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

Ammon E. Posey¹, Kiersten M. Ruff¹, Tyler S. Harmon¹, Scott L. Crick¹, Aimin Li³, Marc I. Diamond², Rohit V. Pappu¹*

¹Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA
²Center for Alzheimer’s and Neurodegenerative Diseases, University of Texas, Southwestern Medical Center, Dallas, TX 75390, USA
³Department 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
Steady-state tryptophan fluorescence data analysis
Right-angle static light scattering data analysis
Thioflavin T fluorescence measurements
TEM image analysis
All Atom Simulations
Coarse-Grained Langevin Dynamics Simulations
Coarse-Grained Model Architecture
Coarse-Grained Model Energy Function
Calculation of Fraction Bound from Simulations

SUPPORTING FIGURES AND TABLE
Figure S1: The C38 region is required for modulation of Htt-NTF aggregation by profilin
Figure S2: Aggregate size distributions determined by analysis of TEM images
Figure S3. Profilin binds the polyP tracts in the C38 region of Htt-NTFs
Figure S4. Hill and Scatchard analyses of profilin/P₁₁₃, profilin/C₃₈ and profilin/Q₄₀-C₃₈ binding
Figure S5. Hill and Scatchard analyses of profilin/Q₃₀-C₃₈ and profilin/N₁₇-Q₃₀-C₃₈ binding.
Figure S6: Test of three distinct auxiliary interaction models by coarse-grained simulations
Figure S7: Pfn:Pfn dimerization architecture

Table S1 – Measured values of c₅ and dissociation constants (Kₐ) for peptide constructs as a function of profilin concentration

SUPPORTING REFERENCES
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 

\[ F = \frac{F_{\text{max}} [L]}{K_d + [L]} \]

where \(F_{\text{max}}\) is the fluorescence maximum recovered from fitting the raw data and \([L]\) is the concentration of Q40-C38. The data were then converted to fraction bound, where fraction bound is defined as: \(F/F_{\text{max}}\). 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.
\[
F = F_{\text{max}} \frac{([M] + [L] + K_d)}{2[M]} \left( [M] + [L] + K_d \right) - \frac{\sqrt{[M] + [L] + K_d}^2 - 4[M][L]}{2[M]}
\]

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 \(c_0\) 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 $T=[288 \text{ K}, 293 \text{ K}, 298 \text{ K}, 305 \text{ K}, 310 \text{ K}, 315 \text{ K}, 320 \text{ K}, 325 \text{ K}, 335 \text{ K}, 345 \text{ K}, 360 \text{ K}, 375 \text{ K}, 390 \text{ K}, 405 \text{ K}]$. Each simulation consisted of $6.15 \times 10^7$ 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 $5.15 \times 10^7$ steps. Temperature swaps were proposed every $5 \times 10^4$ 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$_{11}$-QLPQPQQPLLPQPQ-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

$$F_i = -\nabla W_{\text{eff},i} - \frac{m_i}{\gamma_i} v_i + R_i. \quad (1)$$

Here, $W_{\text{eff}}$ is the effective energy function described below (Equation (2)). The second term denotes the frictional force which is proportional to the velocity of bead $i$, $v_i$. The damping term, $\gamma_i$, is calculated using $\gamma_i = C \frac{m_i}{6\pi \eta R_i}$. Here, $m_i$ is the mass of bead $i$, $\eta=6.29 \times 10^{-4}$ 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-Ribiére conjugate gradient algorithm (6). Then, $10^6$ 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 $R_g = R_0 N^\nu$. Here, $R_g$ is the radius of gyration, $N$ is the number of glutamine residues in the cluster, $\nu = 0.33$ for globules, and $R_0 = 2.85$ is the pre-factor. The pre-factor $R_0$ was determined by fitting the $<R_g>$ from all atom simulations of $Q_{22}$ and $Q_{40}$ to the equation $<R_g> = R_0 N^{0.33}$.

The coarse-grained architecture of the C38 stretches is $p-p_{Pfn}-p-I-I-I-I-I-p_{Pfn}-p$. 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 $I$ 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, $W_b$, $W_\theta$, $W_{LJ}$, $W_{\text{Primary}}$, 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

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

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

Here, $N_b$ and $N_\theta$ 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 $\theta_{0i}$, respectively. The Boltzmann inversion procedure (7) was used to extract $K_i$, $L_i$, $b_{0i}$, and $\theta_{0i}$ 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 \( \theta_{0i} \):

\[
\begin{align*}
b_{0i} &= \langle b_i \rangle, \quad K_i = \frac{k_B T}{\langle b_i^2 \rangle - \langle b_i \rangle^2}, \\
\theta_{0i} &= \langle \theta_i \rangle, \quad L_i = \frac{k_B T}{\langle \theta_i^2 \rangle - \langle \theta_i \rangle^2}.
\end{align*}
\]

Here, \( b_i \) and \( \theta_i \) 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, \( \theta_{0i} \) 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

\[
W_{LJ} = \sum_{i=1}^{N_{nb}} \sum_{j<i} 4\varepsilon_{ij}\left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right], \quad r_{ij} < r_c
\]

\[
= 0, \quad r_{ij} \geq r_c.
\]

Here, \( N_{nb} \) is the number of beads, \( r_{ij} \) is the distance between beads \( i \) and \( j \), \( \sigma_{ij} \) is the distance at which the inter-bead potential is zero, and \( \varepsilon_{ij} \) 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 \( \sigma_{ij} = \left( \frac{2R_g^i + 2R_g^j}{2} \right) / 2 \text{ Å} \),
where $R_g^i$ is the $R_g$ extracted from all atom simulations or the Pfn crystal structure for bead $i$ as explained above, $\epsilon_p=0.01$ 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,

$$W_{\text{Primary}} = w_G + w_{LJ},$$

$$w_G = \sum_{i=1}^{N_{\text{pPfn}}} \sum_{j=1}^{N_{\text{pPfn}}} -\epsilon_G \exp \left[ -\frac{(r_{ij} - r_G)}{2\sigma_G} \right], \quad r_{ij} < r_G^{\text{cutoff}},$$

$$w_{LJ} = \sum_{i=1}^{N_{\text{pPfn}}} \sum_{j=1}^{N_{\text{EV}}} 4\epsilon_{LJ} \left[ \left( \frac{\sigma_{LJ}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{LJ}}{r_{ij}} \right)^{6} \right], \quad r_{ij} < r_{LJ}^{\text{cutoff}}.$$

Here, $N_{\text{pPfn}}$, $N_{\text{pPfn}_{p}}$, $N_{\text{pPfn}_{Pfn}}$, and $N_{\text{pPfn}_{EV}}$ correspond to the number of $p_{\text{Pfn}}$ beads, the number of Pfn$_p$ beads, the number of $p$ and $p_{\text{Pfn}}$ beads, and the number of Pfn$_{EV}$ beads, respectively. For $w_G$, $r_{ij}$ is the distance between beads $i$ and $j$, $\epsilon_G=6$ kcal/mol is the interaction strength, $r_G=0$ Å is the position of the well, $\sigma_G=2.27$ Å controls the width of the well, and $r_G^{\text{cutoff}}=2.5R_{\text{pPfn}}$, where $R_{\text{pPfn}}$ is the radius of the C38 $p_{\text{Pfn}}$ bead. The width of the well was chosen such that the energy for two $p_{\text{Pfn}}$ beads to bind the Pfn$_p$ bead on a single Pfn molecule was half that of the energy for a single $p_{\text{Pfn}}$ bead to bind the Pfn$_p$ bead. This choice was made in order to disfavor a single Pfn$_p$ from binding multiple $p_{\text{Pfn}}$ beads. Explicitly, in the dimer case $r_{p_{\text{Pfn}},p_{\text{Pfn}}} = R_g$ since $\sigma_{p_{\text{Pfn}},p_{\text{Pfn}}} = 2R_g$, where $R_g$ is the radius of the $p_{\text{Pfn}}$ bead. This yields
For \( w_{LJ}, r_{ij} \) is the distance between beads \( i \) and \( j \), \( \epsilon_{LJ} = 0.01 \text{ kcal/mol} \) is the interaction strength, \( \sigma_{LJ} \) is the distance at which the inter-bead potential is zero, and \( r_{cutoff}^{LJ} = 2.5\sigma_{LJ} \). Here, \( \sigma_{LJ} \) was set to the radius of the Pf\(_{nEV}\) bead. This choice prevented excluded volume effects between Pf\(_{nEV}\) beads and \( p \) and Pf\(_{nP}\) beads from disfavoring Pf\(_{nP}\)-Pf\(_{nP}\) 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:Pf\(_{n}\) interaction, a linker:Pf\(_{n}\) interaction, and a Pf\(_{n}^{\prime}\):Pf\(_{n}\) interaction. The details of each model will be discussed below.

The auxiliary interaction potential for the polyQ:Pf\(_{n}\) interaction is modeled as a Lennard-Jones potential between the polyQ bead and the Pf\(_{nEV}\) 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 Pf\(_{n}\). Explicitly, the model takes the form

\[
W_{\text{Auxiliary}}^{\text{polyQ:Pf}n} = \sum_{i=1}^{N_{\text{polyQ}}} \sum_{j=1}^{N_{\text{Pf}nEV}} \sum_{\text{polyQ:Pf}n} 4\epsilon_{\text{polyQ:Pf}n} \left[ \left( \frac{\sigma_{\text{polyQ:Pf}n}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{\text{polyQ:Pf}n}}{r_{ij}} \right)^{6} \right], \quad r_{ij} < r_{cutoff}^{\text{polyQ:Pf}n}
\]

\[
= 0, \quad r_{ij} \geq r_{cutoff}^{\text{polyQ:Pf}n}.
\]
In Equation (4), \( N_{\text{polyQ}} \) 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, \( R_{\text{polyQ}} \) is the radius of the polyQ bead. The cutoff, \( r_{\text{cutoff}}^{\text{polyQ:Pfn}} \), 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 Pfn_{linker}, was set to have the same mass and radius as the Pfn_{\rho} 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 Pfn_{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,

\[
W^{\text{linker:Pfn}}_{\text{Auxiliary}} = W^b_{\text{linker:Pfn}} + W^\theta_{\text{linker:Pfn}} + W^\text{LJ}_{\text{linker:Pfn}},
\]

\[
W^b_{\text{linker:Pfn}} = \sum_{i=1}^{N_{\text{Pfn}}} K_i \left( b_i - b_{0i} \right)^2 / 2,
\]

\[
W^\theta_{\text{linker:Pfn}} = \sum_{i=1}^{N_{\text{Pfn}}} L_i \left( \theta_i - \theta_{0i} \right)^2 / 2,
\]

\[
W^\text{LJ}_{\text{linker:Pfn}} = \sum_{i=1}^{N_{\text{Pfn}}} \sum_{j=1}^{N_{\text{Pfn}_{\text{linker}}}} 4\varepsilon_{\text{linker:Pfn}} \left[ \left( \frac{\sigma_{\text{linker:Pfn}}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{\text{linker:Pfn}}}{r_{ij}} \right)^6 \right], r_{ij} < r_{\text{cutoff}}^{\text{linker:Pfn}},
\]

\[
= 0, r_{ij} \geq r_{\text{cutoff}}^{\text{linker:Pfn}}.
\]

Here, \( N_{\text{Pfn}} \), \( N_{\text{linker}} \) and \( N_{\text{Pfn}_{\text{linker}}} \) denote the number of Pfn molecules, non-polyP linker beads, and Pfn_{linker} beads, respectively. For \( W^\text{LJ}_{\text{linker:Pfn}} \), \( r_{ij} \) 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,
\( R_{\text{Pfn}_{\text{linker}}} \) and \( R_i \) are the radii for the Pfn_{\text{linker}} and \( i \) beads, respectively. The cutoff, \( r_{\text{cutoff}_{\text{linker}:\text{Pfn}}} \), was set to 2.5\( \sigma_{\text{linker}:\text{Pfn}} \).

In Equation (5), \( w_{b_{\text{linker}:\text{Pfn}}} \) and \( w_{\theta_{\text{linker}:\text{Pfn}}} \) account for the bond length and bond angle potentials between Pfn_{\text{linker}}:Pfn_{\text{EV}} beads and Pfn_{P_{\text{Pfn}}}:Pfn_{\text{EV}:Pfn_{\text{linker}}} beads, respectively. The equilibrium bond length (\( b_{0i} \)) and force constant (\( K_i \)) between Pfn_{\text{linker}} and Pfn_{\text{EV}} were set to the sum of the radii of the two Pfn beads and 10 kcal/mol/Å\(^2\), respectively. The equilibrium angle (\( \theta_{0i} \)) 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, Pfn_{dimer}, was added to the architecture (see Figure S6). 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 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,
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 $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}$ and $\text{Pfn}_{2}^{\text{EV}}:\text{Pfn}_{2}^{\text{dimer}}$. Here, m1 and m2 correspond to molecule 1 and molecule 2, respectively. The equilibrium bond length ($b_{0i}$) and force constant ($K_i$) between $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}$ and $\text{Pfn}_{2}^{\text{EV}}:\text{Pfn}_{2}^{\text{dimer}}$ were set to 6.47 Å and 10 kcal/mol/Å$^2$, respectively. The new angles defined are $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}:\text{Pfn}_{2}^{\text{dimer}}$ and $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}:\text{Pfn}_{2}^{\text{EV}}$. The equilibrium angle ($\theta_{0i}$) and force constant ($L_i$) for $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}:\text{Pfn}_{2}^{\text{dimer}}$ and $\text{Pfn}_{1}^{\text{EV}}:\text{Pfn}_{1}^{\text{dimer}}:\text{Pfn}_{2}^{\text{EV}}$ were set to 180 degrees and 10 kcal/mol/radians$^2$, respectively. These choices allow for the distance between the two Pfn beads of a dimer to vary while the distance between the two PfnEV beads is held relatively fixed.

We allow for variability in the distance between the two Pfn beads since the site of Pfn dimerization is not known. However, in order to restrict the two Pfn beads from coming too close we changed the Lennard-Jones potential parameters between two Pfn beads. The Lennard-Jones potential between two Pfn beads is defined by $w_{LJ}^{\text{Pfn:Pfn}}$ in Equation (6). Here, $N_{\text{Pfn}}$ is the number of Pfn beads, $r_{ij}$ is the distance between beads $i$ and $j$, $\varepsilon_{\text{Pfn:Pfn}} = 0.01$ and $\sigma_{\text{Pfn:Pfn}}$. 

\[ W^{\text{Auxiliary}}_{\text{Pfn:Pfn}} = w_b^{\text{Pfn:Pfn}} + w_{\theta}^{\text{Pfn:Pfn}} + w_{LJ}^{\text{Pfn:Pfn}}, \]

\[ w_b^{\text{Pfn:Pfn}} = \sum_{i=1}^{N_{b}^{\text{Pfn:Pfn}}} \frac{K_i \left( b_i - b_{0i} \right)^2}{2}, \]

\[ w_{\theta}^{\text{Pfn:Pfn}} = \sum_{i=1}^{N_{\theta}^{\text{Pfn:Pfn}}} \frac{L_i \left( \theta_i - \theta_{0i} \right)^2}{2}, \]

\[ w_{LJ}^{\text{Pfn:Pfn}} = \sum_{i=1}^{N_{\text{Pfn}}^{\text{Pfn}}} \sum_{j \neq i}^{N_{\text{Pfn}}^{\text{Pfn}}} 4 \varepsilon_{\text{Pfn:Pfn}} \left[ \left( \frac{\sigma_{\text{Pfn:Pfn}}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{\text{Pfn:Pfn}}}{r_{ij}} \right)^{6} \right], \]

\[ r_{ij} < r_{cutoff}^{\text{Pfn:Pfn}}, \]
kcal/mol is the strength of the interaction, $\sigma_{\text{Pfn,Pfn}} = 10$ Å is the distance at which the inter-bead potential is zero, and $r_{\text{cutoff}}^{\text{Pfn,Pfn}} = 2.5\sigma_{\text{Pfn,Pfn}}$.

**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 $p_{\text{Pfn}}$ 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 $p_{\text{Pfn}}$ bead if the distance between the two beads was less than $2R_{\text{Pfn}}^{\text{Pfn}}$, where $R_{\text{Pfn}}^{\text{Pfn}}$ is the radius of the $p_{\text{Pfn}}$ 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.
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 \( \frac{\Delta F}{F_{\text{ref}}} \) where \( \Delta F = (F_X - F_{\text{ref}}) \). Here, \( F_X \) is the intracellular FRET for the construct in question and \( F_{\text{ref}} \) is the FRET measured for N17-Q72-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-Q72-C38-CFP/YFP aggregation by 30%, as described previously (12). However, the aggregation of N17-Q72-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/P₁₁, profilin/C₃₈ and profilin/Q₄₀-C₃₈ binding. (a, b, c) Hill analyses of profilin/peptide binding with corresponding linear fits. The analyses of four independent trials for profilin/P₁₁ and profilin/C₃₈ are shown to demonstrate reproducibility. (a) Profilin/P₁₁ 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 Q₄₀-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 Q₄₀-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 P₁₁ 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/Q₄₀-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\textsubscript{30}-C\textsubscript{38} and profilin/N\textsubscript{17}-Q\textsubscript{30}-C\textsubscript{38} binding. Hill (a, b) and Scatchard (c, d) analyses were carried out for the binding of profilin to Q\textsubscript{30}-C\textsubscript{38} and N\textsubscript{17}-Q\textsubscript{30}-C\textsubscript{38}. (a, b) The Hill analysis gives a less than optimal linear fit for profilin/Q\textsubscript{30}-C\textsubscript{38} and profilin/N\textsubscript{17}-Q\textsubscript{30}-C\textsubscript{38}, 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\textsubscript{30}-C\textsubscript{38} and profilin/N\textsubscript{17}-Q\textsubscript{30}-C\textsubscript{38}, 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 $M \times X = 630$, 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

<table>
<thead>
<tr>
<th>Profilin concentration (µM)</th>
<th>0</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_S$</td>
<td>Q$<em>{40}$-C$</em>{38}$</td>
<td>0.2948 ± 0.1</td>
<td>3.2 ± 0.4</td>
<td>4.3 ± 1.0</td>
<td>6.4 ± 0.7</td>
</tr>
<tr>
<td></td>
<td>Q$<em>{40}$-C$</em>{38}$</td>
<td>-</td>
<td>11.6 ± 2.2</td>
<td>8.8 ± 1.7</td>
<td>5.6 ± 1.4</td>
</tr>
<tr>
<td></td>
<td>C$_{38}$</td>
<td>-</td>
<td>-</td>
<td>50.7 ± 3.5</td>
<td>-</td>
</tr>
<tr>
<td>$K_d$</td>
<td>P$_{11}$</td>
<td>-</td>
<td>-</td>
<td>106.5 ± 9.1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Q$<em>{30}$-C$</em>{38}$</td>
<td>-</td>
<td>-</td>
<td>20.5 ± 2.7</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>N17-Q$<em>{30}$-C$</em>{38}$</td>
<td>-</td>
<td>-</td>
<td>5.9 ± 1.3</td>
<td>-</td>
</tr>
</tbody>
</table>

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

SUPPORTING REFERENCES