Spatiotemporal Proteomic Profiling of Huntington’s Disease Inclusions Reveals Widespread Loss of Protein Function

Authors
Fabian Hosp, Sara Gutiérrez-Ángel, Martin H. Schaefer, ..., Rüdiger Klein, Irina Dudanova, Matthias Mann

Correspondence
idudanova@neuro.mpg.de (I.D.), mmann@biochem.mpg.de (M.M.)

In Brief
Hosp et al. use quantitative proteomics to describe the soluble and insoluble proteome of several brain regions in a mouse model of Huntington’s disease at various stages of disease progression. Their findings suggest that widespread sequestration of proteins into mutant huntingtin inclusion bodies contributes to HD pathogenesis.

Data and Software Availability
PXD004973

Highlights
- Spatiotemporally resolved brain proteome of wild-type and HD mice
- Quantitative characterization of huntingtin inclusion bodies in vivo
- Sequestration correlates with protein expression levels and specific sequence features
- Resupplying sequestered proteins ameliorates HTT-induced toxicity and inclusion size
Spatiotemporal Proteomic Profiling of Huntington’s Disease Inclusions Reveals Widespread Loss of Protein Function

Fabian Hosp,1,9 Sara Gutiérrez-Ángel,2,3 Martin H. Schaefer,3,4 Jürgen Cox,5 Felix Meissner,9 Mark S. Hipp,7,8 F.-Ulrich Hartl,7,8 Rüdiger Klein,2,8 Irina Dudanova,2,9 Martin H. Schaefer,3,4 Jürgen Cox,2,8 Irina Dudanova,2,9 F.-Ulrich Hartl,7,8 Rüdiger Klein,2,8 Irina Dudanova,2,9 and Matthias Mann1,10,*

1Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
2Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
3Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
4Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
5Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
6Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
7Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
8Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
9Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
10Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

*Correspondence: idudanova@neuro.mpg.de (I.D.), mmann@biochem.mpg.de (M.M.)

https://doi.org/10.1016/j.celrep.2017.10.097

SUMMARY

Aggregation of polyglutamine-expanded huntingtin exon 1 (HttEx1) in Huntington’s disease (HD) proceeds from soluble oligomers to late-stage inclusions. The nature of the aggregates and how they lead to neuronal dysfunction is not well understood. We employed mass spectrometry (MS)-based quantitative proteomics to dissect spatiotemporal mechanisms of neurodegeneration using the R6/2 mouse model of HD. Extensive remodeling of the soluble brain proteome correlated with insoluble aggregate formation during disease progression. In-depth and quantitative characterization of the aggregates uncovered an unprecedented complexity of several hundred proteins. Sequestration to aggregates depended on protein expression levels and sequence features such as low-complexity regions or coiled-coil domains. In a cell-based HD model, overexpression of a subset of the sequestered proteins in most cases rescued viability and reduced aggregate size. Our spatiotemporally resolved proteome resource of HD progression indicates that widespread loss of cellular protein function contributes to aggregate-mediated toxicity.

INTRODUCTION

Cellular environment is characterized by very high protein concentrations, increasing the danger of aggregation. Thus, maintenance of protein solubility is a fundamental aspect of cellular homeostasis. Aggregates play a prominent role in many neurodegenerative diseases (NDDs) (Ross and Poirier, 2004). Misfolding and aggregation can disrupt cellular function in two ways, loss of endogenous protein function or gain of a novel toxic function, and there is evidence that both processes occur in disease (Winkler et al., 2008). The aggregation cascade involves multiple intermediate conformations, such as globular and fibrillar structures, up to the late-stage inclusion bodies (IBs) (Knowles et al., 2014).

Huntington’s disease (HD) is an autosomal dominant NDD characterized by neuropsychiatric and motor impairments (Saudou and Humbert, 2016). It is caused by a CAG repeat expansion in exon 1 of the huntingtin (HTT) gene, which leads to an expanded polyglutamine (polyQ) stretch in the N terminus of the huntingtin protein (The Huntington’s Disease Collaborative Research Group, 1993). A hallmark of this disorder is the appearance of cytoplasmic and intranuclear huntingtin aggregates (Davies et al., 1997; DiFiglia et al., 1997), which interfere with several cellular processes, such as proteostasis, transcription, vesicular trafficking, and energy metabolism (Orr and Zoghbi, 2007; Saudou and Humbert, 2016).

Obtaining insights into the composition of huntingtin IBs is of particular importance in understanding HD mechanisms. So far, it has been challenging to characterize the protein composition of IBs, because they remain insoluble even in very high detergent concentrations. Efforts to purify late-stage polyQ aggregates culminated in the identification of only a few tens of proteins, mainly heat shock factors (Mitsui et al., 2002), components of the ubiquitin-proteasome pathway (Doi et al., 2004), and certain transcription factors (Doi et al., 2004; Shimohata et al., 2000). Given the size of the IBs—up to several micrometers (Gutekunst et al., 1999)—it is likely that they contain far more proteins. Moreover, the quantitative composition of IBs and its variation among brain regions is almost entirely unknown.

Mass spectrometry (MS)-based quantitative proteomics is powerful technology for systems-wide analysis of complex cellular processes (Aebesold and Mann, 2016). To date,
proteomic approaches exploring HD pathogenesis have focused either on expression level changes or on soluble huntingtin interactions (Culver et al., 2012; Langfelder et al., 2016; Shirasaki et al., 2012). In contrast, knowledge of the interplay between soluble and insoluble aspects of the proteome is limited (Baldo et al., 2012), particularly when following disease progression over time and in differentially affected brain regions.

Here, we investigate molecular neurodegeneration signatures in a mouse HD model by MS-based quantitative proteomics with spatiotemporal resolution. We describe the soluble and insoluble proteomes of four brain regions, as well as the cerebrospinal fluid proteome at three time points of disease progression. Acid-based hydrolysis of the aggregates allows us to characterize them in depth by quantitative MS. Our data enable comparison of the degree of protein sequestration from the soluble pool to IBs over time, as well as analysis of sequence features of the sequestered proteins. Functional follow-up in a cellular HD model reveals that increasing expression levels of sequestered proteins in many cases restores cellular function and alters the nature of HttEx1 aggregates, suggesting that widespread loss of protein function due to sequestration contributes to HD pathogenesis.

RESULTS

R6/2 transgenic mice express N-terminal exon 1 of polyQ-expanded huntingtin under control of the human huntingtin promoter (Mangiarini et al., 1996). For our analysis of neurodegeneration-associated proteome changes, we selected an early time point at 5 weeks, before disease onset; an intermediate time point at 8 weeks, on the verge of visible motor deficits; and a late time point at 12 weeks, at the end of the lifespan of R6/2 mice (Carter et al., 1999; Davies et al., 1997). We prepared soluble and insoluble extracts from four brain regions (striatum, cortex, hippocampus, and cerebellum), which show differential vulnerability to HD, with striatum being most severely affected and cerebellum remaining relatively spared until advanced disease stages (Vonsattel and DiFiglia, 1998). To relate the proteome changes in the brain to those in a proximal body fluid, cerebrospinal fluid was also analyzed (Figure 1).

Spatiotemporal Brain Proteome Resource of Wild-Type Mice

Recently, we compiled an in-depth cell-type- and region-specific proteomic catalog of the adult mouse brain with a depth of ~13,000 proteins (Sharma et al., 2015). Building on that analysis by using the peptide identifications as a library (Supplemental Experimental Procedures), we now identified 12,498 proteins in the soluble proteome alone using single-run liquid chromatography-tandem mass spectrometry (LC-MS/MS). Of these, we chose 8,455 proteins by stringent filtering for valid values (Figure 2A; Table S1A; Supplemental Experimental Procedures). Correlation analysis indicated a high degree of reproducibility between replicates (R = 0.84 overall; R = 0.90 for biological replicates) (Figure S1A). First, we investigated the soluble proteome of wild-type (WT) controls. Protein abundances spanned seven orders of magnitude, with only 217 proteins amounting to 50% of the total protein mass (Figure 2A). Gene Ontology (GO) biological processes correlated with overall protein abundances in expected ways (Figure 2A; Table S1A; Supplemental Experimental Procedures). Correlation analysis indicated a high degree of reproducibility between replicates (R = 0.84 overall; R = 0.90 for biological replicates) (Figure S1A). First, we investigated the soluble proteome of wild-type (WT) controls. Protein abundances spanned seven orders of magnitude, with only 217 proteins amounting to 50% of the total protein mass (Figure 2A). Gene Ontology (GO) biological processes correlated with overall protein abundances in expected ways (Figure 2A; Tables S1B and S1C). High-abundant proteins tended to have a smaller molecular mass (Figure 2B), as in other tissues (Wiśniewski
et al., 2014), presumably reflecting evolutionary constraints on biosynthesis costs (Warringer and Blomberg, 2006). Our approach uncovered several brain region-specific protein clusters (Lein et al., 2007; Sharma et al., 2015). For instance, proteins involved in dopaminergic signaling were enriched in the striatum, which receives extensive dopaminergic innervation (Figure 2C; Table S1D) (Gerfen and Surmeier, 2011). As expected for the narrow time window of the analysis compared to the lifespan of WT animals, little changes in protein expression were detected at different time points (Figure S1B) (Walther and Mann, 2011). Collectively, our soluble brain region data of WT mice provide a comprehensive resource for the community, with copy numbers and spatiotemporal resolution for almost 8,500 proteins.

Extensive Brain Proteome Remodeling during Disease Progression in R6/2 Mice

Principal-component analysis (PCA) on the spatiotemporally resolved soluble proteomes of R6/2 and WT mice showed clear separation of samples with two major effects (Figure 3A). First, R6/2 mice after disease onset (8 and 12 weeks) were separated from other samples, reflecting extensive protein remodeling during pathogenesis (Figure 3A). Second, cerebellar proteomes were strongly separated from other brain regions. Only a few proteins were responsible for driving separation in the PCA (Figure S1B) (Walther and Mann, 2011). Concordant with the PCA results, the three individual groups exhibited similar numbers of significant protein changes. The combination of age and genotype contributed almost as much as age and genotype separately, indicating that age and genotype jointly drove the extensive proteome remodeling in our dataset. To investigate spatiotemporal changes of the soluble proteome in more detail, we tested for differences in any functional annotations to the background protein distribution and observed that protein annotations changed spatiotemporally and in accordance with known features of HD progression (Figure S2B; Supplemental Experimental Procedures) (Seiger et al., 2012). One example is the increased expression of proteins involved in neurotransmitter secretion at early disease stages in hippocampus and striatum (8 weeks) followed by a decrease at later stages (12 weeks) (Chen et al., 2013). Another finding is the loss of protein expression associated with the calcineurin complex in the cortex (Gratuz et al., 2015). We have compiled a list of annotation changes across brain regions and time (Table S1E) and find many potentially interesting clues for the selective spatial disease progression, such as reduced cortical expression of annotations associated with amino acid import and ion transmembrane transporter activity or reduced striatal expression of annotations associated with mRNA splicing (Figure S2B). Hence, our dataset provides a rich resource of brain region-specific changes in R6/2 mice to formulate testable hypotheses about HD pathogenesis.

Expression profiles of R6/2 and WT animals at 5 weeks were largely identical, although specific protein clusters already exhibited different expression (Figure 3C). Just three weeks later, expression of mutant HttEx1 caused extensive remodeling of the proteome with marked up- and downregulation of thousands of proteins (Figure 3C, colored clusters). The two upregulated clusters were enriched in proteins of the chaperonin containing TCP1 complex (cluster 2, p < 2.0E–4, BH-FDR) and proteasome accessory complex (cluster 3, p < 4.0E–3), whereas the two downregulated clusters were enriched in proteins associated with the ribosome (cluster 1, p < 9.5E–4) and energy transport.
across the mitochondrial electron transfer chain (cluster 4, \( p < 9.0 \times 10^{-4} \)) (Table S1F). This is consistent with the deficient energy metabolism and impairment of the ubiquitin-proteasome system in HD (Acuña et al., 2013; Ortega and Lucas, 2014). Our analysis now provides the underlying protein changes in the R6/2 model on a proteome-wide scale.

A hallmark of HD neuropathology is the selective degeneration of striatal medium spiny neurons (MSNs), which has been linked to transcriptional dysregulation of cyclic AMP (cAMP) and CREB signaling (Vonsattel and DiFiglia, 1998; Wyttenbach et al., 2001). We found Pde10a, a cAMP or cyclic guanosine monophosphate (cGMP)-hydrolyzing enzyme highly enriched in MSNs, to be markedly downregulated in R6/2 striatum (Figure 3D). It is one of the earliest and most significantly downregulated gene products in HD patients and a promising target to restore cyclic nucleotide (cNMP) signaling in affected neurons (Giampa, et al., 2010). Furthermore, Pde10a ligands are used in positron emission tomography (PET) to assess the extent of disease and predict conversion to HD (Russell et al., 2014).

We next asked whether other proteins follow the same profile and thus might serve as markers of HD progression. Following the Euclidean distance of the top 20 expression profiles most similar to Pde10a revealed a set of proteins with interrelated functions, such as cNMP metabolism, neurotransmitter, and specifically dopaminergic signaling (Figure 3D; Figure S2C). The first group included Pde10a-related Pde1b, as well as Darpp-32, which is widely used as a marker for both MSNs and the attenuation of dopaminergic signaling in HD (Jiang et al., 2011). Further candidates were also associated with dopamine metabolism, such as Drd1, a dopamine receptor expressed in MSNs whose striatal loss correlates with cognitive decline (Chen et al., 2013). Adenylate cyclase 5 (Adcy5) is predominantly expressed in the striatum and has been linked to rarer forms of chorea and dystonia (Carapito et al., 2015). A transcriptomics and proteomics study identified 17 of our 21 potential markers among the top 50 hub genes of a striatal module negatively associated with CAG repeat length in a different HD mouse model (Langfelder et al., 2016). Thus, our data define a set of MSN-specific proteins that can be used to assess cNMP-linked dopaminergic imbalances.

**Inflammatory Profile in the Cerebrospinal Fluid Proteome**

The cerebrospinal fluid is an invaluable source in assessing pathological alterations of the CNS (Kroksveen et al., 2011). We next asked whether HD signatures in the soluble brain proteome correlate with changes in the cerebrospinal fluid. A total of 778 cerebrospinal fluid proteins were quantified with our single-shot workflow after filtering out likely contaminants (Table S2A; Supplemental Experimental Procedures). Unsupervised...
hierarchical clustering separated R6/2 and WT samples, indicating changes in protein secretion to the cerebrospinal fluid upon expression of mutant HttEx1 (Figure 4A). Among the most dysregulated proteins we found peptidase regulators increased in the cerebrospinal fluid of R6/2 mice (Figures 4B and 4C; Table S2B). The presence of proteolytic enzymes reflects inflammatory processes, and our data suggest that several members of the cathepsin family dominantly contribute to these processes, because they are major drivers of the PCA separation between R6/2 and WT samples (Ossovskaya and Bunnett, 2004). Finally, we compared correlations of protein expression changes between the brain and the cerebrospinal fluid proteome (Figure 4D). Although expression changes in the brain over time correlated between R6/2 mice (R > 0.5), the correlation between soluble brain proteomes and cerebrospinal fluid was low (R < 0.1) (Figure 4D). This indicates that global expression changes in the soluble proteome are not necessarily reflected in the cerebrospinal fluid proteome.

In-Depth Characterization of the Insoluble Proteome

In Vivo

Huntingtin-containing IBs are a hallmark of HD, but the identity and quantity of their constituent proteins are largely unknown. In agreement with previous studies (Davies et al., 1997; Meade et al., 2002), brains of R6/2 mice already displayed widespread IBs at 5 weeks, with striatal IBs being significantly smaller than cortical ones (Figures S3A–S3C).

Although purification of late-stage aggregates is straightforward due to their extreme insolubility, this biophysical property also makes them resistant to proteolytic digestion, a precondition for MS analysis. Concentrated formic acid can dissolve polyQ aggregates (Hazeki et al., 2000; Kim et al., 2016), and here we demonstrate that it is compatible with tissue-based quantitative proteomics (Figure 1; Supplemental Experimental Procedures). This purification approach, when coupled to state-of-the-art MS, identified several hundred proteins and revealed extensive protein sequestration to the insoluble fraction in the R6/2 brain, including endogenous mouse Htt (Figure S4; Table S3). For quantitative analysis of the aggregate composition, we used the iBAQ (intensity-based absolute quantification) algorithm to estimate protein abundances (Schwanhäusser et al., 2011). Only 10 proteins constituted more than 50% and the top 50 proteins constituted more than 75% of the aggregate mass in 12-week-old R6/2 striatum (Figure 5A). Apart from histones and RNA-binding proteins, these included proteins involved in neuronal plasticity, as well as myelin components.

Figure 4. Altered Cerebrospinal Fluid Proteome in R6/2 Mice

(A) Hierarchical clustering of cerebrospinal fluid protein expression reveals good separation of R6/2 and WT mice; row bars indicate filters for known GO annotations related to secretion (see Supplemental Experimental Procedures for details).

(B) PCA projections of all cerebrospinal fluid samples show good separation between WT and R6/2 mice.

(C) Corresponding PCA loadings of (B) reveal strong inflammatory response in R6/2 animals. The color bar displays the 1D annotation score with the two most enriched annotations for R6/2 mice; the annotation score indicates the center of the protein distribution of each significant annotation category relative to the overall distribution of values. p value, BH-FDR corrected.

(D) Correlation of protein expression changes between different disease stages is high within the soluble proteome (purple inlay) but low for the comparison of the soluble proteome with the cerebrospinal fluid (green inlay) of R6/2 mice. See also Table S2.
Co-immunostainings for the myelin protein Plp1, aggregated HttEx1, and a neuronal marker revealed co-localization of Plp1 with aggregated HttEx1 in many neuronal IBs, confirming that myelin proteins are true components of neuronal inclusions (Figures S3D and S3E).

Most sequestered proteins contributed little to overall aggregate mass. There was a general tendency for abundant cellular proteins to also be among the more abundant aggregate proteins (median R = 0.40 across all conditions) (Figure S5). GO analysis revealed many proteins associated with native huntingtin function and the proteostasis network to be enriched in the R6/2 insoluble fraction, including known interactors and aggregation modifiers of polyQ-expanded huntingtin. Several members of the TCP-1 ring complex (TRiC) and the Hsp40, Hsp70, and Hsp90 families made up ~1.5% of the insoluble fraction mass. All five members of the collapsin response mediator protein (Crmp) family, also known as dihydropyrimidinase-related proteins (Dpysl), were also enriched (Figure 5B). A study identified Crmp1 as a suppressor of huntingtin toxicity (Stroedicke et al., 2015). Apart from confirming these known aggregate constituents, our data provide a large number of new proteins with potential links to HD (Table S3A).

We hypothesized that the extensive soluble proteome remodeling in HD should be reflected in the insoluble proteome.

Figure 5. In-Depth Characterization of PolyQ Aggregates
(A) Distribution of iBAQ values for 12 week R6/2 striata. Pie chart distribution of annotations for the top 50 proteins.
(B and C) Proteins with dysregulated soluble expression (color coding from Figure 3C) are enriched in R6/2 insoluble fractions, superimposing enrichment for all 12 week brain regions together. The most enriched insoluble proteins and endogenous Htt is indicated.
(D–G) Significant enrichment of proteins with longer polyQ (D) and LCRs (E), more CCDs (F), and higher molecular weight (MW) (G) in 12 week R6/2 over WT striata. Mann-Whitney U test.

See also Figures S3–S5 and Table S3.
Mapping our soluble proteome data to the insoluble fraction revealed a large number of proteins that were downregulated in the soluble proteome and enriched in the aggregates (blue and cyan in Figures 5B and 5C; Figure S4), suggesting widespread loss of protein function by sequestration. Conversely, we also found many proteins that were upregulated in both the soluble and the aggregate proteomes (orange and red in Figures 5B and 5C; Figure S4; Table S3B). Among the latter, the TRiC chaperonin and the GO term “proteasome accessory complex” were highly overrepresented (p < 4.0E−5 and p < 3.2E−3 BH-FDR, respectively) (Table S3E), confirming specific upregulation of chaperones that interact with aggregates but become entangled with them. Potential loss-of-function candidates (reduced in the soluble proteome and increased in the insoluble proteome) often had a significantly higher abundance compared to proteins upregulated in the soluble proteome (Figure S6A), consistent with a sequestration mechanism that depletes the cellular pool of these proteins, thereby impeding cellular function.

Next, we asked whether sequestration into the aggregates correlated with certain biophysical features. We focused on polyQ length, low-complexity regions (LCRs), and coiled-coil domains (CCDs), motifs known to modulate protein aggregation (Fiumara et al., 2010; Kato et al., 2012; Li and Li, 2004; Schaefer et al., 2012). Insoluble fractions from R6/2 striata contained significantly more aggregation-prone proteins than WT controls, with a robust effect size for CCDs and small effect sizes for polyQ and LCRs (Figures 5D–5F; Figures S6B–S6D). In addition, R6/2 samples were significantly enriched for proteins with higher molecular weight (Figure 5G; Figure S6E). Larger proteins tend to be less thermodynamically stable, which may explain why their folding is compromised under conditions of conformational stress (Sharma et al., 2012).

To quantify the degree of sequestration, we estimated absolute protein abundance using the iBAQ algorithm and compared the amount of each protein in the soluble and insoluble proteome (Figure 6A; Figure S6F). Extracellular matrix (ECM) proteins had the least soluble proportion, reflecting their insolubility and demonstrating efficient enrichment of insoluble proteins in our protocol (Figure 6A, marked in red). Endogenous Htt was also recruited into the aggregates (Figure 6A, marked in pink). In concordance with the increase in aggregate size, we observed increased sequestration of protein mass over time, hence progressively depleting the pool of functional proteins. In 8-week-old animals, a full 80% of proteins were at least in a 1:1 ratio in the cortical IBs compared to the soluble pool.

To determine whether the IB composition varies over time and across brain regions, we tested for differences in any functional annotations to the background protein distribution (Figure 6B; Table S3F; Supplemental Experimental Procedures) and found spatiotemporal changes of protein annotations, reflecting many known features of HD pathogenesis. Transcriptional and epigenetic dysregulation is an early event in HD, and we found that IBs at 5 weeks were already significantly enriched in DNA- and chromatin-binding elements (FDR < 0.05). Conversely, late-stage IBs at 8 to 12 weeks were enriched in proteins associated with glutamate receptor signaling and synaptic transmission, correlating with the onset of motor phenotypes (FDR < 0.05, respectively). Similarly, we observed spatial changes such as enrichment of SMAD-binding proteins in cortical IBs (FDR < 0.05), or depletion of proteins involved in Wnt signaling pathways in striatal IBs (FDR < 0.05), linked to synaptic degeneration (Galli et al., 2014). Our proteomic analysis therefore demonstrates that protein sequestration is both brain region and time specific and may link dynamic changes in the IBs to the phenotype.
Overexpressing Loss-of-Function Candidates Rescues PolyQ Length-Dependent Toxicity and Alters IB Size

We identified several hundred proteins that were depleted from the soluble pool and increased in insoluble aggregates. To test whether this represents widespread loss of protein function, we selected a group of 18 candidates for functional follow-up based on several criteria, including their profiles in the proteomics experiments and association with neuronal functions but no previous link to HD (Figure S7A; Table S4). For several candidates, reduction in protein levels in the soluble fraction was confirmed by western blot (Figure 7A; Figure S7B). Among the selected candidates was the amyotrophic lateral sclerosis (ALS)-associated protein hnRNPA2B1. ALS-linked hnRNPA2B1 mutations lead to an accumulation of insoluble protein in the nucleus (Martinez et al., 2016). Immunostaining in R6/2 striatum demonstrated a nuclear accumulation of hnRNPA2B1 in most MSNs, as well as its co-localization with HttEx1 IBs (Figures S7C–S7E).

For the follow-up of the selected proteins, we used inducible neuron-like PC12 cells with stably integrated, GFP-tagged HttEx1 with either 23 or 74 glutamines (HD-Q23 and HD-Q74 cells). Induction of HttEx1 led to polyQ length-dependent cell death (Figure S7F) (Wyttenbach et al., 2001). Quantitative proteomics revealed that all candidates were expressed in this cell line with medium to high abundance (Figure S7G), and induction of HttEx1 did not substantially alter their expression (Figure S7H). Furthermore, in most cases, transfection of a candidate increased its protein levels compared to non-transfected controls (Figure S7I). Overexpression of individual candidates significantly improved the viability of HD-Q74 cells in 12 of 18
cases (BH-FDR < 0.05), with a mean survival increase of 40% (Figure 7B, lower panel). Three proteins (Ap2a2, Mapt, and Nfasc) that increased the viability of control HD-Q23 cells (Figure 7B, upper panel) and two proteins that repeatedly exhibited low transfection rates (Dctn1 and Sbf1) were excluded from further analyses. To distinguish between a general effect on cell survival and a specific involvement in mutant HttEx1 toxicity, we tested whether the candidates rescued the viability of non-induced HD-Q74 cells under starvation and observed a rescue with only one protein, Cnp (Figure 7C). Thus, the effects of the remaining candidates were specific for mutant HttEx1 toxicity.

We next asked whether the HttEx1 toxicity-modifying proteins also had an effect on aggregation. Overexpression of the candidates did not change the number of GFP+ foci per cell (Figure S7J) but did reduce their size in 10 of 11 cases (Figure 7D). Moreover, the percentage of small foci (<1 μm in diameter) increased significantly in 8 of the 11 cases, whereas the percentage of large foci (>2.5 μm) was significantly reduced in all cases (Figures 7E and 7F). As an orthogonal approach, we employed membrane filter trap assay, which also did not reveal a significant change in the aggregation load (Figure S7K). Finally, the soluble proportion of most cellular proteins increased significantly after overexpression of the candidates in HD-Q74 cells, as well as, in most cases, the soluble proportion of the candidate (Figure 7G). Thus, overexpression of the proteins sequestered to HttEx1 IBs ameliorates mutant HttEx1 toxicity and decreases aggregate size, but not load.

DISCUSSION

We applied quantitative proteomics to investigate molecular neurodegeneration signatures in an in vivo model of HD. Employing a recently published proteomic resource of the mouse brain (Sharma et al., 2015) and a single-run LC-MS/MS workflow, we characterized the brain proteome of the R6/2 mice and WT controls to a depth of more than 12,000 proteins. Spatiotemporal characterization of the soluble proteome during disease progression uncovered extensive alterations in brain regions vulnerable to HD. The data provide a resource to the community, which is available via a user-friendly database (http://maxqb.biochem.mpg.de/mxdb/). The value of such data is illustrated by the identification of several MSN-specific proteins whose signaling is compromised in HD and that are already used for PET imaging. These clinically relevant proteins emerge from our analysis and thus validate the R6/2 model, which is one of the best characterized and most widely used HD mouse models. Comparative studies demonstrated its extensive similarity with full-length Huntingtin models at pathological, transcriptional, and electrophysiological levels (Cumberworth et al., 2013; Kato et al., 2012). CCDs are prominently involved in protein-protein interactions and can interact with polyQ proteins, promoting their aggregation (Fiumara et al., 2010). RNA-binding proteins, which figure prominently in the aggregates, are particularly rich in LCRs. CCDs and LCRs are molecular recognition motifs, regulating oligomerization of higher-order structures among both RNAs and proteins (Cumberworth et al., 2013; Kato et al., 2012). LCRs are molecular recognition motifs, regulating oligomerization of higher-order structures among both RNAs and proteins (Cumberworth et al., 2013; Kato et al., 2012). The formation of such reversible structures allows increase of local concentrations of relevant interactors. However, because LCRs also promote protein binding promiscuity and aggregation propensity, this comes at the cost of undesirable, non-productive interactions, leading to several diseases (Cumberworth et al., 2013). We have shown in a cell model of HD that LCR domains co-aggregate with mutant HTT and other LCR-containing proteins, exacerbiating aggregate formation (Kim et al., 2016). Our current results suggest that LCRs are also responsible for promoting co-aggregation in vivo. Several elements of the protein folding machinery also contain LCRs (Hageman et al., 2010). It is tempting to speculate that the presence of these sequences in both aggregation-prone and aggregation-counteracting proteins represents a balance that has co-evolved over time.

Our insoluble fraction showed little overlap with recently published protein aggregation sets. Of all insoluble proteins in our dataset, less than 5% overlapped with interactors of artificial...
amyloid-like beta sheet proteins (Olszcha et al., 2011), and less than 1% were shared with either RNA-binding proteins containing prion-like domains (Li et al., 2013) or stress granules (Jain et al., 2016). The little overlap between these aggregation-associated proteins indicates that proteins from our insoluble dataset specifically co-aggregate with mutant HttEx1.

The integrative aspect of jointly analyzing the soluble and insoluble proteome, combined with the spatiotemporal resolution, allowed us to elucidate molecular neurodegeneration signatures that have not yet been linked to HD. Hundreds of proteins were downregulated in the soluble proteome and upregulated in the insoluble proteome. This prevalent sequestration of proteins into the aggregates suggests widespread loss of protein function in HD. Although it has been reported that IBs are protective by sequestering toxic soluble oligomers (Arrasate et al., 2004; Haass and Selkoe, 2007), the substantial depletion of soluble proteins observed here suggests a major impairment of protein homeostasis in the cell.

In accordance with this hypothesis, our rescue experiments demonstrated that overexpression of loss-of-function candidates ameliorated HttEx1 toxicity in a cellular HD model. The molecular mechanism underlying the observed increase in cell viability requires further analysis. It is possible that each of the selected candidates is necessary for cell survival and overexpression simply re-supplies the cell with an essential factor. In addition, a protein with a strong affinity for HttEx1 aggregates may occupy much of the aggregates’ interaction surface and thereby reduce sequestration of endogenous proteins, as suggested by the increase of the soluble proportion of most proteins upon transfection of the candidates into HD-Q74 cells. The latter scenario may also explain why many of our candidates reduced IB size. Our data therefore suggest that overexpression of candidate proteins could increase cell viability by interfering with multistage aggregate formation. Soluble interactors of both native and mutant NDD proteins have been identified as disease modifiers (Hosp et al., 2015; Kaltenbach et al., 2007). Most proteins in the insoluble aggregates, however, are most likely not associated with the native function of the disease protein. Given the high proportion of toxicity mediators in our validation set, it is likely that the total set of potential loss-of-function proteins contains further candidates that would be interesting to study in the context of HD.

In summary, we provide a rich resource comprising (1) changes in the proteome upon disease progression, (2) protein copy numbers, (3) degree of protein sequestration to the IBs, and (4) correlation of proteomes across different brain regions, time, and aspects of the brain and cerebrospinal fluid proteome. Altogether, these data paint a quantitative picture of the dynamic aggregate proteome in relation to the brain proteome. This highlights the power of integrative approaches to elucidate molecular mechanisms of HD, helping to bridge the gap between identification of disease-associated pathways and their corresponding phenotypes.

**EXPERIMENTAL PROCEDURES**

Further details and an outline of resources used in this work can be found in Supplemental Experimental Procedures.
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and four tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2017.10.097.

AUTHOR CONTRIBUTIONS

F.H. performed wet lab and MS experiments and analyzed the data. S.G.-A. performed wet lab and imaging experiments and analyzed the data. I.D. performed cerebrospinal fluid extraction and mouse brain dissection. M.H.S. and J.C. developed bioinformatic tools. M.S.H., F.M., and F.-U.H. provided reagents and advised research. F.H., S.G.-A., I.D., R.K., and M.M. interpreted the data. F.H. and M.M. conceived the project and wrote the manuscript. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

We thank our colleagues at the Max Planck Institutes of Biochemistry and Neurobiology, especially G.H.H. Borner and D. Hornburg for help and fruitful discussions; S. Kroiss, K. Mayr, I. Paron, G. Sowa, and B. Splettstösser for excellent technical assistance; and K. Schulz-Trieglaff and the MPIB animal facility for help with the R6/2 breeding and caretaking. We acknowledge the gift of the PC12 HD cell lines from D. Rubinsztein (University of Cambridge, UK). The research leading to these results has received funding from the Munich Cluster for Systems Neurology “SyNergy,” the European Research Council Synergy grant “ToPAG—Toxic protein aggregation in neurodegeneration” (ERC-2012-SyG_318987-ToPAG), and the Max Planck Society for the Advancement of Science.

Received: June 23, 2017
Revised: September 13, 2017
Accepted: October 24, 2017
Published: November 21, 2017

REFERENCES


