linker:Pfn}}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{\text{linker:Pfn}}}{r_{ij}} \right)^{6} \right], \; r_{ij} < r_{\text{linker:Pfn}}^{\text{cutoff}} \\
&= 0, \; r_{ij} \geq r_{\text{linker:Pfn}}^{\text{cutoff}}.
\end{aligned}
\tag{5}
$$

Here, N_{Pfn}, N_{linker} and $N_{\text{Pfn}_{\text{linker}}}$ denote the number of Pfn molecules, non-polyP linker beads, and $\text{Pfn}_{\text{linker}}$ beads, respectively. For $w_{LJ}^{\text{linker:Pfn}}$, r_{ij} is the distance between beads *i* and *j*, $\varepsilon_{\text{linker:Pfn}}$ is the strength of the interaction, which was set to either 1, 1.5, or 2 kcal/mol, and $\sigma_{\text{linker:Pfn}} = \left(2R^{\text{Pfn}_{\text{linker}}} + 2R^{l}\right)/2$ Å is the distance at which the inter-bead potential is zero. Here,

$R^{\mathrm{Pfn_{linker}}}$ and R^{l} are the radii for the $\mathrm{Pfn_{linker}}$ and *l* beads, respectively. The cutoff, $r_{\mathrm{linker:Pfn}}^{\mathrm{cutoff}}$, was set to $2.5\sigma_{\mathrm{linker:Pfn}}$.

In Equation (5), $w_b^{\mathrm{linker:Pfn}}$ and $w_\theta^{\mathrm{linker:Pfn}}$ account for the bond length and bond angle potentials between $\mathrm{Pfn_{linker}}$:$\mathrm{Pfn_{EV}}$ beads and Pfn_p:$\mathrm{Pfn_{EV}}$:$\mathrm{Pfn_{linker}}$ beads, respectively. The equilibrium bond length (b_{0i}) and force constant (K_i) between $\mathrm{Pfn_{linker}}$ and $\mathrm{Pfn_{EV}}$ were set to the sum of the radii of the two Pfn beads and 10 kcal/mol/Å^2, respectively. The equilibrium angle (θ_{0i}) and force constant (L_i) between all three Pfn beads were set to 180 degrees and 10 kcal/mol/radians^2, respectively. This architecture was chosen to account for the experimental observation that the K_d for the C38:Pfn interaction is twice that of the K_d for the P_{11}:Pfn interaction. This result implies the linker does not engage in interactions with the same Pfn molecule that interacts with one of the two polyP segments within C38. Thus, in order to be consistent with experimental results, if the linker was to interact with Pfn, then it would have to interact on the opposite face of the polyP binding pocket. This architecture restricts the linker and a polyP segment of the same C38 from interacting with the same Pfn molecule. Additionally, the opposite face of the polyP binding pocket on Pfn is enriched in polar residues which may engage in interactions with the polar and hydrophobic residues of the linker (11).

For the Pfn:Pfn auxiliary interaction, stable Pfn dimers were created to bypass kinetic problems associated with Pfn:Pfn dimerization. To construct stable Pfn dimers an additional Pfn bead, $\mathrm{Pfn_{dimer}}$, was added to the architecture (see Figure S6). $\mathrm{Pfn_{dimer}}$ was set to have the same mass and radius as the Pfn_p bead. The rest of the mass was distributed to the $\mathrm{Pfn_{EV}}$ bead. The Pfn:Pfn auxiliary potential contains bond length, bond angle, and Lennard-Jones potentials in order to account for the stable dimer architecture. Specifically,

$$
\begin{aligned}
W_{\text{Auxiliary}}^{\text{Pfn:Pfn}} &= w_b^{\text{Pfn:Pfn}} + w_\theta^{\text{Pfn:Pfn}} + w_{LJ}^{\text{Pfn:Pfn}}, \\
w_b^{\text{Pfn:Pfn}} &= \sum_{i=1}^{N_b^{\text{Pfn:Pfn}}} \frac{K_i\left(b_i - b_{0i}\right)^2}{2}, \\
w_\theta^{\text{Pfn:Pfn}} &= \sum_{i=1}^{N_\theta^{\text{Pfn:Pfn}}} \frac{L_i\left(\theta_i - \theta_{0i}\right)^2}{2}, \\
w_{LJ}^{\text{Pfn:Pfn}} &= \sum_{i=1}^{N_{\text{Pfn}_p}} \sum_{j<i}^{N_{\text{Pfn}_p}} 4\varepsilon_{\text{Pfn:Pfn}} \left[\left(\frac{\sigma_{\text{Pfn:Pfn}}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{\text{Pfn:Pfn}}}{r_{ij}} \right)^{6} \right], \; r_{ij} < r_{\text{Pfn:Pfn}}^{\text{cutoff}} \\
&= 0, \; r_{ij} \geq r_{\text{Pfn:Pfn}}^{\text{cutoff}}.
\end{aligned}
\tag{6}
$$

Here, $N_b^{\text{Pfn:Pfn}}$ and $N_\theta^{\text{Pfn:Pfn}}$ are the number of new bond lengths and bond angles that need to be defined in order to create a stable dimer, respectively. The new bonds defined are $\text{Pfn}_{\text{EV}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m1}}$, $\text{Pfn}_{\text{EV}}^{\text{m2}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$, and $\text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$. Here, m1 and m2 correspond to molecule 1 and molecule 2, respectively. The equilibrium bond length (b_{0i}) and force constant (K_i) between $\text{Pfn}_{\text{EV}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m1}}$, $\text{Pfn}_{\text{EV}}^{\text{m2}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$, and $\text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$ were set to 6.47 Å and 10 kcal/mol/Å^2, respectively. The new angles defined are $\text{Pfn}_{\text{EV}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$ and $\text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}} : \text{Pfn}_{\text{EV}}^{\text{m2}}$. The equilibrium angle (θ_{0i}) and force constant (L_i) for $\text{Pfn}_{\text{EV}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}}$ and $\text{Pfn}_{\text{dimer}}^{\text{m1}} : \text{Pfn}_{\text{dimer}}^{\text{m2}} : \text{Pfn}_{\text{EV}}^{\text{m2}}$ were set to 180 degrees and 10 kcal/mol/radians^2, respectively. These choices allow for the distance between the two Pfn_p beads of a dimer to vary while the distance between the two Pfn_{EV} beads is held relatively fixed. We allow for variability in the distance between the two Pfn_p beads since the site of Pfn dimerization is not known. However, in order to restrict the two Pfn_p beads from coming too close we changed the Lennard-Jones potential parameters between two Pfn_p beads. The Lennard-Jones potential between two Pfn_p beads is defined by $w_{LJ}^{\text{Pfn:Pfn}}$ in Equation (6). Here, N_{Pfn_p} is the number of Pfn_p beads, r_{ij} is the distance between beads i and j, $\varepsilon_{\text{Pfn:Pfn}} = 0.01$

kcal/mol is the strength of the interaction, $\sigma_{\mathrm{Pfn,Pfn}} = 10$ Å is the distance at which the inter-bead potential is zero, and $r^{\mathrm{cutoff}}_{\mathrm{Pfn:Pfn}} = 2.5\sigma_{\mathrm{Pfn:Pfn}}$.

Calculation of Fraction Bound from Simulations: In order to determine which Q_{40}-C38 clusters Pfn preferentially bound, the fraction of Pfn_p beads bound to C38 p_{Pfn} beads was calculated. We only consider this interaction in our calculation given that this is the interaction the steady-state tryptophan fluorescence experiments measure. A Pfn_p bead was said to be bound to a C38 p_{Pfn} bead if the distance between the two beads was less than $2R^{p_{\mathrm{Pfn}}}$, where $R^{p_{\mathrm{Pfn}}}$ is the radius of the p_{Pfn} bead. The fraction bound for a given simulation was averaged over the last 38.2 ns of the simulation, i.e., the post-equilibration time.

SUPPORTING FIGURES AND TABLE

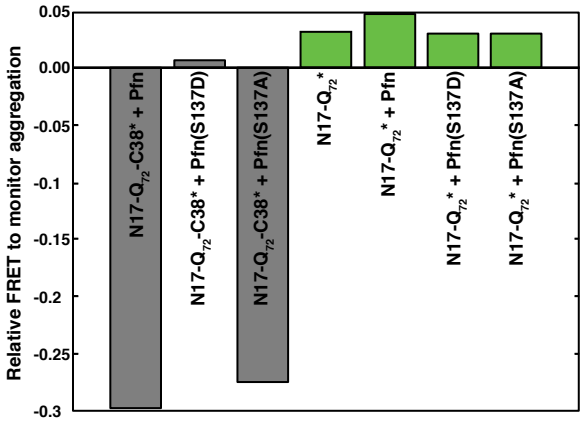

Figure S1: The C38 region is required for modulation of Htt-NTF aggregation by profilin. FRET-based intracellular aggregation assay. All constructs have either a CFP or YFP fused to the C-terminus as denoted by the asterisk. The bar heights quantify the relative FRET, which is defined as ($\Delta F/F_{ref}$) where $\Delta F=(F_X–F_{ref})$. Here, F_X is the intracellular FRET for the construct in question and F_{ref} is the FRET measured for N17-Q_{72}-C38-CFP/YFP in the absence of profilin. The green bars are data for Htt-NTF constructs without the C38 module. The more negative the bars, the stronger the suppressive effects of profilin. Wild type profilin reduced N17-Q_{72}-C38-CFP/YFP aggregation by 30%, as described previously (12). However, the aggregation of N17-Q_{72}-CFP/YFP was unaffected by profilin overexpression (first and second green bars). The ability to suppress intracellular aggregation by profilin was blocked by the phosphomimic mutation S137D and was preserved by the S137A mutation (second and third gray bars, respectively) whereas these mutations had no effects on constructs lacking C38 (third and fourth green bars). Pfn refers to profilin whereas Pfn(S137D) and Pfn(S137A) refer to variants of profilin with Ser 137 replaced by Asp or Ala to mimic phosphorylation or non-phosphorylation, respectively.

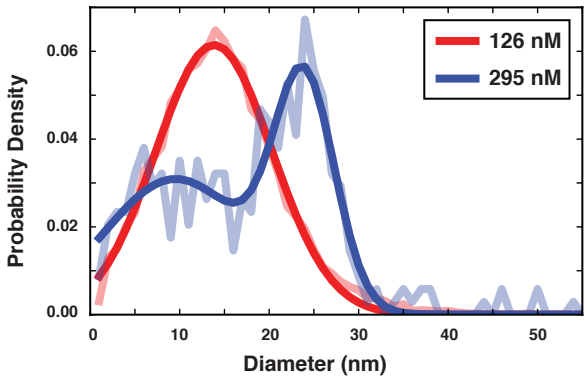

Figure S2: Aggregate size distributions determined by analysis of TEM images. Histograms of Q_{40}-C38 particle/aggregate sizes observed at 126 nM (M-phase, red) and at 295 nM (S-phase, blue), quantified using a custom image analysis algorithm. The raw histograms are shown as lighter shades of each color, while fits to Gaussian curves are shown as darker shades. The histograms represent the value of the probability density function at each bin (bin size = 1 nm), normalized such that the integral over the range is 1. 22,384 M-phase particles and 342 S-phase particles were sampled for the histograms and the sizes of particles at these two concentrations were found to be distinct, with a mean particle size of 13.83 +/- 0.06 for the single peak from the 126 nM sample (red), and 9.53 +/- 0.83 and 23.94 +/- 0.24 for the two peaks from the 295 nM sample (blue). Mean particle sizes were determined by fitting each histogram to a Gaussian (126 nM sample) or sum of two Gaussian curves (295 nM sample). The two peaks from the 295 nM sample likely represent M- and S-phase species, respectively. Quantification of the monomer and small oligomer sizes are likely a significant overestimate because monomers and some small oligomers approach the size of one or a few pixels and are therefore unavoidably filtered out as noise by the image analysis algorithm.

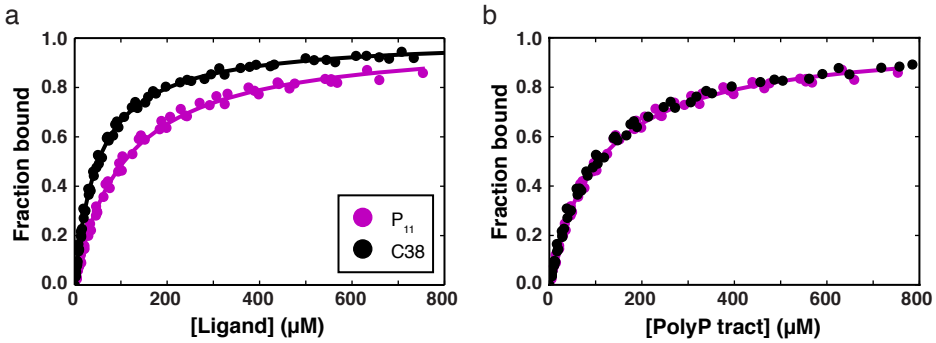

Figure S3. Profilin binds the polyP tracts in the C38 region of Htt-NTFs. (**a**) Binding isotherms of profilin with P_{11} (purple) or C38 (black) peptides. Fits to the data yield apparent K_d values of 106.5 μM and 50.7 μM, respectively. (**b**) When the peptide concentration is adjusted to reflect the concentration of polyP tracts (P_{11} is a single tract, whereas C38 contains two polyP tracts), then the binding isotherms overlay, indicating that the difference in apparent K_d is due solely to the number of binding sites. The profilin concentration was 5 μM.

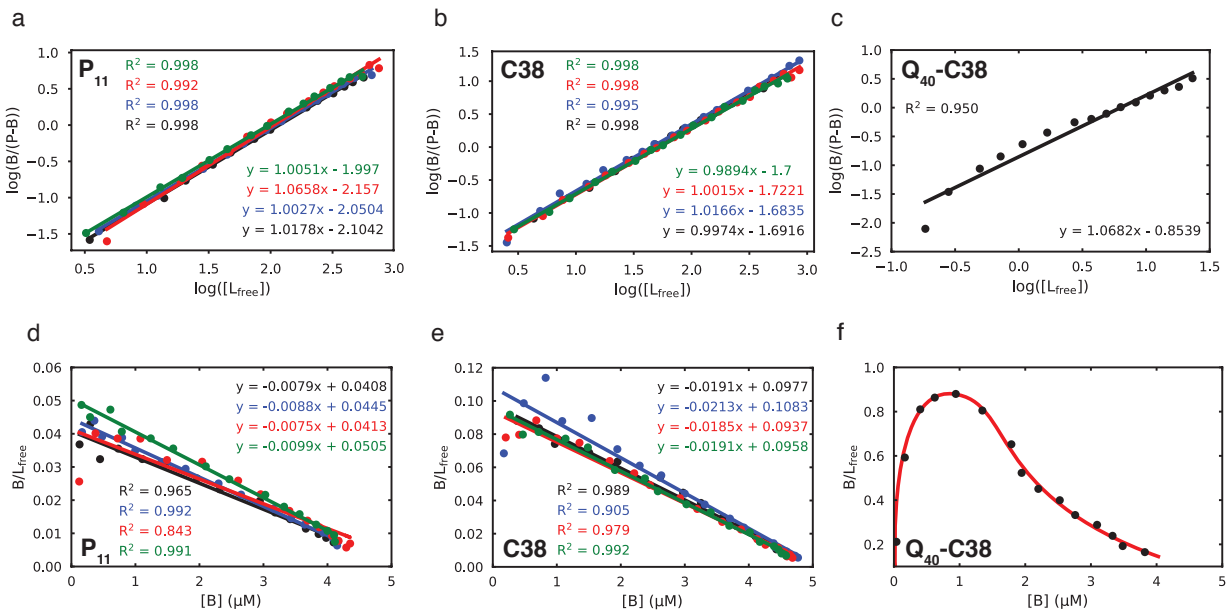

Figure S4. Hill and Scatchard analyses of profilin/P_{11}, profilin/C38 and profilin/Q_{40}-C38 binding. (**a, b, c**) Hill analyses of profilin/peptide binding with corresponding linear fits. The analyses of four independent trials for profilin/P_{11} and profilin/C38 are shown to demonstrate reproducibility. (**a**) Profilin/P_{11} served as a control since it only has a single binding site, and this was confirmed by the slope of each of the four trials being essentially equal to one. (**b**) For profilin/C38, the slope of each of the four trials was also essentially equal to one, indicating a lack of cooperativity between the polyproline modules in C38. (**c**) The slopes of Q_{40}-C38 trials were near one, but the fits were less robust, as can be seen in the selected fit (one trial that is representative of the other independent Q_{40}-C38 trials). (**d, e, f**) Given this less than optimal fit, we also performed a Scatchard analysis for clarification. (**d**) A linear (as opposed to concave) trend in a Scatchard analysis is indicative of the absence of cooperativity between binding sites, as can be seen in the P_{11} control. (**e**) The Scatchard analysis and corresponding linear fits of profilin/C38 confirm the absence of cooperativity between polyproline binding sites in C38. (**f**) The profilin/Q_{40}-C38 Scatchard plot exhibits a clear concave-down shape and intersection with the origin, which together are indicative of positive cooperativity. The red line is drawn to

highlight these characteristics. In all of the panels, “B” refers the concentration of bound species, “P” is the concentration of Profilin, and “L_{free}” is the concentration of free (unbound) ligand (i.e., peptide).

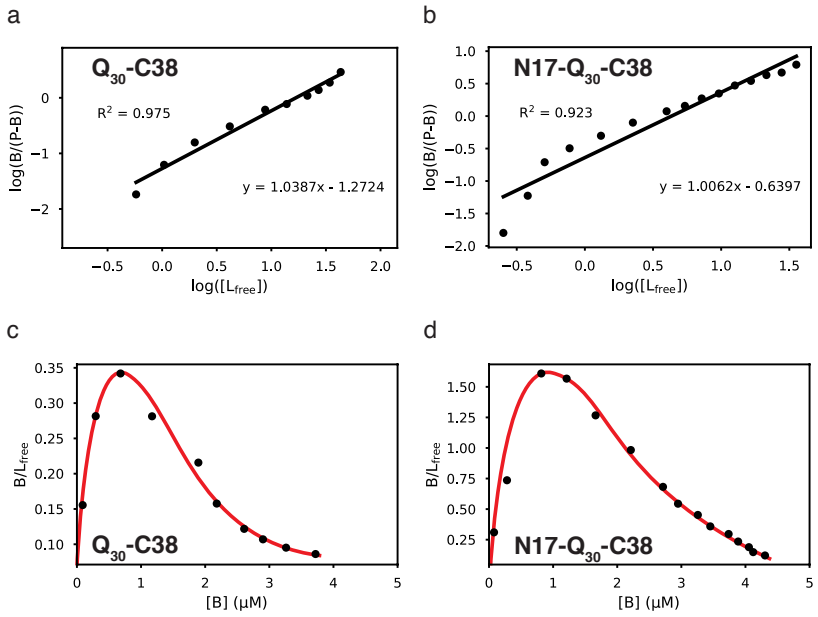

Figure S5. Hill and Scatchard analyses of profilin/Q_{30}-C38 and profilin/N17-Q_{30}-C38 binding. Hill (a, b) and Scatchard (c, d) analyses were carried out for the binding of profilin to Q_{30}-C38 and N17-Q30-C_{38}. (a, b) The Hill analysis gives a less than optimal linear fit for profilin/Q_{30}-C38 and profilin/N17-Q_{30}-C38, respectively. Similar results were obtained for additional trials, but a single representative trial is shown for the sake of clarity. (c, d) Scatchard analysis of profilin/Q_{30}-C38 and profilin/N17-Q_{30}-C38, respectively, shows clear downward concavity and intersection with the origin (again in both cases) consistent with cooperative binding. The red line is drawn to highlight these characteristics. Similar results were obtained for additional trials, but a single representative trial is shown for the sake of clarity. In all of the panels, “B” refers the concentration of bound species, “P” is the concentration of Profilin, and “L_{free}” is the concentration of free (unbound) ligand (i.e., peptide).

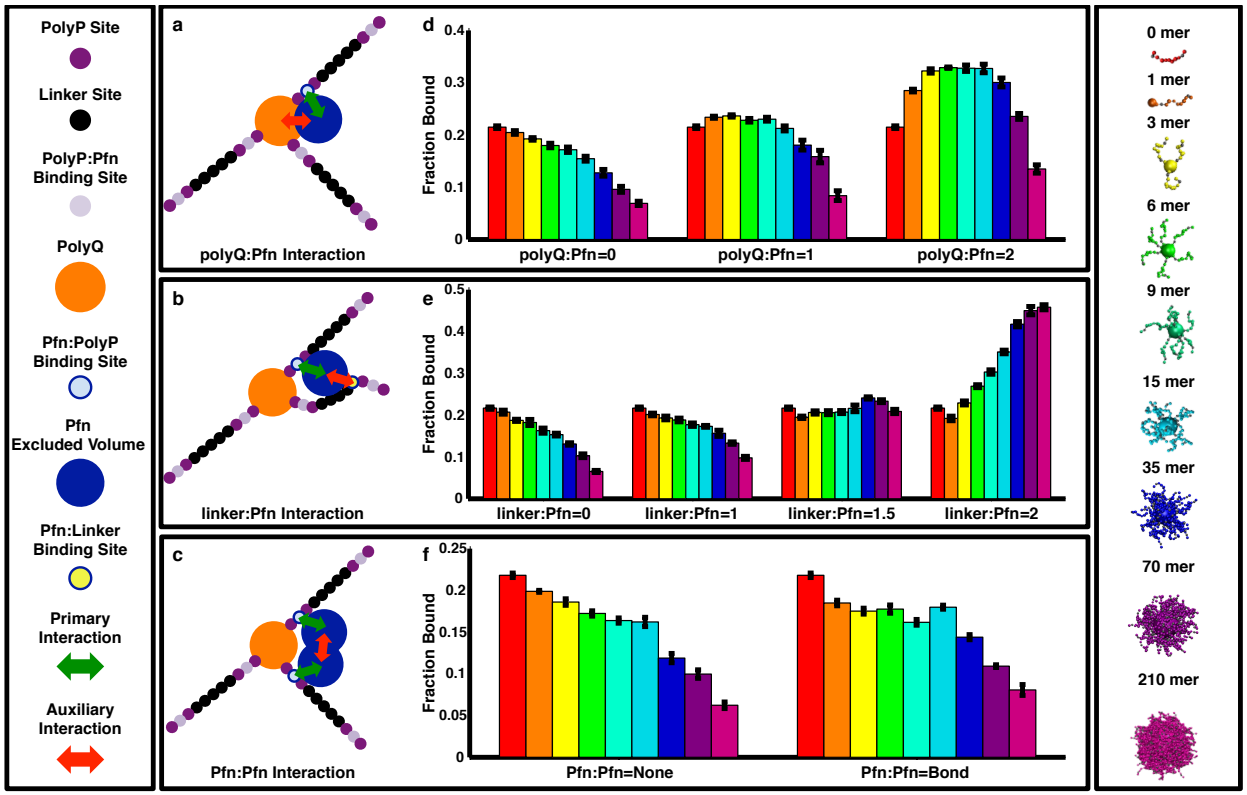

Figure S6: Test of three distinct auxiliary interaction models by coarse-grained simulations. (**a**, **b**, **c**) Visual representations of the architectures used for the polyQ:Pfn, linker:Pfn, and Pfn:Pfn auxiliary interaction models, respectively. Pfn and Q_{40}-C38 molecules are represented by a set of coarse-grained beads connected by flexible springs. Each bead type is defined in the left legend. Model details are described in the Methods section. (**d**, **e**, **f**) Fraction of profilin molecules bound to a polyP site through the primary Pfn:PolyP interaction for the polyQ:Pfn, linker:Pfn, and Pfn:Pfn auxiliary interaction models, respectively. Each bar denotes the fraction bound for a simulation consisting of M Q_{40}-C38 clusters of size X such that M x X=630, where X=1, 3, 6, 9, 15, 35, 70, or 210 (see Methods for details). The bar colors correspond to the cluster sizes given by the corresponding colors in the right legend. Error bars denote the standard error of mean calculated over five independent simulations. A cluster size of 0 indicates a C38 molecule without the polyQ domain. For the polyQ:Pfn and linker:Pfn

auxiliary interaction models the strength of the auxiliary interaction potential is given in kcal/mol below each plot. For the Pfn:Pfn auxiliary interaction model, the auxiliary Pfn:Pfn interaction is modeled by creating a bond between two Pfn beads (see Methods). Only the polyQ:Pfn auxiliary interaction model is found to be consistent with the two experimental results: (1) c_S shifts to higher concentrations in the presence of Pfn, which implies as Q_{40}-C38 cluster sizes become larger, Pfn binding must become weaker, and (2) K_d of Pfn binding C38 decreases for Q_{40}-C38 compared to C38, which implies Pfn must bind Q_{40}-C38 clusters better than C38.

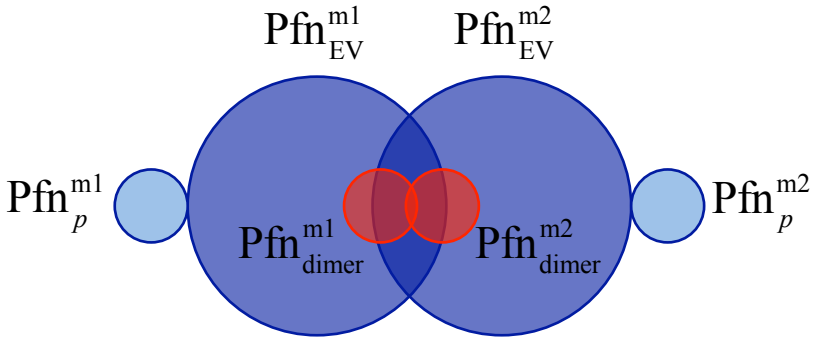

Figure S7: Pfn:Pfn dimerization architecture. Here, dark blue circles correspond to Pfn_{EV} beads, light blue circles correspond to Pfn_p beads, and red circles correspond to Pfn_{dimer} beads.

Table S1 – Measured values of c_S and dissociation constants (K_d) for peptide constructs as a function of profilin concentration

		Profilin concentration (μM)				
		0	**1**	**5**	**10**	**20**
c_S	Q_{40}-C38	0.2948 ± 0.1	3.2 ± 0.4	4.3 ± 1.0	6.4 ± 0.7	6.3 ± 3.3
K_d	Q_{40}-C38	-	11.6 ± 2.2	8.8 ± 1.7	5.6 ± 1.4	5.4 ± 1.0
	C38	-	-	50.7 ± 3.5	-	-
	P_{11}	-	-	106.5 ± 9.1	-	-
	Q_{30}-C38	-	-	20.5± 2.7	-	-
	N17-Q_{30}-C38	-	-	5.9 ± 1.3	-	-

All values in units of μM ± standard deviation for 3 - 4 independent trials

SUPPORTING REFERENCES

1. Lu, X., and Murphy, R. M. (2014) Synthesis and disaggregation of asparagine repeat-containing peptides. *Journal of Peptide Science* **20**, 860-867
2. Crick, S. L., Ruff, K. M., Garai, K., Frieden, C., and Pappu, R. V. (2013) Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20075-20080
3. Radhakrishnan, A., Vitalis, A., Mao, A. H., Steffen, A. T., and Pappu, R. V. (2012) Improved atomistic Monte Carlo simulations demonstrate that poly-L-proline adopts heterogeneous ensembles of conformations of semi-rigid segments interrupted by kinks. *J. Phys. Chem. B* **116**, 6862-6871
4. Vitalis, A., and Pappu, R. V. (2009) ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **30**, 673-699
5. Das, R. K., Crick, S. L., and Pappu, R. V. (2012) N-terminal segments modulate the alpha-helical propensities of the intrinsically disordered basic regions of bZIP proteins. *J Mol Biol* **416**, 287-299
6. Plimpton, S. (1995) Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* **117**, 1-19
7. Moore, T. C., Iacovella, C. R., and McCabe, C. (2014) Derivation of coarse-grained potentials via multistate iterative Boltzmann inversion. *J Chem Phys* **140**, 224104
8. Kayatekin, C., Matlack, K. E., Hesse, W. R., Guan, Y., Chakrabortee, S., Russ, J., Wanker, E. E., Shah, J. V., and Lindquist, S. (2014) Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12085-12090
9. Ripaud, L., Chumakova, V., Antonin, M., Hastie, A. R., Pinkert, S., Korner, R., Ruff, K. M., Pappu, R. V., Hornburg, D., Mann, M., Hartl, F. U., and Hipp, M. S. (2014) Overexpression of Q-rich prion-like proteins suppresses polyQ cytotoxicity and alters the polyQ interactome. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18219-18224
10. Wear, M. P., Kryndushkin, D., O'Meally, R., Sonnenberg, J. L., Cole, R. N., and Shewmaker, F. P. (2015) Proteins with Intrinsically Disordered Domains Are Preferentially Recruited to Polyglutamine Aggregates. *PLoS One* **10**, e0136362
11. Mahoney, N. M., Rozwarski, D. A., Fedorov, E., Fedorov, A. A., and Almo, S. C. (1999) Profilin binds proline-rich ligands in two distinct amide backbone orientations. *Nat. Struct. Biol.* **6**, 666-671
12. Shao, J., Welch, W. J., Diprospero, N. A., and Diamond, M. I. (2008) Phosphorylation of profilin by ROCK1 regulates polyglutamine aggregation. *Mol Cell Biol* **28**, 5196-5208