SUPPLEMENTARY INFORMATION

Identifying polyglutamine protein species *in situ* that best predict neurodegeneration

Jason Miller[1,2,3,Ω], Montserrat Arrasate[1,4,Σ,Ω], Elizabeth Brooks[1,4,#], Clare Peters-Libeu[1,5,π], Justin Legleiter[1∞], Danny Hatters[1,5,+], Jessica Curtis[1,4], Kenneth Cheung[1,4,§], Preethi Krishnan[1,4,¥], Siddhartha Mitra[1,3,6], Kartika Widjaja[1,4], Benjamin A. Shaby[7,θ], Gregor P. Lotz[1], Yvonne Newhouse[1,5], Emily Mitchell[8,9], Alex Osmand[10], Michelle Gray[11,Φ], Vanitha Thulasiramin[12], Frederic Saudou[13], Mark Segal[14], X. William Yang[11], Eliezer Masliah[15], Leslie M. Thompson[8,9,16], Paul J. Muchowski[1,4,6,17,18], Karl H. Weisgraber[1,5,19], and Steven Finkbeiner[1,4,6,18,20*]

[1]Gladstone Institute of Neurological Disease, the Taube Koret Center for Huntington's Disease Research, and the Hellman Family Foundation Program in Alzheimer's Disease Research, San Francisco, California 94158, USA
[2]Chemistry and Chemical Biology Program, [3]Medical Scientist Training Program, [4]Neuroscience Program, University of California, San Francisco, California 94143, USA
[5]Gladstone Institute of Cardiovascular Disease, San Francisco, California 94158, USA
[6]Biomedical Sciences Program, University of California, San Francisco, California 94143-0452, USA
[7]Department of Statistical Science, Cornell University, Ithaca, New York 14853, USA
[8]Department of Biological Chemistry, [9]Department of Psychiatry and Human Behavior, University of California, Irvine, California 92697, USA
[10]Department of Medicine, University of Tennessee Graduate School of Medicine, Knoxville, Tennessee 37920, USA
[11]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, Department of Psychiatry and Biobehavioral Sciences, Brain Research Institute, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA
[12]Bio-Rad Laboratories, Fremont, California 94555, USA
[13]Institute Curie, 2CNRS UMR 146, 91405 Orsay, France
[14]Division of Biostatistics, University of California, San Francisco, California 94143-0560, USA
[15]Departments of Neurosciences and Pathology, University of California, San Diego, California 92093, USA
[16]Department of Neurobiology and Behavior, University of California, Irvine, California 92697, USA
[17]Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA
[18]Department of Neurology, University of California, San Francisco, California 94143-0114, USA
[19]Department of Pathology, University of California, San Francisco, California 94143-0506, USA
[20]Department of Physiology, University of California, San Francisco, California 94143-2610, USA

[Ω] Authors made equal contributions to this work.
[*] To whom correspondence should be addressed to: S.F. (sfinkbeiner@gladstone.ucsf.edu)

SUPPLEMENTARY METHODS

Generation of monoclonal antibodies

Six mice were immunized against a fusion protein encoding glutathione-S-transferase fused at its carboxyl terminus to an amino-terminal 171 fragment of htt containing a stretch of 66 glutamines (GST-171-Q_{66}), and 480 hybridomas were prepared. Supernatants containing secreted monoclonal antibodies were screened for their selective immunoreactivity against mutant htt (mHtt) by a native sandwich enzyme-linked immunosorbent assay (ELISA). Those with the highest levels of immunoreactivity were tested *in situ*. Large-scale production of 3B5H10 was carried out at the National Cell Culture Center in a 3B5H10 hybridoma line grown in Hyclone serum-free medium.

Plasmids

Expression plasmids encoding an N-terminal fragment of htt fused to GFP (pGW1-Htt^{ex1}-(Q_{46} or Q_{97})) were derived from pcDNA3.1-based plasmids[1] by subcloning into pGW1-CMV (British Biotechnologies). Htt^{ex1}-Q_{17} or Q_{46}-CyPet or YPet plasmids were derived from cloning CyPet or YPet PCR product (CyPet and YPet plasmids optimized for mammalian codons were a gift of Patrick Daugherty at UC, Santa Barbara) into pcDNA3.1-based htt^{ex1} plasmids[1] and then subcloning into pGW1-CMV. Plasmid constructions were confirmed by sequencing. Htt^{ex1}-(Q_{17},Q_{72})-eGFP have been described[2].

Cell culture and transfection

Primary cultures of rat striatal neurons were prepared from embryos (E18–20) and transfected with plasmids (5–6DIV) as described[3,4]

(http://www.gladstone.ucsf.edu/gladstone/php/?sitename=finkbeiner). Typically, neurons were co-transfected with pGW1-mRFP and a version of pGW1-Htt^{ex1}-(Q_{17}, Q_{25}, Q_{47}, Q_{72}, or Q_{97})-GFP in a ~1:1 molar ratio, with a total of 1–4 mg of DNA in each well of a 24-well plate. After transfection, neurons were maintained in serum-free medium. HEK293 cells were maintained and transfected with plasmids as described[5].

Immunocytochemistry

All immunocytochemistry was performed as described[5]. MAbs MW1, MW7, and MW8 have been described[6]. The α-oligomer antibody has also been described[7]. α-c-Myc chicken polyclonal antibody was obtained from Aves Labs Inc. (Tigard, OR). EM48 was obtained from Chemicon/Millipore (Billerica, MA) (Mab5374). Cy3- and Cy5-conjugated secondary antibodies were acquired from Jackson ImmunoResearch (West Grove, PA). For 3B5H10 labeling in **Figure 1e**, Alexa-647 was directly conjugated to 3B5H10 with an antibody labeling kit, according to the manufacturer's instructions from Invitrogen/ Molecular Probes (Carlsbad, CA).

Quantitative fluorescence confocal microscopy with a Zeiss LSM 510 microscope for **Figure 2a** and **Supplementary Figure 7** was performed on neurons fixed 24 h after transfection. Lasers were given 2 h to warm up and stabilize before any measurements were made; additionally, reference samples were periodically imaged during each experiment to confirm consistency in image intensity. For each antibody, one photomultiplier tube (PMT) detector was set to a single gain and offset setting for Htt^{ex1}-Q_n-eGFP detection, and another PMT detector was set to a different single gain and offset setting for antibody detection (via Cy5 or Alexa647 labeled secondary antibody). PMT gain and offset settings were dictated by the detection range necessary to image, with a single setting, Htt^{ex1}-Q_n-eGFP or antibody fluorescence across all

polyQ lengths. While each antibody vs. Htt^{ex1}-Q_n-eGFP graph presented in **Figure 2a** has a single PMT gain and offset setting for each fluorescence channel, those settings are not consistent across antibodies. Therefore, absolute levels of GFP or Cy5 fluorescence cannot be compared across graphs in **Figure 2a**. In contrast, absolute levels of GFP and Cy5 fluorescence for **Supplementary Figure 7** have been normalized to each other and therefore can be compared across graphs.

Each neuron subjected to quantitative fluorescent imaging was selected randomly. To avoid effects of photobleaching, each neuron selected for imaging was exposed to epifluorescent light for less than 5 s before a single LSM image was acquired. Images were acquired using a 63×, 1.4 N.A. oil immersion lens. Levels of htt and antibody were then extracted from images by drawing a region of interest across the neuronal cell body and recording average pixel intensity using Zeiss LSM 510 software. Background subtraction was deemed unnecessary as PMT offset was set to ensure no detection of background.

Analysis of interaction of 3B5H10 with polyQ by SELDI-TOF-MS

PolyQ ($K_2Q_{39}K_2$) or PACAP peptide (basic peptide) dissolved in 50 mM sodium bicarbonate buffer pH 9.0 was incubated overnight with shaking at 4^{o}C on PS10 ProteinChip Array (Bio-Rad). Peptide solution was removed. Array spots were subsequently blocked with bovine serum albumin (7 mg/ml) in phosphate buffered saline (PBS, pH 7.2) containing 0.1% Tween for 4 h at ~25^{o}C with shaking. Block was removed, and the array spots were washed twice with PBS. Control antibody or 3B5H10 in PBS with 0.1% Triton was added for 4 h with shaking at ~25^{o}C. The antibody solution was removed. The arrays were washed with PBS once for 2 min with shaking, followed by 2 washes with Urea-CHAPS buffer (1 M urea, 0.1% CHAPS, 0.5 M NaCl,

50 mM Tris, pH 7.2) for 2 min each. The arrays were washed twice more with PBS and then once with water before being allowed to dry. Sinapinic acid in 50% acetonitrile and 0.1% TFA were added twice; the arrays were allowed to dry after each addition. Finally, arrays were loaded into a ProteinChip SELDI System (Bio-Rad) for data collection.

Immunogold labeling and electron microscopy

Primary striatal cell cultures were plated in Mat-Tek (Ashland, MA) dishes and transfected with appropriate constructs. Approximately 24 h after transfection, plates were fixed in 4% paraformaldehyde for 15 min. Plates were then fixed in 2% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.4, at room temperature for 2 min, then placed on ice for 30 min, washed with 0.1 M cacodylate buffer and postfixed with 1% osmium tetraoxide for 30 min. After the dishes were washed in 0.1 M cacodylate buffer and double-distilled water, the cells were dehydrated in ethanol series (20, 50, 70, and 90% on ice for 1 min each) and three times in 100% ethanol (dry) at room temperature. The cells were pre-embedded with 50% Durcupan epoxy resin and 50% ethanol (dry) for 30 min and then embedded in Durcupan mix epoxy resin and polymerization at vacuum at 60°C for 48 h.

After the resin was polymerized, cells were detached from the coverslips in the dishes, and small blocks (2 cubic mm) were mounted into plastic cylinders, sectioned with an ultra microtome (Reichert Ultracut E) at 60 nm thickness and collected in nickel grids for immunogold labeling.

The grids were treated with antigen retrieval (sodium periodate saturated in water) for 1 min, washed in water, blocked with 3% BSA in TBS for 30 min and incubated with 3B5H10 (1:100) overnight. Next day, the grids were washed in TBS, blocked with 3% BSA, and

incubated with the secondary antibody IgG-antimouse/10-nm gold particles (AURION Immunogold reagents) for 2 h at room temp. Grids were then washed in TBS and deionized water. Labeling enhancement was performed using silver mixture (AURION R-gent SE-EM) for 25 min, followed by extensive washing with distilled water and then contrasted.

The immunostained grids were post-stained using saturated uranyl acetate solution in 50% ethanol for 20 min at room temperature, washed in distilled water, and placed in bismuth nitrate solution for 10 min, followed by a final wash in deionized water. The immunolabeled grids were analyzed with a Zeiss EM10 electron microscope.

Immunohistochemistry

BACHD, R6/2, and YAC-SCA3 tissue sections were immunostained with the same protocol. Biotinylation of 3B5H10 was performed using N-hydroxysuccinimide-(polyethylene glycol)$_4$-biotin (Quanta Biodesign, Columbus OH), following the manufacturer's recommendations, to a density of 3 molecules of biotin per molecule of IgG (b^3-3B5H10). Immunohistochemical staining was performed under standard conditions[8]: no blocking was used other than pretreatments with sodium borohydride and detergent, except that 1% normal mouse serum was included with the antibody to block potential F_c-receptor interactions. b^3-3B5H10 was incubated with free-floating 30-35μm sections overnight at concentrations ranging from 30 to 100 ng/ml. Aggressive "antigen retrieval" was performed to dissociate hydrogen-bonded polyglutamine aggregates by incubating some of the sections in 88–98% formic acid (3 × 10 min) before pretreatments and incubation with b^3-3B5H10. DAB staining was preceded by tyramide signal amplification, except for the YAC-SCA3 tissue staining. YAC-SCA3 tissue was additionally thionin/Nissl counterstained.

Robotic microscope imaging system

The robotic microscope imaging system has been described[9], and a second-generation system was constructed and used with some modifications. The systems are based on inverted Nikon microscopes (TE300 Quantum or TE2000E/Perfect Focus). Images were detected and with Photometrics CoolSNAP HQ or Hammatsu Orca II 12/14 bit, digital, cooled CCD cameras and digitized with Media Cybernetics Image Pro or Universal Imaging Metamorph software. Stage movements and focusing were executed with computer-controlled stepper motors. Fluorescence excitation and emission filters were moved into or out of the optical path with each program loop by two 10-position filter wheels (Sutter Instruments) under computer control. Computer commands that perform and coordinate automated stage movements, filter wheel movements, and focusing were generated with software programs that combine custom-designed and commercially available algorithms as described[9].

Image analysis

Measurements of htt expression, IB formation, and survival were extracted from files generated with automated imaging by automated analysis programs or by visual inspection as described[2,9]. Htt expression was measured 24 h after transfection. For statistical analysis, survival time was defined as the imaging time point at which a cell was last seen alive. A step-by-step description of our image analysis for this study appears in **Figure 3** and **Supplementary Figure 6** online.

Hierarchical Bayesian statistical analysis of predictive power for α-htt antibodies

For a step-by-step narrative description of how we combine immunocytochemistry, automated microscopy, Cox models, and hierarchical Bayesian analysis to determine which α-htt antibodies predict neurodegeneration, see **Figure 3** and **Supplementary Figure 6**. The following is the mathematical formulation of the analysis.

Dataset used:

For survival analysis (**Fig. 3a** at top; **Supplementary Fig. 6b**(iv)), 1508 neurons were analyzed. Each neuron contained one of five constructs (eGFP-only, Htt^{ex1}-Q_{17}-eGFP, Htt^{ex1}-Q_{46}-eGFP, Htt^{ex1}-Q_{72}-eGFP, Htt^{ex1}-Q_{97}-eGFP) such that approximately 300 neurons were transfected with each construct.

For calculation of linear regression coefficients (**Fig. 2a**; **Fig. 3a** at bottom; **Supplementary Fig. 6b**(iii)), neurons containing one of the five constructs from above (eGFP-only, Htt^{ex1}-Q_{17}-eGFP, Htt^{ex1}-Q_{46}-eGFP, Htt^{ex1}-Q_{72}-eGFP, Htt^{ex1}-Q_{97}-eGFP) and stained with one of four antibodies (3B5H10, MW1, MW7, EM48) were analyzed (17–48 neurons neurons for each unique transfected construct/antibody condition). The eGFP-only construct served as a negative control for the expression of antigen. However, none of the antibodies stained neurons transfected with only eGFP to any detectable level at the photomultiplier tube gain and offset settings used.

To employ an accurate hierarchical Bayesian statistical model, the variance of the data around the regression lines must be consistent across all stages of the hierarchical statistical analysis. This can only be achieved if eGFP fluorescence levels from neurons used for regression coefficient analysis (**Fig. 2a**; **Fig. 3a** at bottom; **Supplementary Fig. 6b**(iii)) and those used for survival analysis (**Fig. 3a** at top; **Supplementary Fig. 6b**(iv)) are normalized to each other.

Thus, all data from **Figure 2a** and **Figure 3a** (at top) have been normalized using mean Q_{46} expression levels for each independent experiment as the normalizing denominator.

Pre-model testing of data:

Multiple forms of regression (including penalized splines) were applied to the antibody vs. htt^{ex1} graphs in **Figure 2a**. Linear regression proved to be as good a fit as other non-linear fits tested. Therefore, it was deemed appropriate to use linear regression in the model presented below.

A review of notation for the mathematical analysis:

- The $A \sim B$ notation should be read as "*A* is distributed as *B*."
- *N(a,b)* means "normal with mean a and variance b."
- Bold face indicates a matrix or vector
- $a \in \mathsf{B}$ is read as "a is a member of set B"
- The notation $\mathbf{1}_{\{ a \in \mathsf{B} \}}$ has a value 0 if $a \notin \mathsf{B}$ and 1 if $a \in \mathsf{B}$.

The model:

Let $h_i(t), i = 1,\ldots,1508$ be the hazard function corresponding to the vector $\mathbf{x}_i$, where $\mathbf{x}_i = (x_{i\,3B5H10}, x_{i\,MW1}, x_{i\,MW7}, x_{i\,EM48})'$ is the level of the four antibodies in neuron i. Let $u_i, i = 1,\ldots,1508$ be the measured level of eGFP or Htt^{ex1}-Q_n-eGFP in neuron i from the survival data in Step 1 of **Figure 3**. G_j is the group of neurons with polyQ length j (eGFP-only (Q_0), Q_{17}, Q_{46}, Q_{72}, Q_{97}). Finally, let $z_{m_{kj}}, i = 1,\ldots,M_{kj}$ (with M_{kj} the number of neurons with

polyQ length j and antibody k measured from the linear regression experiments in **Figure 2a**) be the measured amount of antibody k in neuron m_{jk} with polyQ length j, and let $y_{m_{kj}}, i = 1,\ldots,M_{kj}$ be that neuron's eGFP fluorescence.

Regarding the parameters in the model, α_{kj} is the regression coefficient corresponding to antibody k and polyQ length j, and β_k is the regression (Cox) coefficient corresponding to antibody k. The σ^2s are variances. The model is as follows:

$$h_i(t) = h_0(t)\exp\left\{\sum_{k=1}^{4}\beta_k x_{ik}\right\} \quad (1)$$

$$x_{ik} \sim N\left(\sum_{j=1}^{5}\alpha_{kj}u_i\mathbf{1}_{\{u_i\in \mathsf{G}_j\}}, \sum_{j=1}^{5}\sigma_{kj}^2\mathbf{1}_{\{u_i\in \mathsf{G}_j\}}\right) \quad (2)$$

$$z_{m_{kj}} \sim N\left(\alpha_{kj}y_{m_{kj}}, \sigma_{kj}^2\right) \quad (3)$$

Line (1) is the standard Cox proportional hazards equation for multivariate analysis (**Supplementary Fig. 6b**(v-vi)). Line (2) gives the linear predictions for the four antibody levels based on the measured eGFP fluorescence u_i from each neuron in the survival data from the top of **Figure 3a** (see **Fig. 3b**; **Supplementary Fig. 6b**(iv)). Line (3) is the linear regression analysis applied to **Figure 2a** (**Supplementary Fig. 6b**(ii-iii)). Note that variances (estimation error) appear in both lines (3) and (2) and will propagate through to line (1) during the hierarchical Bayesian analysis. Because no antibody stained neurons containing eGFP-only (Q_0), variance for j = GFP-only (Q_0) is made to be some very small constant such that $\mathbf{x}_i \approx 0$ for j = GFP-only (Q_0).

Because the analysis applied to the hierarchy of Lines (1)–(3) will be Bayesian, a few further steps are needed in the model. Bayesian analysis is a statistical technique that allows one

to infer or update one's "degree of belief" in a hypothesis or parameter value in light of new experimental information. For our purposes, Bayesian analysis starts with an assumed/given "prior" probability distribution for a parameter (say β (Cox coefficient) or α (linear regression coefficient)) of interest. We may be as specific or non-specific as we want with the assumed/given "prior" distribution. Evidence from experiments is then used to update the prior distributions with new probability distributions for the parameter. After a sufficient number of iterative updates, Bayesian analysis produces a "posterior" distribution, which gives the probability distribution for the parameter of interest taking into account all experimental results. Therefore, to perform Bayesian analysis on the model in Lines (1)–(3) above, we set prior distributions for the model's parameters as follows:

$$\beta_k \sim Laplace(0, \sigma_\beta^2) \quad (4)$$

$$\alpha_{kj} \sim N(0, \sigma_\alpha^2) \quad (5)$$

$$\sigma_{kj}^2 \sim Inverse-gamma(a, b) \quad (6)$$

The Laplace prior in Line (4) gives a Bayesian formulation of the LASSO variable selection/shrinkage criterion (see figure caption for **Supplementary Fig. 6b**)[10]. The prior distribution in Line (5) is a less stringent variable shrinkage prior than that specified in Line (4). In practice, we choose σ_α^2 to be very large so the prior on Line (5) has no effect (i.e., the α's don't actually get shrunken). Similarly, a and b in Line (6) are chosen to be very large so that the prior distribution in Line (6) will have no effect.

We add one final layer in the hierarchy. We consider σ_β^2 to be a parameter instead of a hyperparameter. In other words, we did not specify how much to shrink the Cox coefficients

(β's in Line (4)), but rather, we put a vague prior on σ_β^2 and let the data tell us how much to shrink the Cox coefficients (Line (7)).

$$\sigma_\beta^2 \quad \sim \quad Inverse-gamma(a,b) \qquad (7)$$

One last technical consideration was how to represent the baseline hazard function ($h_0(t)$) in Line (1). We model the function nonparametrically using a gamma process prior, which is the standard Bayesian approach since it was introduced by Kalbfleisch[11].

The analysis:

The above model was subjected to 50,000 iterations with three starting positions for an appropriately defined Markov chain. Convergence was achieved for all parameters. As is customary, we disregarded the first 10,000 iterations of the analysis as a "burn-in" period to protect against possible bias caused by starting conditions. The mean of the posterior distributions for each antibody's Cox coefficient (β) are presented in **Table 1**. Since a positive Cox coefficient in our model predicts increased risk of death, the most appropriate analysis of significance for the Cox coefficients is the probability of each antibody's Cox coefficient being less than 0. This significance test is also presented in **Table 1**. **Figure 4** presents the entire posterior distribution results from the Bayesian hierarchical statistical analysis of the four α-htt antibodies.

Concentration-corrected Förster resonance energy transfer (N-FRET)

To measure concentration-corrected FRET by the method of Xia and Lui (N-FRET)[12], neurons were transfected with Htt^{ex1}-Q_n-CFP (donor only), Htt^{ex1}-Q_n-YFP (acceptor only), or both Htt^{ex1}-Q_n-CFP and Htt^{ex1}-Q_n-YFP (FRET sample). Donor or acceptor cross-talk coefficients were

determined from averaging values from 20–40 donor only or acceptor only neurons, respectively. With these donor and acceptor cross-talk coefficients, N-FRET values were then calculated from the FRET samples (10-40 neurons per condition) by taking average per pixel signal for regions of interest encompassing the neuronal cell body. The entire analysis was carried out using the Zeiss FRET Macro available as an add-on software package to the Zeiss LSM 510 system. After acquiring its N-FRET value, the neuron was then subjected to the acceptor photobleach method of FRET[13]. Results from acceptor photobleach FRET were exactly consistent with our N-FRET results (data not shown). Images from **Figure 5a** are representative of donor fluorescence, acceptor fluorescence, and N-FRET corrected values for neurons co-transfected with Htt^{ex1}-Q_n-CFP and Htt^{ex1}-Q_n-YFP. For the donor and acceptor fluorescence images for Htt^{ex1}-Q_{25} in **Figure 5a** (and only for these images), neurons in the micrograph were cropped onto a black background since areas outside the neuron had a grey haze. No identifiable background features or surrounding neurons were cropped out of the picture during the process, and original, uncompressed, uncropped images were retained and are available upon request. N-FRET measurements with identical procedures were also acquired from striatal neurons transfected with Htt^{ex1}-Q_n-mCFP, Htt^{ex1}-Q_n-mYFP, or both constructs (gift from O. Onodera) and from striatal neurons transfected with Htt^{ex1}-Q_n-CyPet, Htt^{ex1}-Q_n-YPet, or both constructs (cloned as per above from CyPet and YPet constructs provided as a gift by P. Daugherty).

Dot blotting with α-oligomer antibody

When we failed to detect α-oligomer staining of striatal neurons transfected with mutant htt^{ex1}, we were concerned that the fixation process or the detergents used during immunocytochemistry may have diminished α-oligomer binding. To test for this possibility, we spotted pre-aggregated

Htt^{ex1}-Q_{53} reagent, used in our AFM experiments (see below), on nitrocellulose. We then subjected the nitrocellulose membrane to the exact same protocol as we use for immunocytochemitry of neurons, including the same fixation conditions, permeabilization conditions, and detergent concentrations. Under this protocol, the α-oligomer antibody positively stained the dot blot.

Filter retardation assays

HEK293 cells transfected with Htt-171-(Q_{17}, Q_{40}, Q_{68}, Q_{89}, or Q_{142})-FLAG plasmids were lysed on ice for 30 min in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM $MgCl_2$, 0.5% (w/v) Nonidet P-40 (NP40), 1 mM EDTA with protein inhibitors PMSF, leupeptin, pepstatin, aprotenin and antipain. After centrifugation (20,000g at 4°C), the pellet with insoluble material was resuspended in 20mM Tris-HCl (pH 8.0), 15 mM $MgCl_2$ and incubated at 37°C for 1 h. A sample from this insoluble material was diluted in 2% SDS and loaded on a cellulose acetate membrane (Osmonics INC Acetate Plus 0.22 μm 142 mm) previously rinsed three times with 2% SDS and placed on a slot-blot apparatus (Amersham slot blot manifold (Hoefer PR 648)). After loading the samples, wells were washed again three times with 0.1% SDS. Finally, the membrane was blotted with α-FLAG and 3B5H10 antibodies.

Aggregation analysis of htt by agarose gel electrophoresis with western blotting

PC12 cells stably expressing a truncated exon1 of htt fused to GFP (Q_{103}; no polyproline region) and under the control of an ecdysone inducible-expression system[14] were grown in six-well plates. Cells were grown in DMEM with 10% horse serum, 5% fetal bovine serum, and Zeocin and G418 as selection markers at 37°C (all reagents from Invitrogen). Expression of htt was

induced with 5 μM ponasterone and harvested at 48 h after induction. Three independent experiments were performed. At 48 h post-transfection, cells were lysed in RIPA buffer (10 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA (pH 8.0), 1% NP40, 0.5% SDS) containing Complete Protease Inhibitor (Roche Diagnostics). *DC* Protein Assay (Bio-Rad) was performed to determine protein concentration. 30 μg lysate was added in a 1:1 ratio to loading buffer without reducing agent (150 mM Tris pH 6.8, 33% glycerol, 1.2% SDS) and loaded without heating onto a 1% agarose gel containing 0.1% SDS and run until the dye front had migrated at least 12 cm to allow for maximum resolution of aggregates from the dye front. The proteins were then semi-dry blotted (Owl HEP-1) onto a PDVF membrane in transfer buffer (192 mM glycine, 25 mM Tris-base, 0.1% SDS, 15% MeOH). This blot was blocked for 1 h in 5% BSA (EMD) in Tris-buffered saline containing 0.1% Tween-20 at room temperature. The blot was the probed with either 3B5H10 or α-GFP antibody (Clontech). Peroxidase-conjugated AffiniPure goat anti-mouse secondary (Jackson ImmunoResearch) was used for 1 h at room temperature. Blots were detected using PICO detection reagent (Pierce).

Atomic force microscopy

GST-Htt^{ex1}-Q_{53} protein was purified as described[15]. Before each experiment, Htt^{ex1}-Q_{53} was centrifuged at 20,000*g* for 30 min at 4^{o}C to remove preexisting aggregates. Htt^{ex1}-Q_{53} (20 μM) was then incubated alone or with 3B5H10 (20 μM) in buffer A (50 mM Tris-HCl, pH 7, 150 mM NaCl, 1 mM DTT). Precission protease (4 units/100 μg fusion protein) (GE Healthcare Biosciences, Piscataway, NJ) was added to cleave off GST and initiate aggregation. Samples were incubated at 37^{o}C and centrifuged at 1,400 rpm for the duration of the experiment. At 1, 5, 8, and 24 h after addition of protease, a 5-μl sample was deposited on freshly cleaved mica (SPI

supplies, West Chester, PA). After 1 min, the substrate was tilted and washed with 200 μl of ultrapure water by allowing the wash to gently flow over the sample. The sample was then dried under a gentle stream of air. For experiments on preformed oligomers or fibrils, solutions of Htt^{ex1}-Q_{53} were allowed to aggregate for ≥ 5 h after the removal of GST. Then, 3B5H10 was added so that the final concentration of Htt^{ex1}-Q_{53} was 20 μM and the ratio of 3B5H10 to Htt^{ex1}-Q_{53} monomer was 1:10, 1:5, or 1:1. These samples were deposited on mica at 0, 1, and 3 h after the addition of 3B5H10 as described earlier.

AFM experiments were performed using a MFP3D scanning probe microscope (Asylum Research, Santa Barbara, CA). For *ex situ* experiments (**Fig. 5d-f; Supplementary Fig. 10a**), images were taken with a silicon cantilever with a nominal spring constant of 40 N/m and resonance frequency of ~300 kHz. Typical imaging parameters were: drive amplitude 150–500 kHz with set points of 0.7–0.8 V, scan frequencies of 2–4 Hz, image resolution 512 by 512 points, and scan size of 5 μm. All experiments were performed in triplicate.

For *in situ* AFM experiments (**Supplementary Fig. 10b** online), solutions containing preformed fibrils of Htt^{ex1}-Q_{53} were allowed to rest on mica until several fibrils were present on the surface. Then, the substrate was washed with buffer A to remove proteins remaining in solution. The deposited fibrils were imaged in clean buffer as a control or in the presence of 2.5 μM 3B5H10. Images were taken with V-shaped oxide-sharpened silicon nitride cantilever with a nominal spring constant of 0.5 N/m. Scan rates were set at 1 Hz with cantilever drive frequencies at ~8–12 kHz.

Statistical tests

P values (one-sided) were determined using the Student (unpaired) t-test for **Figures 5a,d; Supplementary Figure 8**. Spearman's correlation coefficients were used to compare curves from **Figure 5e-g**.

Chemical cross-linking and blotting of htt^{ex1}

GST-Htt^{ex1}-Q_{53} protein was purified as described[15]. GST-Htt^{ex1}-Q_{53} (27 μM) was then incubated with Precission protease (4 units/100 μg fusion protein) (GE Healthcare Biosciences, Piscataway, NJ) to cleave off GST and initiate aggregation. After 30 min of incubation on a shaker (1400 rpm) at 37°C, glutaraldehyde was added to a final concentration of 0.3%. Agitated incubation at 37 °C was continued for 5 min, followed by the addition of glycine in PBS to a final concentration of 0.43 M. Samples were immediately mixed with 4× NuPage sample buffer (Invitrogen) and run at 150V on NuPage Bis-Tris 4–12% gradient gels with MOPS buffer. Transfer onto PVDF was at 30V for 5 h. Blotting dilution for MW7 (supernatant concentrate from Developmental Studies Hybridoma Bank) was 1:2500 and for 3B5H10 (from a starting concentration of 3 mg/ml) was 1:20,000.

Thio-Htt^{ex1}-Q_{39}-His_6 purification

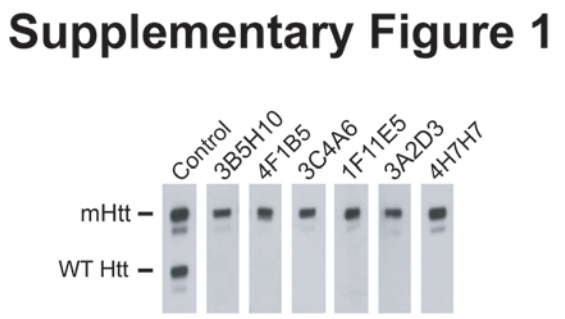
Thio-Htt^{ex1}-Q_{39}-His_6 was produced and purified as described[16]. Protein was produced in BL21(DE3) cells. Cells were pelleted by centrifugation at 2500 rpm for 15 min at 4°C and then resuspended in 5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 8. Extracts were prepared by sonication and spun at 17,000 rpm for 20 min at 4°C. Supernatant was incubated with nickel beads (GE Healthcare Biosciences, Piscataway, NJ), and bound proteins were eluted in 1 M

imidazole. Eluate was further purified by DEAE ion exchange HPLC (Tosoh Bioscience, South San Francisco, CA). Proteins were stored stably for weeks at 4^{o}C in 5 mM Tris-HCl pH 8.

Analytical ultracentrifugation (AU) sedimentation equilibrium analysis

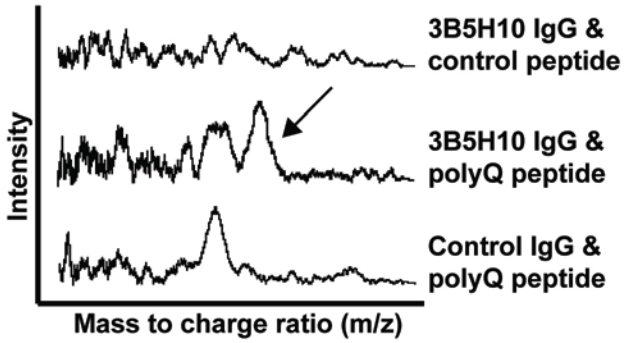
Absorbance (OD_{280}) measurements from AU, taken at 0.001-cm radial increments, were fitted to several different self-association and non-interacting models with the Ultrascan 8 software (http://www.ultrascan.uthscsa.edu/). The buffer density (1.00058 g/ml) was determined with Ultrascan 8. Likewise, the partial specific volume (0.728 ml/g) and extinction coefficient (denatured E_{280nm}=92,880 $mol^{-1}cm^{-1}$) of the complex were determined based on amino acid composition. Confidence intervals for the best-fit [complex]:[dimer of complex] equilibrium constant (K_d = 340 nM) were estimated with Monte Carlo analysis software (10,000 simulations) (95% confidence intervals of 240 and 490 nM).

SUPPLEMENTARY RESULTS



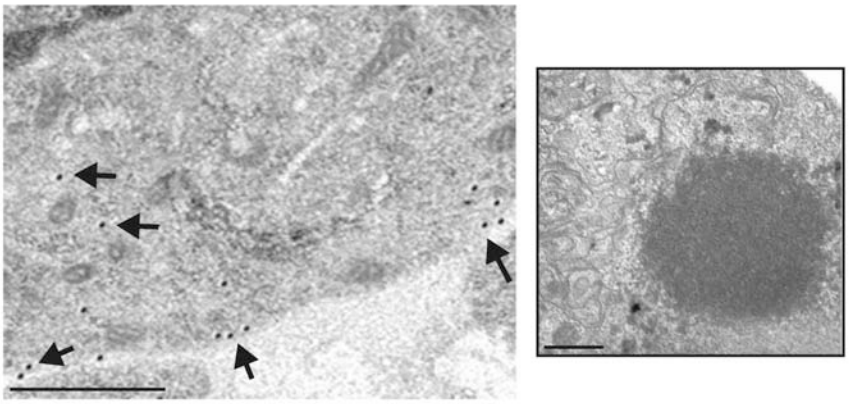
Supplementary Figure 1 Screen for monoclonal antibodies against mutant htt (mHtt). Six of 480 hybridomas secreted mAbs that showed preferential immunoreactivity in an ELISA against mHtt and were tested further. HEK293 protein extracts[13] containing an N-terminal 480-amino acid fragment of htt with either Q_{17} (WT Htt) or Q_{68} (mHtt) were combined and blotted with each of the six mAbs. The 4C8 antibody (control), which recognizes a common epitope (amino acids 443–457), demonstrates comparable loading of wt-htt and mHtt.

Supplementary Figure 2



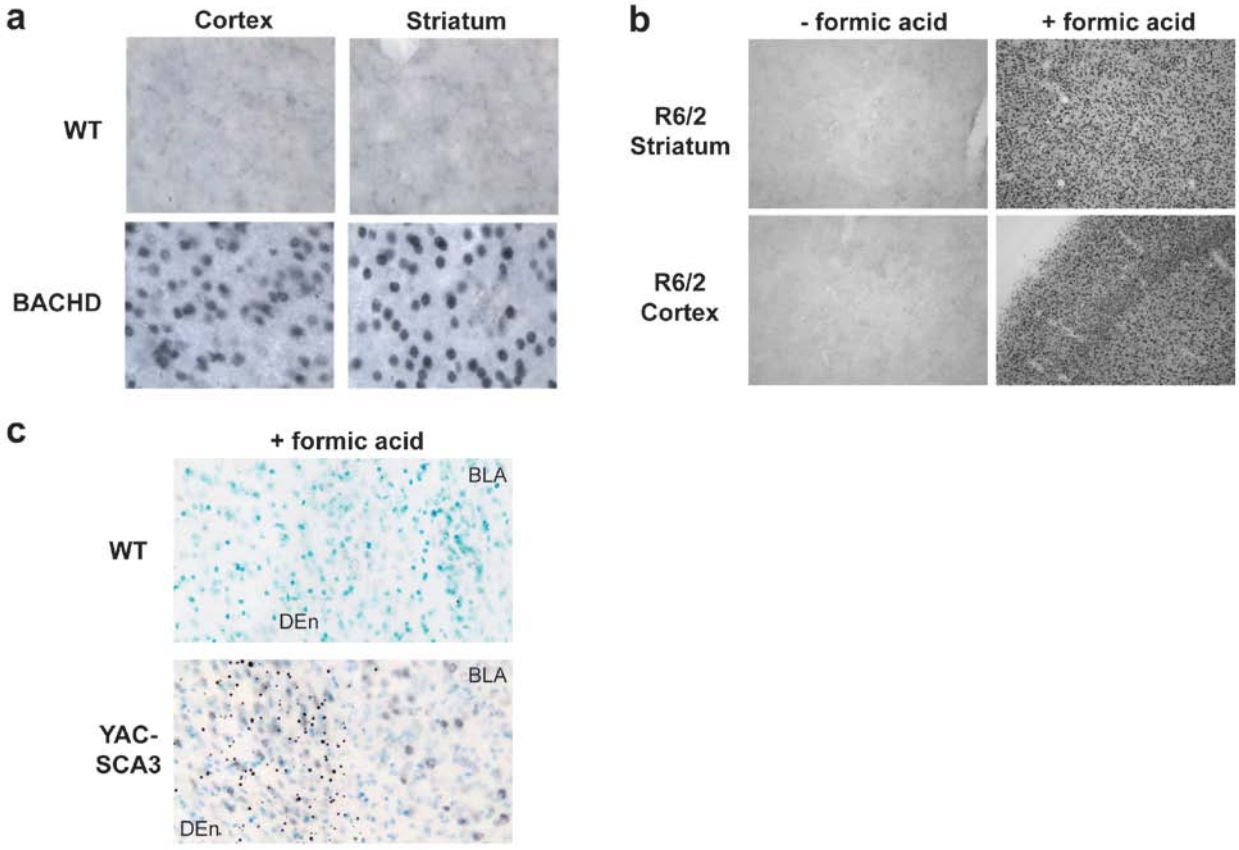
Supplementary Figure 2 3B5H10 binds pure polyQ peptide. Analysis of interaction of 3B5H10 IgG with polyQ by SELDI-TOF-MS. PolyQ (K2 Q_{39}K2) or PACAP peptide (basic peptide as a control) were incubated on a PS10 Protein Chip Array (Bio-Rad) and control IgG antibody or 3B5H10 IgG were added. Arrays were loaded into a Protein Chip SELDI System (Bio-Rad) for data collection. Analysis of intensity peak patterns shows a unique peak (see arrow) that corresponds to the mass-to-charge (m/z) ratio expected for a protein complex formed by the polyQ peptide and 3B5H10 IgG antibody.

Supplementary Figure 3



Supplementary Figure 3 3B5H10 immunogold staining of striatal neurons transfected with Htt^{ex1}-Q_{97}-eGFP. By electron microscopy, 3B5H10 labels diffuse mHtt in the cytoplasm with little staining of htt in organelles and no staining of IBs. Scale bar=1 μm.

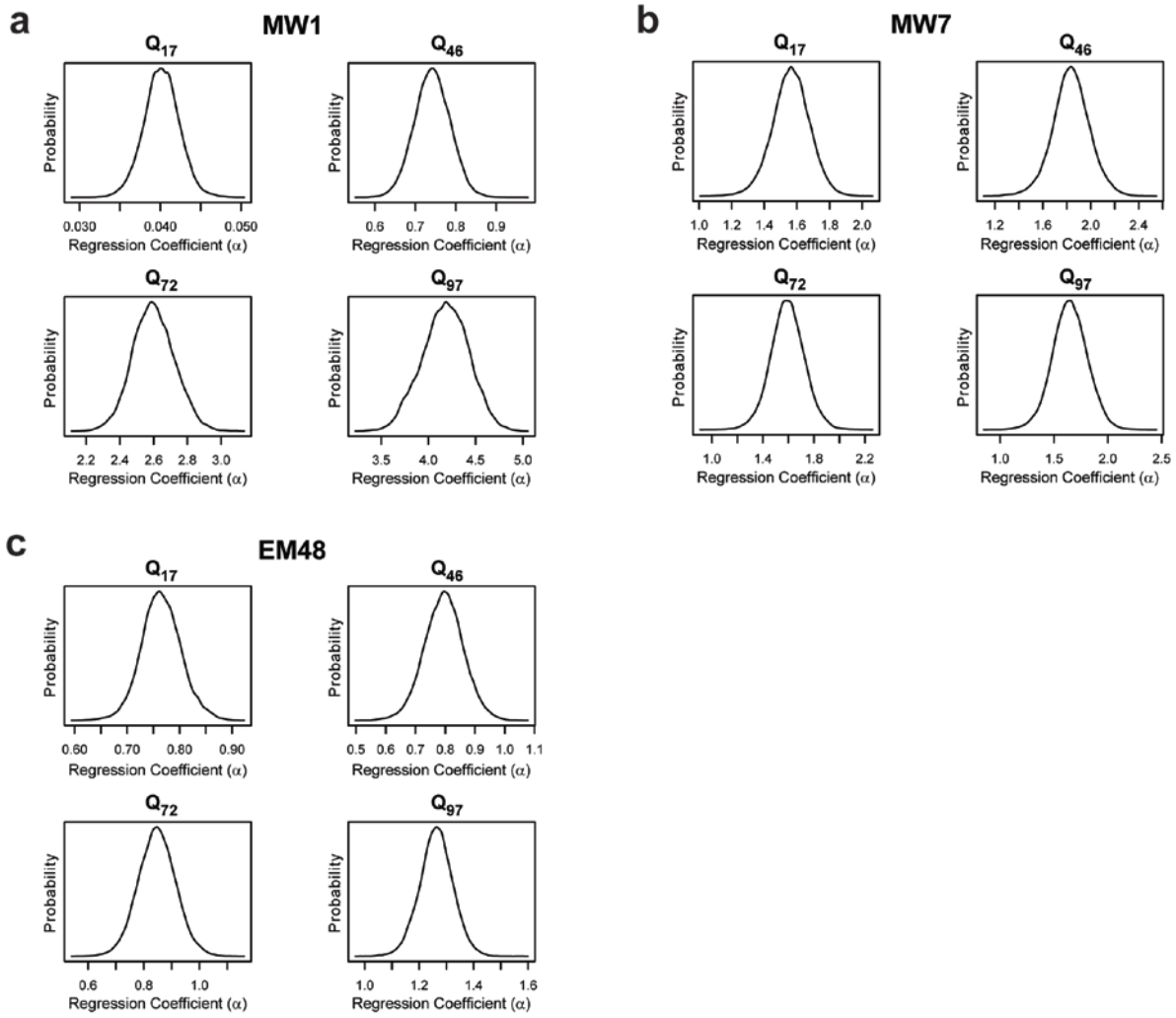
Supplementary Figure 4



Supplementary Figure 4 3B5H10 staining of brain tissue from animal models of polyQ disease. (**a**) Striatal and cortical brain sections from 12-month-old BACHD mice and strain control mice were stained with 3B5H10. BACHD mice have a full-length mHtt (Q_{97}) human transgene with the human endogenous promoter. By 12 months, large inclusions in the cortex and tiny aggregates in the striatum are detectable in this mouse model with EM48 antibody[17]. However, 3B5H10 fails to recognize these inclusions, instead staining diffuse mHtt. When aggressive antigen retrieval using formic acid was employed, 3B5H10 staining of BACHD tissue revealed more prominent staining of neuropil aggregates (data not shown). (**b**) Striatal and cortical brain sections from 15-week-old R6/2 mice, which highly overexpress the exon1 fragment of human

mHtt (~Q_{150}),[18] were stained with 3B5H10 with or without strong antigen retrieval (30 total min of 90% formic acid treatment). By 15 weeks of age, essentially all mHtt is aggregated into inclusion bodies (IBs) in this mouse model. 3B5H10 fails to recognize these aggregates unless strong antigen retrieval unmasks otherwise buried epitopes in the IBs. (**c**) Dorsal endopiriform claustrum (DEn) and basolateral amygdala (BLA) brain sections from adult YAC-SCA3 mice and strain control mice were stained with 3B5H10. YAC-SCA3 mice have a full-length mutant ataxin-3 (Q_{84}) human transgene with the human endogenous promoter[19]. Sections were treated with strong formic acid antigen retrieval and counterstained with thionin (blue). 3B5H10 fails to stain littermate control brains but robustly stains YAC-SCA3 sections, including both diffuse ataxin-3 and aggregates. Again, staining of aggregates is likely attributable to unmasking of normally-buried epitopes by formic acid treatment.

Supplementary Figure 5



Supplementary Figure 5 Bayesian regression coefficient (α) probability distributions for graphs of Htt^{ex1}-eGFP versus α-htt antibody levels (from **Figure 2a**)**.**

(**a**) MW1

(**b**) MW7

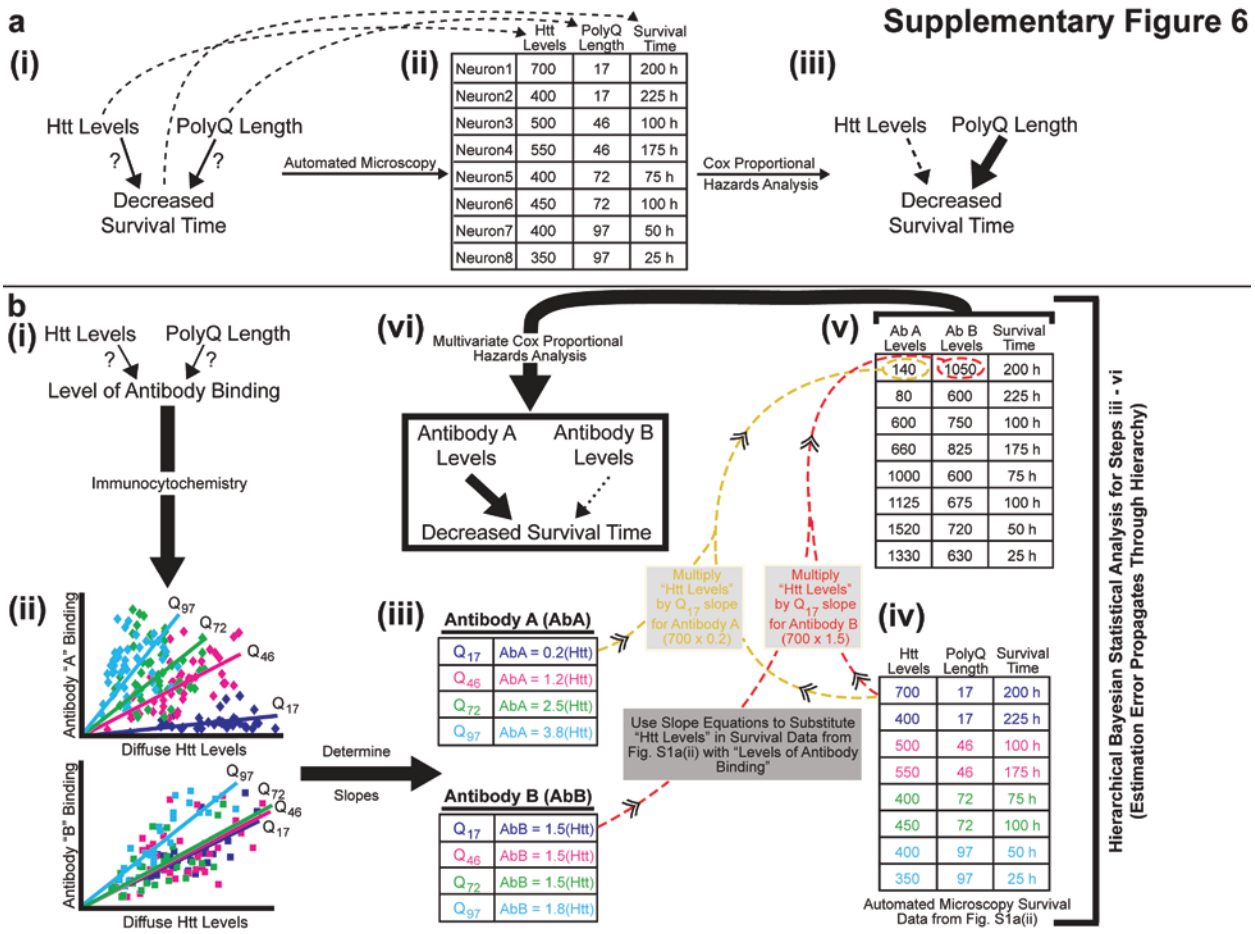
(**c**) EM48

From the above plots, it's clear that MW1 recognizes htt in a polyQ-dependent manner. Why, then, does MW1 fail to predict neurodegeneration in a multivariate Cox analysis that includes 3B5H10?

A key part of our Cox analysis is the ability to simultaneously compare the prognostic value of multiple risk factors (staining by each antibody). In this "multivariate" analysis, the epitope recognized by 3B5H10 becomes the only positive predictor of neurodegeneration in our striatal model of HD. This is because the model is built to find the minimum combination of antibodies that can maximally explain variations in survival. In other words, if there were information encoded in MW1 staining that did not overlap with information encoded in 3B5H10 staining and this non-overlapping information explained neurodegeneration, then MW1 would become a positive predictor of toxicity. Therefore, the interpretation of 3B5H10 being the only positive predictor of death in multivariate analysis is that 3B5H10 independently contains all the necessary information to account for toxicity.

When doing verification testing for the Cox model, we discovered that MW1 staining increases too fast as the polyQ stretch increases to account for the more tempered increase in neuron death rates as the polyQ stretch expands. The multivariate Cox analysis is likely picking up on this discordance.

Our multivariate analysis does not preclude MW1 as a predictor of toxicity. In fact, if we drop 3B5H10 out of the Cox analysis and simply compare MW1, MW7, and EM48, then MW1 becomes a positive predictor of death. This is because MW1 staining contains information that explains survival which is not represented among the remaining antibodies (i.e., MW7 and EM48) to which it is being compared.



Supplementary Figure 6 Method for quantitatively determining which htt species best predicts neurodegeneration.

(**a**) Assessing quantitative relationships between neuronal survival time and the risk factors of polyQ length and diffuse htt levels.

Previously, we combined automated microscopy with Cox proportional hazards analysis to determine how well particular risk factors predict an outcome of interest in a primary striatal neuron model of HD[2,9]. The system was used here to determine how well polyQ length or diffuse htt levels predict striatal neuron survival.

To begin, we diagram the relationships we were interested in quantifying (**Supplementary Fig. 6a**(i)). Do "Htt Levels" or "PolyQ Length" determine "Survival Time?" If so, what are their respective contributions? Next, we utilized a previously described automated microscope system to individually track thousands of neurons over time[9]**,** quantifying the risk factors of interest (PolyQ Length and Htt Levels) in each neuron and then recording the outcome of interest for that neuron (Survival Time: the time point at which each neuron dies[2]). Because our previous work showed that levels of diffuse mHtt predicted neuronal degeneration and were sufficient to cause neuronal death without IB formation[2], we focused our analysis of the present study on neurons that *did not* form an IB; the risk factor "Htt Levels" therefore refers to levels of diffuse htt.

We next measured "Htt Levels" (at ~24 h after transfection of htt into neurons) and "Survival Time" for each neuron from the images we collected, as described[2,9]. We recorded these values, along with the "PolyQ Length" of the construct transfected into each neuron, in a table (**Supplementary Fig. 6a**(ii) – values are illustrative only). We previously validated the measurements of the fluorescence of the eGFP tag fused to htt^{ex1} as an accurate estimate of htt^{ex1} levels ("Htt Levels") in neurons[2]. We have also shown previously that intraneuronal htt levels 24–48 h after transfection among neurons that do not form IBs are linearly related to average intraneuronal levels of htt over the course of the experiment and can be used as a correlate of those levels for the purpose of Cox analysis. Bivariate Cox proportional hazards analysis was then applied to these data to determine whether any of the measured variables (e.g., "Htt Levels", "PolyQ Length") significantly determine the outcome (e.g., "Survival Time") (**Supplementary Fig. 6a**(iii)) – the stronger the association, the wider the arrow).

Univariate Cox analysis of our automated microscopy data previously revealed "Htt Levels" as a major negative predictor of "Survival Time" for populations of neurons with the same mutant "PolyQ Length"[2]. Therefore, we considered the possibility that "Htt Levels" are the major negative predictor of neuronal survival for any neuron with mHtt, regardless of "PolyQ Length." To test this possibility, we reanalyzed our automated microscopy data combining neurons from all "PolyQ Length" categories and performing univariate Cox analysis with "Htt Levels" as the sole risk factor. This analysis revealed that "Htt Levels" alone *do not* significantly predict "Survival Time," because neurons transfected with htt containing longer polyQ expansions have shorter "Survival Times" but lower steady state "Htt Levels."

These findings motivated us to look for a better unifying predictor (risk factor) of neuronal "Survival Time" than "PolyQ Length" or "Htt Levels." We reasoned that an epitope of a malfolded species of htt, whose abundance presumably is related to "PolyQ Length" and "Htt Levels," but which might be more specifically measured by some conformation-specific α-htt antibody, could be just such a unifying predictor of neurodegeneration (see **Supplementary Fig. 6b**).

(**b**) Assessing quantitative relationships between neuronal "Survival Time" and the risk factors of 3B5H10, MW1, EM48, and MW7 binding to diffuse mHtt *in situ*.

Because antibody staining requires the fixation of neurons, we could not directly utilize the longitudinal tracking methodology described in **Supplementary Figure 6a** to determine how well certain epitopes recognized by α-htt antibodies predict neuronal survival. As a result, we related antibody staining *in situ* to neuronal "Survival Time" by the approach below (**Supplementary Figure 6b** reads in counterclockwise manner).

To begin, we ask, "How do the risk factors in **Supplementary Figure 6a** ("Htt Levels" or "PolyQ Length") relate to a new outcome of interest ("Level of Antibody Binding")?" (**Supplementary Fig. 6b**(i) – the risk factor "Htt Levels" refers to levels of diffuse htt as before). To answer this question, we performed quantitative immunocytochemistry (**Supplementary Fig. 6b**(ii)). Neurons transfected with Htt^{ex1}-(Q_n)-eGFP constructs are fixed (at the same post-transfection time (~24 h) as "Htt Levels" are measured for the separate set of neurons undergoing the automated microscopy survival experiment in **Supplementary Figure 6a**) and labeled with an α-htt antibody and a fluorescently labeled secondary antibody (Cy5 or Alexa647). For each neuron, the fluorescent signal from the secondary antibody (*y* axis) and the fluorescent signal from the eGFP tag attached to htt^{ex1} (*x* axis) are recorded and plotted[2] (**Supplementary Fig. 6b**(ii); see **Fig. 2a** for actual plots of the four antibodies tested in this study). Linear regression analysis of these data yield equations with slopes that quantitatively relate "Htt Levels" to "Level of Antibody Binding" *in situ* for versions of htt^{ex1} with different polyQ lengths (**Supplementary Fig. 6b**(iii)). Importantly, multiple types of non-linear regression curves were fitted to the data, including penalized splines, but linear regression with a zero intercept proved to be as good as other models tested.

With the linear regression "calibration" equations from **Supplementary Figure 6b**(iii), the "Level of Antibody Binding" for any live neuron can be estimated from its "Htt Levels" and "PolyQ Length." Thus, we are able to estimate how much antibody binding would have occurred if the neurons from the automated microscopy experiments in **Supplementary Figure 6a**(ii) had been stained. First, the survival data from **Supplementary Figure 6a**(ii) are recopied as **Supplementary Figure 6b**(iv). Next, the values for "Htt Levels" in each neuron, whose survival has been measured, are replaced with estimates of the amount of each epitope in that neuron by

multiplying "Htt Levels" by the linear regression "calibration" coefficients from **Supplementary Figure 6b**(iii). Thus, the resulting table in **Supplementary Figure 6b**(v) contains the survival outcome measure for each neuron and the new risk factors whose predictive value we wish to test—the amounts of binding for each antibody in each neuron. With the new survival data table in **Supplementary Figure 6b**(v) we then used multivariate Cox analysis to determine whether binding by any of these antibodies predicted decreased "Survival Time" (**Supplementary Fig. 6b(**vi)).

In implementing the methodology described in **Supplementary Figure 6b** we considered four statistical concerns: (1) estimation error, (2) the potential collinearity of risk factors submitted for Cox analysis, (3) the Cox model assumption of proportional hazards, (4) the Cox model assumption of a loglinear relationship between risk factors and outcome.

(**1**) Because the antibody risk factors in the survival data table from **Supplementary Figure 6b**(v) are *estimated* based upon linear regression "calibration" equations, there is an inherent "estimation error" embedded in these values that we must account for during Cox analysis. Indeed, the incorporation of linear regression within a Cox proportional hazards model and the need to propagate "estimation error" through the entire analysis strongly suggest the use of a hierarchical Bayesian statistical approach (see **Supplementary Methods** for schematic of hierarchy)[11,20]. Because we applied a hierarchical Bayesian statistical approach, at the bottom of the hierarchy the linear regression coefficients (**Supplementary Fig. 6b**(iii)) are calculated with an associated variance, and they are named Bayesian regression coefficients (α) (**Fig. 2b**, **Supplementary Fig. 5** online). At the next level up on the hierarchy, the Bayesian regression coefficients (α) are used to calculate antibody binding values for survival data table in

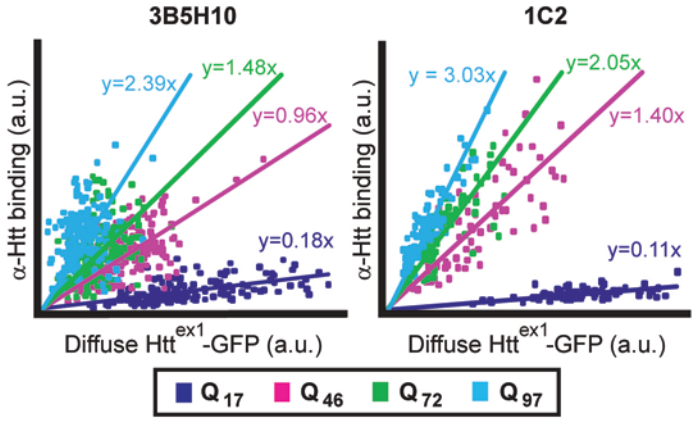
Supplementary Figure 6b(v). At the top level of the hierarchy, estimated antibody binding values (and their associated variance) are then used in Cox analysis (**Supplementary Fig. 6b**(vi)). The Bayesian approach allows steps iii–vi in **Supplementary Figure 6b** to be consolidated into a single analytical step with the statistical significance of the final results (see **Table 1** in main-text) accounting for all "estimation error" in the methodology. Because regression coefficients were calculated using Bayesian analysis, the linear regression equations that appear in **Supplementary Figure 6b**(iii) are illustrative only.

(**2**) A second statistical consideration for the Cox model arises when risk factors are simultaneously evaluated (multivariate analysis); in such analyses, it is assumed that the submitted risk factors are at least partly independent of each other (i.e., not completely collinear). When many highly collinear risk factors are simultaneously entered into the Cox model in a conventional manner, the model breaks down, and results lose biological meaning. To address this concern, we turned to a variable selection technique called LASSO, which helps shrink a pool of risk factors (in our case, all four antibodies) to the minimum number needed to explain an outcome (decreased "Survival Time"). The LASSO technique shrinks the contributions of risk factors excessively collinear with other risk factors, leaving sufficiently independent risk factors as the major weights for multivariate Cox analysis. Thus, LASSO and its Bayesian formulation, the Laplace prior, seek the most parsimonious representation of predictors while still retaining prediction accuracy[10] (see **Supplementary Methods** online for how the Laplace prior was integrated into our hierarchical Bayesian statistical analysis).

(**3**) The third assumption of the Cox model deals with proportional hazards; this assumption states that changes in levels of the risk factors ("Levels of Antibody Staining" for each antibody) should produce proportionate changes in the outcome (decreased "Survival Time"), independent of time. We validated this assumption using graphical approaches[21].

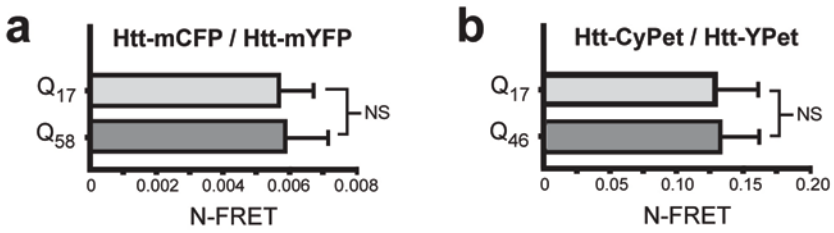
(**4**) And finally a fourth assumption of the Cox model involves the presumed log-linear relationship between risk factors ("Levels of Antibody Staining" for each antibody) and outcome (decreased "Survival Time"). This assumption derives from the functional form of the equation used in Cox models (ln[risk of death] ~ β * antibody level; where β is the Cox coefficient). It is possible to test whether more complicated relationships between risk factors and outcome exist by transforming Cox coefficients (β) into more complicated functional forms called penalized splines and then repeating Cox analysis. Using this methodology[22], we ascertained results consistent with the results from our Cox analysis employing the loglinear assumption.

Supplementary Figure 7



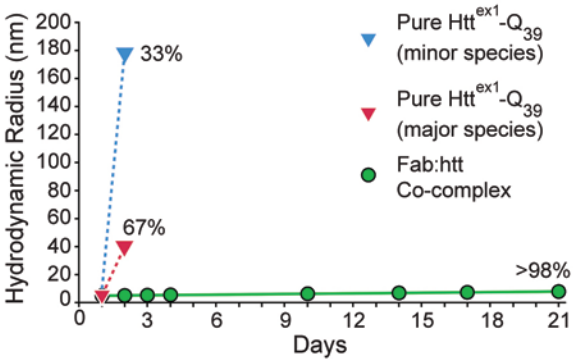
Supplementary Figure 7 Recognition of htt^{ex1} by monoclonal antibodies 3B5H10 and 1C2. Striatal neurons transfected with Htt^{ex1}-(Q_{17},Q_{46},Q_{72}, or Q_{97})-eGFP were fixed at 24 h and subjected to immunocytochemistry with 3B5H10 or 1C2. Fluorescence was measured by confocal microscopy (at least 65 neurons per condition). For this analysis, only neurons without IBs were measured. Linear regression slopes numerically demonstrate that the relationship between htt levels and antibody staining are subtly different for 3B5H10 and 1C2. However, these differences are not large enough to perform reliable survival analysis.

Supplementary Figure 8



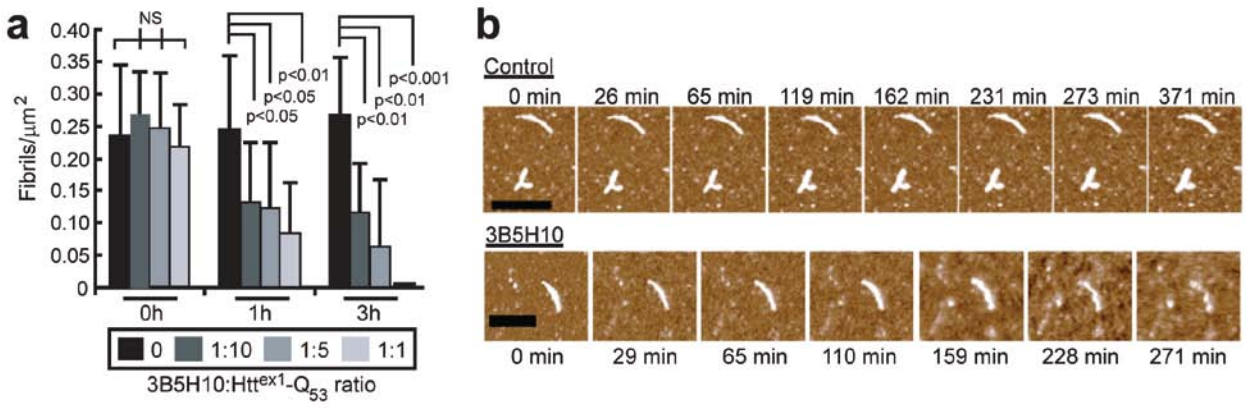
Supplementary Figure 8 Alternative constructs fail to reveal an elevated N-FRET signal for diffuse mHtt in striatal neurons. (**a**) N-FRET was measured in striatal neurons between Htt^{ex1}-Q_{17}-mCFP and Htt^{ex1}-Q_{17}-mYFP and between Htt^{ex1}-Q_{58}-mCFP and Htt^{ex1}-Q_{58}-mYFP, constructs that had previously exhibited an elevated FRET signal for mHtt in a cell line. N-FRET signal from regions of neurons with diffuse mHtt was not significantly different than the signal from neurons with wild-type htt. N-FRET (mean$\pm$s.d.) values are from 32–34 neurons per condition. (**b**) N-FRET was measured in striatal neurons between Htt^{ex1}-Q_{17}-CyPet and Htt^{ex1}-Q_{17}-YPet and between Htt^{ex1}-Q_{46}-CyPet and Htt^{ex1}-Q_{46}-YPet. CyPet and YPet have been evolutionarily optimized for maximal FRET efficiency. N-FRET signal from regions of neurons with diffuse mHtt was not significantly different than the signal from neurons with wild-type htt. N-FRET (mean$\pm$s.d.) values are from 25 neurons per condition.

Supplementary Figure 9



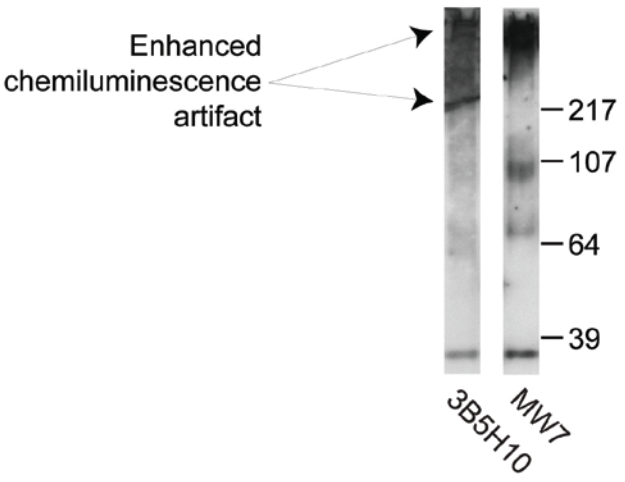
Supplementary Figure 9 3B5H10 prevents mHtt aggregation by dynamic light scattering (DLS). Freshly purified Thio-Htt^{ex1}-Q_{39}-His_6 was incubated alone or with 3B5H10 Fab. After 2 days, Thio-Htt^{ex1}-Q_{39}-His_6 aggregated into two predominant species. After 3 days, the aggregates were too large to monitor with DLS. By contrast, the addition of 3B5H10 Fab to the solution of Thio-Htt^{ex1}-Q_{39}-His_6 kept particle size small, homogenous, and stable for at least 1 month.

Supplementary Figure 10



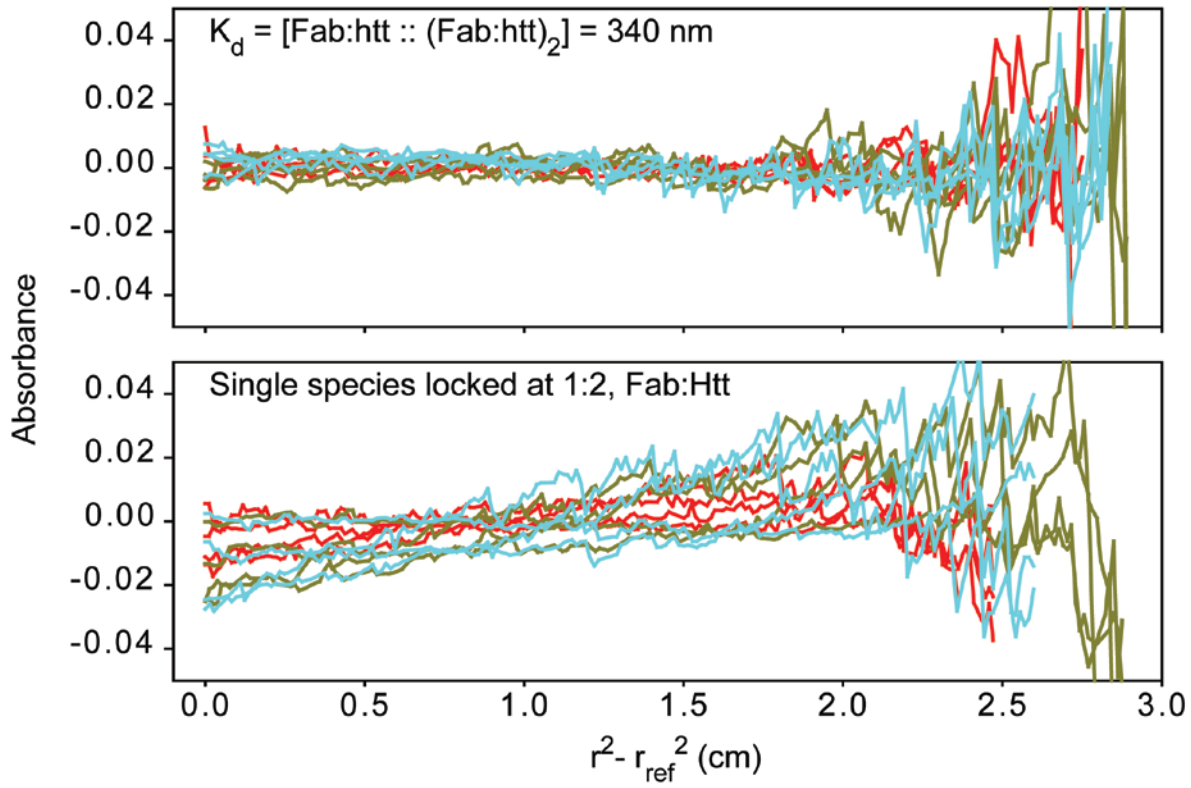
Supplementary Figure 10 3B5H10 dissolves pre-aggregated Htt^{ex1}-Q_{53} fibrils. (**a**) Pre-aggregated fibrils of Htt^{ex1}-Q_{53} were monitored with AFM in the absence of 3B5H10 (■) or with it added (3B5H10:Htt^{ex1}-Q_{53} molar ratios: ■=1:10, ■1:5, ■1:1). Compared with buffer control, the presence of 3B5H10 led to a significant dose- and time-dependent reduction in the number of fibrils. (**b**) Time-dependent dissolution of individual pre-aggregated fibrils of Htt^{ex1}-Q_{53} incubated with buffer or 2.5 μM 3B5H10. Scale bars=500 nm.

Supplementary Figure 11



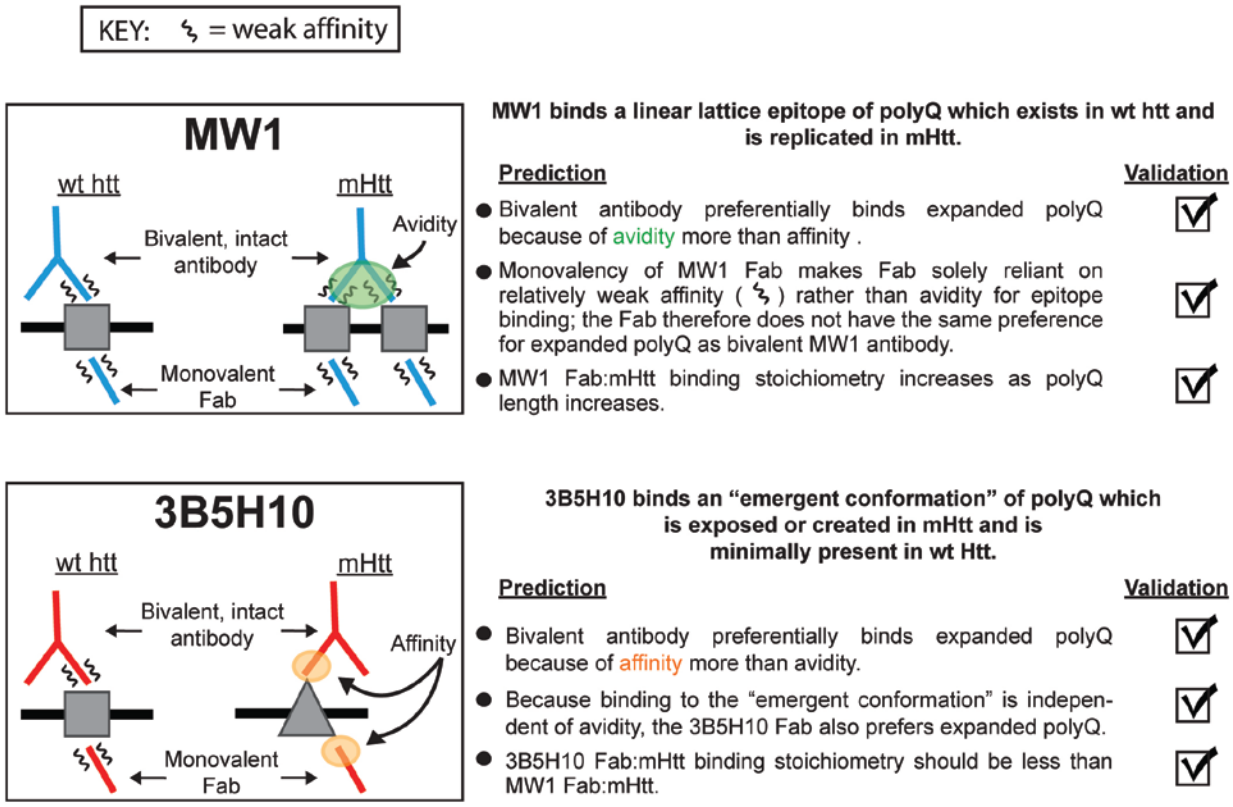
Supplementary Figure 11 Full western blot of a Htt^{ex1}-Q_{53} solution containing monomers and small oligomers chemically cross-linked and probed with MW7 or 3B5H10. The 3B5H10 lane demonstrates an enhanced chemoluminescence (ECL) artifact towards the top of the gel, but the MW7 lane (from a sister blot) does not contain the ECL artifact. Gel ladder measured in kDa.

Supplementary Figure 12



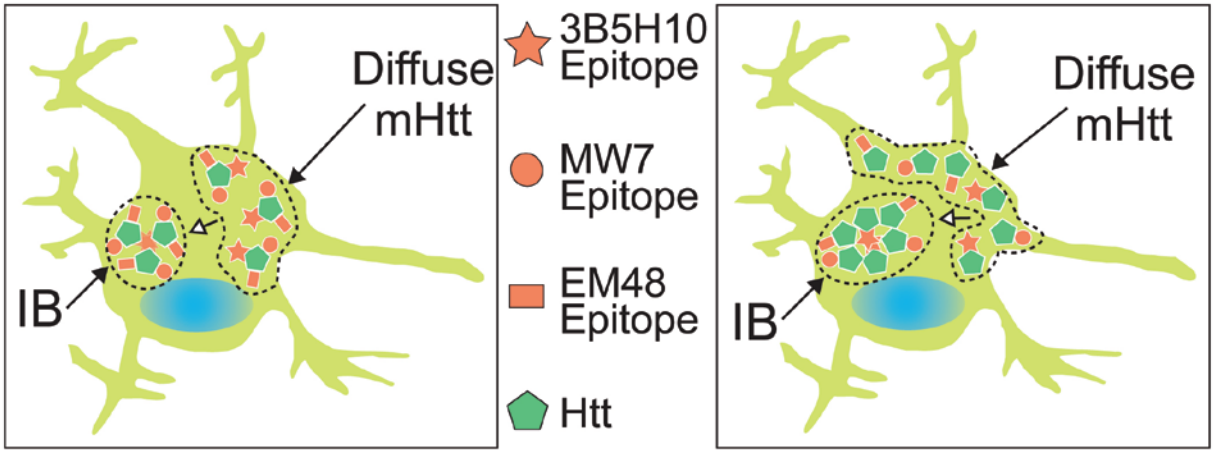
Supplementary Figure 12 Equilibrium sedimentation analytical ultracentrifugation (AU) analysis fails to support a model in which 3B5H10 Fab binds a dimer of mHtt. The AU dataset from **Figure 6d** was fit to two models. The first model, in which the size of the Fab:htt complex was allowed to float, was best explained by 3B5H10 Fab binding to htt^{ex1} in a 1:1 ratio. At high concentrations, the complex dimerizes. The even distribution of residuals for the first model suggests no bias in the fit (top graph). In contrast, for the second model, we fixed the mass of the complex to represent 3B5H10 Fab bound to a dimer of htt^{ex1} (1:2 ratio). The residuals for this fit demonstrate significant skew, suggesting the fit is inappropriate (bottom graph).

Supplementary Figure 13



Supplementary Figure 13 3B5H10 and MW1 recognize mHtt via different mechanisms. MW1 recognizes expanded polyQ as a "linear lattice," in which the antibody binds weakly to a relatively unstructured epitope of wt polyQ (represented by squares on left side). As the polyQ stretch expands into a length associated with disease (mHtt), the unstructured epitope repeats (represented by two squares). Since antibodies are bivalent, the presence of two epitopes in tandem results in increased avidity-based binding by MW1. In contrast, 3B5H10 recognizes expanded polyQ as an "emergent conformation." In this model, the polyQ structure recognized by 3B5H10 (represented by triangles on the left side) is minimally present in wt-htt and emerges as the polyQ stretch expands into the mutant range. The predictions of these two models and the data that support these predictions are presented at right.

Supplementary Figure 14



Supplementary Figure 14 IBs may be neuroprotective by sequestering and masking mHtt epitopes that predict neuronal death. The epitope recognized by 3B5H10 predicts neuronal death and is available only in diffuse forms of mHtt. In contrast, the epitopes recognized by EM48 and MW7 are poor predictors of neurodegeneration and available in diffuse mHtt as well as IBs. Given that IB formation can be a beneficial coping response to mHtt[2], findings in this study suggest that IBs might predict neuronal survival, in part, by masking or refolding epitopes of mHtt that predict neurodegeneration. All epitopes (in this case, the epitopes recognized by 3B5H10, MW7, and EM48) may simultaneously exist on the same molecule of mHtt (left panel). Aggregation into IBs could (1) orient each 3B5H10 epitope toward the center of the IB, which could mask it from further interaction, (2) produce the staining pattern seen in **Figure 1e**, (3) account for the improved survival associated with IB formation[2], and (4) leave the epitopes that do not predict neurodegeneration (e.g., those recognized by MW7, EM48) available on the IB surface for antibody staining. Alternatively, each epitope may exist on a separate molecule of mHtt. As monomers of htt associate with each other, the 3B5H10 epitope (right panel) is masked and an identical outcome occurs as in the left panel. In either case (left or right panel),

aggregation of htt into IBs could neutralize the 3B5H10 epitope by burying it so that it is inaccessible to 3B5H10 or other intracellular targets or by refolding the polyQ stretch so that it no longer contains the epitope recognized by 3B5H10.

SUPPLEMENTARY REFERENCES

1. Kazantsev, A., Preisinger, E., Dranovsky, A., Goldgaber, D. & Housman, D. Insoluble detergent-resistant aggregates form between pathological and nonpathological lengths of polyglutamine in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11404–11409 (1999).
2. Arrasate, M., Mitra, S., Schweitzer, E.S., Segal, M.R. & Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805–810 (2004).
3. Saudou, F., Finkbeiner, S., Devys, D. & Greenberg, M.E. Huntingtin acts in the nucleus to induce apoptosis, but death does not correlate with the formation of intranuclear inclusions. *Cell* **95**, 55–66 (1998).
4. Finkbeiner, S., *et al.* CREB: A major mediator of neuronal neurotrophin responses. *Neuron* **19**, 1031–1047 (1997).
5. Brooks, E., Arrasate, M., Cheung, K. & Finkbeiner, S.M. Using antibodies to analyze polyglutamine stretches. *Methods Mol. Biol.* **277**, 103–128 (2004).
6. Ko, J., Ou, S. & Patterson, P.H. New anti-huntingtin monoclonal antibodies: Implications for huntingtin conformation and its binding proteins. *Brain Res. Bull.* **56**, 319–329 (2001).
7. Kayed, R., *et al.* Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486–489 (2003).
8. Osmand, A.P., Berthelier, V. & Wetzel, R. Imaging polyglutamine deposits in brain tissue. *Methods Enzymol.* **412**, 106–122 (2006).
9. Arrasate, M. & Finkbeiner, S. Automated microscope system for determining factors that predict neuronal fate. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3840–3845 (2005).
10. Tibshirani, R. Regression shrinkage and selection via the lasso. in *J. Royal Statist. Soc. B*, Vol. 58 267–288 (Blackwell Publishing for the Royal Statistical Society, 1996).
11. Kalbfleisch, J.D. Non-parametric Bayesian analysis of survival data. *J. Royal Statist. Soc. B* **40**, 214–221 (1978).
12. Xia, Z. & Liu, Y. Reliable and global measurement of fluorescence resonance energy transfer using fluorescence microscopes. *Biophys. J.* **81**, 2395–2402 (2001).

13. Karpova, T.S., *et al.* Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal miscroscopy and a single laser. *J. Microscopy* **209**, 56–70 (2003).
14. Apostol, B.L., *et al.* A cell-based assay for aggregation inhibitors as therapeutics of polyglutamine-repeat disease and validation in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5950–5955 (2003).
15. Muchowski, P.J., *et al.* Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7841–7846 (2000).
16. Bennett, M.J., *et al.* A linear lattice model for polyglutamine in CAG-expansion diseases. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11634–11639 (2002).
17. Gray, M., *et al.* Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. *J. Neurosci.* **28**, 6182–6195 (2008).
18. Mangiarini, L., *et al.* Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* **87**, 493–506 (1996).
19. Cemal, C.K., *et al.* YAC transgenic mice carrying pathological alleles of the MJD1 locus exhibit a mild and slowly progressive cerebellar deficit. *Hum. Mol. Genet.* **11**, 1075–1094 (2002).
20. Bustamante, C.D., *et al.* The cost of inbreeding in *Aradposis*. *Nature* **416**, 531–534 (2002).
21. Kleinbaum, D.G. & Klein, M. *Survival analysis: A self-learning text*, (Springer-Verlag, New York, 1996).
22. Kauermann, G. Penalized spline smoothing in multivariable survival models with varying coefficients. *Comp. Stat. Data Anal.* **49**, 169–186 (2004).