Published in final edited form as:

Structure. 2020 September 01; 28(9): 1035–1050.e8. doi:10.1016/j.str.2020.06.008.

The Polyglutamine Expansion at the N-Terminal of Huntingtin Protein Modulates the Dynamic Configuration and Phosphorylation of the C-Terminal HEAT Domain

Taeyang Jung[1,2,3,11], **Baehyun Shin**[4,5,11], **Giorgio Tamo**[6,11], **Hyeongju Kim**[1,11], **Ravi Vijayvargia**[4,5,12], **Alexander Leitner**[7], **Maria J. Marcaida**[6], **Juan Astorga-Wells**[8,9], **Roy Jung**[4,5], **Ruedi Aebersold**[7,10], **Matteo Dal Peraro**[6,*], **Hans Hebert**[2,3,*], **Ihn Sik Seong**[4,5,*], **Ji-Joon Song**[1,13,*]

[1]Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), KI for the BioCentury, Daejeon 34141, Korea

[2]School of Engineering Sciences in Chemistry, Biotechnology and Health, Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, 141 52 Huddinge, Sweden

[3]Department of Biosciences and Nutrition, Karolinska Institutet, 141 83 Huddinge, Sweden

[4]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

[5]Department of Neurology, Harvard Medical School, Boston, MA 02114, USA

[6]Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

[7]Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, 8093 Zürich, Switzerland

[8]Department of Medical Biochemistry & Biophysics, Karolinska Institutet, 171 65 Solna, Sweden

[9]HDxperts AB, 183 48 Täby, Sweden

[10]Faculty of Science, University of Zürich, Zürich, Switzerland

[11]These authors contributed equally

[12]Present address: Department of Biochemistry, The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat 390 002, India

[13]Lead Contact

[*]Correspondence: matteo.dalperaro@epfl.ch (M.D.P.), hansh@kth.se (H.H.), iseong@mgh.harvard.edu (I.S.S.), songj@kaist.ac.kr (J.-J.S.).

AUTHOR CONTRIBUTIONS

T.J., B.S., G.T., M.D.P., H.H., I.S.S., and J.-J.S. conceived the idea. T.J. and H.K. performed cryo-EM. G.T. performed integrative modeling. B.S., R.V., and R.J. performed for PTM analysis. A.L. and R.A. performed XL-MS. M.J.M. performed SAXS experiments. J.A.-W. performed HDX-MS. All authors reviewed the data, and T.J., B.S., I.S.S., and J.-J.S. mainly wrote the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2020.06.008.

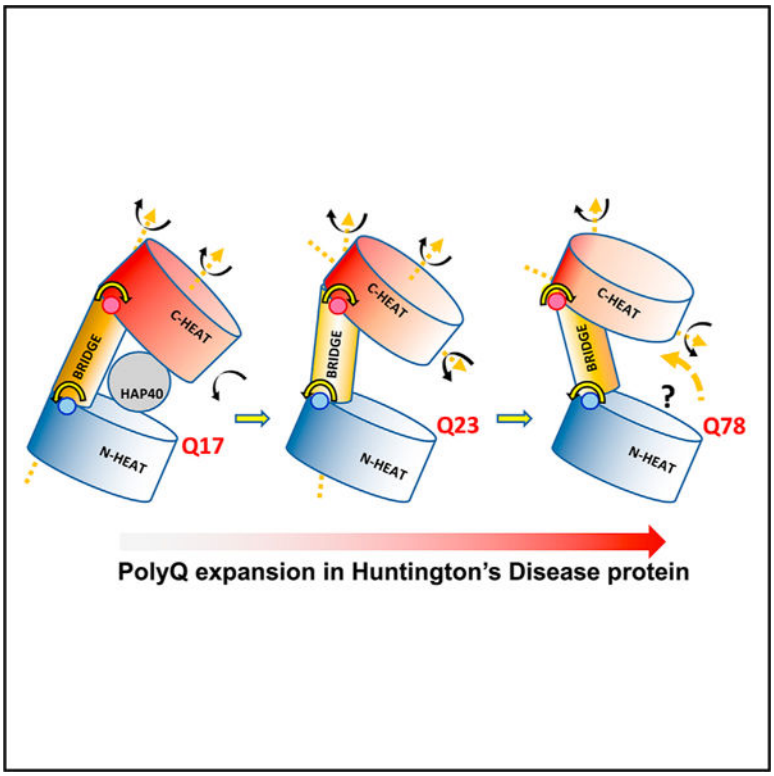
DECLARATION OF INTERESTS

The authors declare no competing interests.

SUMMARY

The polyQ expansion in huntingtin protein (HTT) is the prime cause of Huntington's disease (HD). The recent cryoelectron microscopy (cryo-EM) structure of HTT-HAP40 complex provided the structural information on its HEAT-repeat domains. Here, we present analyses of the impact of polyQ length on the structure and function of HTT via an integrative structural and biochemical approach. The cryo-EM analysis of normal (Q23) and disease (Q78) type HTTs shows that the structures of apo HTTs significantly differ from the structure of HTT in a HAP40 complex and that the polyQ expansion induces global structural changes in the relative movements among the HTT domains. In addition, we show that the polyQ expansion alters the phosphorylation pattern across HTT and that Ser2116 phosphorylation in turn affects the global structure and function of HTT. These results provide a molecular basis for the effect of the polyQ segment on HTT structure and activity, which may be important for HTT pathology.

Graphical Abstract



In Brief

Jung et al. demonstrate that the polyQ expansion in Huntington's disease protein results in global conformational change and phosphorylation pattern using integrative structural approaches.

INTRODUCTION

Huntington's disease (HD) is a progressive autosomal-dominant neurodegenerative disease that is diagnosed by an uncontrolled motor movement known as chorea, cognitive disorder

and depression (Margolis and Ross, 2003; Walker, 2007). HD is caused by a CAG triplet repeat expansion in the huntingtin (*HTT*) gene, leading to an abnormal polyglutamine (polyQ) expansion in the huntingtin (HTT) protein (Myers, 2004). However, the mechanism by which the expansion of only 30–40 glutamine repeats alters the function of HTT, invoking HD, is still not understood, not to mention the normal function of the HTT protein. Multiple studies have revealed that the length of the expanded repeat is proportionally related to the severity of disease (Lee et al., 2012; McNeil et al., 1997; Rubinsztein et al., 1996; Walker, 2007). These data suggest that the polyQ expanded mutant HTT causes disease with a gain of function (Gusella and MacDonald, 2000; Nucifora et al., 2001). HTT is expressed in most tissues with various expression levels and is involved in a number of cellular functions (Sapp et al., 1997; Walker, 2007). Furthermore, the HTT gene is essential for embryonic development (Jeong et al., 2006). However, there is no clear mechanistic understanding of the function of HTT and the toxic activity conferred on HTT caused by polyQ tract expansion.

The HTT protein contains the HEAT (Huntingtin, Elongation factor 3, protein phosphatase 2A, Target of rapamycin 1) repeat motif (Andrade and Bork, 1995). The structural flexibility of HTT has been suggested in previous structural and biophysical studies (Grinthal et al., 2010; Halder et al., 2015; Vijayvargia et al., 2016; Yoshimura and Hirano, 2016). HEAT motif proteins function as scaffold proteins interacting with other binding partners. Consistent with this, HTT has been shown to interact with various partners and participate in multiple cellular functions (Harjes and Wanker, 2003). The flexibility of HTT may be an important structural feature that may be critical to its function, and it is plausible that polyQ tract expansion at the N terminus could affect the overall structure of HTT. Indeed, we previously performed single-particle electron microscopy (EM) analysis and analyzed intra-molecular interactions of normal and mutant HTT with a panel of highly purified human recombinant HTTs, suggesting polyQ length-dependent overall structural changes (Vijayvargia et al., 2016). Moreover, these structural differences of purified HTT caused by the length of the polyQ tract is suggested to result in differential effects on the enhancement of Polycomb Repressive Complex 2 (PRC2) histone methyltransferase activity in a cell-free assay (Seong et al., 2010). However, further investigation is required to understand the mechanism by which the polyQ length expansion alters the structure of HTT, which may be implicated in HD pathology.

Recently, a high-resolution cryo-EM structure of HTT was determined in complex with HTT-associated protein 40 (HAP40) (Guo et al., 2018). This structure delineated the domain structures of HTT, showing that HTT is composed of the N-HEAT, Bridge, and C-HEAT domains. The Bridge domain is also composed of HEAT repeats connecting the N- and C-HEAT solenoid structures. This structure strongly supports that HTT is a functional multivalent scaffold hub (Biddy et al., 2018).

HAP40 is bound between the N-HEAT and the C-HEAT domains and makes the HTT structure compact, suggesting that HAP40 binding to HTT induces structural changes and that the structure of apo HTT might be substantially different from that of the HAP40 bound form. Although the function of HAP40 in the context of HTT function and HD pathology is not clear, this structure set a platform to further investigate the structural changes of HTT,

which might be caused by polyQ expansion. Although the atomic resolution structure of HTT in the HAP40 complex is available, it is critical to know the structure of apo HTT to understand how the flexible structure of HTT can be altered by polyQ expansion, which might be implicated in the HD pathology.

Here, we investigated the structural differences among HTTs in different states. We chose HTT with 23 glutamines (Q23-HTT) as a wild-type HTT and HTT with 78 glutamines (Q78-HTT) as an HD-causing mutant. We applied an integrative structural and biochemical approach, including cryo-EM, crosslinking mass spectrometry (XL-MS), molecular modeling, small-angle X-ray scattering (SAXS), hydrogen-deuterium exchange mass spectrometry (HDX-MS), phosphorylation analysis, and others, revealing that HTT has a dynamic but modular character and undergoes large conformational changes. Furthermore, we provide a proof of concept that polyQ expansion can induce HTT gain of function via interplay between its phosphorylation status and the conformational rearrangement of the domains of HTT.

RESULTS

Cryo-EM Analysis of Full-Length Q23-HTT in the Apo State

HTT is a modular protein containing HEAT repeats, and the expansion of CAG repeats in the *HTT* gene causing polyglutamine (polyQ) tract expansion in the HTT protein is a direct cause of HD. However, it has neither been shown that polyQ expansion alters the global structure of HTT nor that its conformation is affected by binding partners.

The recent atomic resolution structure (PDB: 6EZ8) of HTT (HTT_{HAP40}) bound to HAP40 showed that HTT is composed of the three domains N-HEAT, Bridge, and C-HEAT, with the Bridge domain connecting the two HEAT domains (Biddy et al., 2018), while HAP40 binds between the N- and C-HEAT domains and seems to compact HTT.

To investigate the structure of HTT in the apo form and its structural change, if any, upon the polyQ expansion, we first determined cryo-EM structures of Q23-HTT. As HTT tends to multimerize over time and we were not able to obtain discrete particles of HTT proteins, we isolated stable monomeric Q23-HTT by 10%–30% sucrose density gradient fractionation in the presence of disuccinimidyl suberate (DSS) crosslinkers (Figure S1). We collected 2,331 micrographs using a microscope with a Volta phase plate (VPP) (Figure 1A and Table S1). Finally 20,825 particles from the micrographs were used for three-dimensional (3D) reconstruction using Relion2.1 to generate a 9.6-Å resolution map estimated by gold-standard Fourier shell correlation (FSC) (Figures 1B, 1C, and S1). However, the secondary structure elements could not be resolved in the cryo-EM map, suggesting that the resolution from FSC is likely overestimated. This might result from relatively weak high-resolution information or/and soft edge masking during the final refinement.

To further validate the cryo-EM structure, we independently processed the identical particle set using cisTEM with an independent initial model generated from cisTEM (Grant et al., 2018). The cross-correlation coefficient (CCC) between the cryo-EM maps generated from

Relion2.1 and cisTEM was 0.94, indicating that our cryo-EM map of Q23-HTT is likely to be free from processing artifacts (Figure S3).

The refined Q23-HTT map shows that the overall dimensions of Q23-HTT are 145 × 90 × 100 Å, consistent with the previous structure obtained by negative-stain single-particle EM (Vijayvargia et al., 2016) (Figure 1C). The structure shows that Q23-HTT could accommodate the three domains N-HEAT, Bridge, and C-HEAT as defined from the high-resolution structure of HTT_{HAP40}. Overall, the global architecture of Q23-HTT is similar to that of HTT_{HAP40}.

As the cryo-EM map at the current resolution did not enable us to independently locate the N terminus of Q23-HTT where the polyQ expansion occurs, we aimed to map the N terminus by determining the cryo-EM structure of N-terminal GFP-tagged HTT (GFP-HTT). GFP is a compact protein and has been utilized for mapping the locations of domains in large complexes in low-resolution cryo-EM maps (Ciferri et al., 2015). We generated GFP-HTT and determined its cryo-EM structure. The final refined 3D map of GFP-HTT was reconstructed from 17,689 particles from 2,226 micrographs collected with an overall resolution of 15.3 Å (Table S1 and Figure S2). Superimposition of GFP-HTT on Q23-HTT (CCC = 0.88) revealed an apparent extra density (Figure S2D), clearly indicating the location of the N terminus of Q23-HTT.

To gain further insight into the structure of Q23-HTT in the apo state, we generated and fitted atomic models to our cryo-EM map of Q23-HTT utilizing the structure of HTT_{HAP40} in complex with HAP40 via an integrative modeling strategy. The HTT_{HAP40} model was split into three rigid domains, N-HEAT, Bridge, and C-HEAT, based on the fact that these three domains are weakly connected by unresolved flexible regions (residues 1,713–1,729 separating N-HEAT and Bridge domains, and residues 2,062–2,092 separating Bridge and C-HEAT domains) in HTT_{HAP40} structure. Similarly XL-MS data on Q23-HTT showed a cluster of short intra-domain contacts especially within N-HEAT and C-HEAT domains (Figure S8), also suggesting independent folding of the three previously defined domains. Therefore, these three domains were considered as the main independent components of HTT and were rigidly fitted into the Q23-HTT density map using a hierarchical multi-body constrained molecular docking strategy (Tamo et al., 2017). First, the best fitting N-HEAT domain from the rigid fitting was selected, which was eventually fully consistent with the location of GFP in GFP-HTT. The Bridge domain and C-HEAT domain were subsequently placed on the map via multi-body fitting. The final Q23-HTT atomic model generated in this first stage contains three distinctive domains, representing HEAT-repeat motifs with a reliable fitting score to the experimental cryo-EM density (CCC value of 0.71) (Figures 2 and S3). Importantly, this final model is the best-ranking solution in terms of CCC of an ensemble of solutions, which uniquely show an arrangement of the C-HEAT domain consistent with the HTT_{HAP40} structure (Figures S3A and S3B).

Conformational Change of the C-HEAT Domain in HTT

Having a Q23-HTT model based on our cryo-EM map, we compared Q23-HTT and HTT_{HAP40} to examine their structural differences (Figure 3A). The overall dimensions significantly differ between the two structures. Q23-HTT is more extended along its longest

axis (the z axis defined in Figure 3B) while it is rather compact in the presence of HAP40 (Figure 3A). To compare the two structures in detail, we aligned Q23-HTT and HTT_{HAP40} by superimposing the N-HEAT domains (Figure 3B). This practice revealed a significant difference between Q23-HTT and HTT_{HAP40} in the relative position of the C-HEAT domain to the N-HEAT domain (Figures 3B and 3C). To evaluate the relative movement among the three domains in greater detail, we first measured the movement of the Bridge domains by anchoring the N-HEAT domains (Figure 3B). With the beginning of the Bridge domain serving as a pivot joint, the Bridge domain of Q23-HTT is rotated about 20° along the y axis (perpendicular to the plane as viewed in Figure 3B) and 60° along the z axis, compared to the Bridge domain of HTT_{HAP40} (Figure 3B, left panel). Next, we measured the movement of the C-HEAT domain by superimposing the Bridge domains (Figure 3B, right panel). The end of the Bridge domain serves as another pivot point, and the C-HEAT domain of Q23-HTT is rotated about 50° and rolled over about 35°, compared with the C-HEAT domain of HTT_{HAP40} (Figure 3B). When these two movements are combined, the C-HEAT domain is rotated 20° at the y axis, 110° at the z axis, and 35° rolled over (Figure 3C). There are two pivot points in HTT at the beginning and end of the Bridge domain as if HTT is a modular structure. With these movements, the C-HEAT domain is consequently moved away from the N-HEAT in Q23-HTT compared with the HTT_{HAP40}. This structural comparison indicates that the positional change of the C-HEAT domain relative to the N-HEAT domain may have an impact on the function of HTT by modulating its interactions with HAP40. In short, HTT exhibits a large C-HEAT domain swing-out movement by utilizing two pivot points at the beginning and end of the Bridge domain, when the N-HEAT domains are aligned. These data show that HTT has a flexible and modular character, and suggest that the conformational change among the domains may play a critical role on the function of HTT such as interacting with other proteins.

Structural Comparison between Q23-HTT and Q78-HTT

The polyQ expansion at the N terminus of HTT leads to HD, and accumulating data suggest that the polyQ expansion alters the global structure of HTT (Cui et al., 2014; Vijayvargia et al., 2016). We previously observed the structural differences between Q23-HTT and Q78-HTT using XL-MS and negative-stain EM (Vijayvargia et al., 2016). However, the precise nature of the structural changes caused by the polyQ expansion remains unknown. To gain detailed insights into the structural differences between Q23- and Q78-HTTs, we aimed to determine the cryo-EM structure of Q78-HTT. Q78-HTT protein was prepared in the same way as Q23-HTT. A total of 1,552 micrographs were collected using a Titan Krios 300 KeV microscope with a Gatan Summit K2 detector using VPP (Figure 4A). A final set of 21,983 particles from the total of 45,038 particles picked were selected for subsequent two-dimensional (2D) class-averaging and 3D reconstruction, generating a cryo-EM map of Q78-HTT at a resolution of about 20 Å (Figures 4A, 4B, and S4A). The overall dimensions of Q78-HTT are similar to those of Q23-HTT, featuring only a slight expansion along the z axis. Despite the relatively low resolution, we were able to clearly locate the position of the N-HEAT and C-HEAT domains. To compare the structure of Q78-HTT with that of Q23-HTT, we fitted the atomic model of each N- and C-HEAT domains onto the Q78-HTT cryo-EM map. While the N- and C-HEAT domains could be positioned within the map, we were not able to place the Bridge domain by rigid body fitting without

deforming its structure, suggesting that there may be some substantial structural changes in the Bridge domain. Therefore, we fitted only the N- and C-HEAT domains into our cryo-EM map, producing an ensemble of solutions all having a similar general arrangement of the domains and spanning a limited interval of CCC values (Figures S3A and S3C). The atomic model optimally fitting the cryo-EM density (CCC value of 0.78) was therefore selected as a representative structure for Q78-HTT (Figure 4B). We then compared the relative orientations of N- and C-HEAT domains of Q23- and Q78-HTTs. When two maps of Q23- and Q78-HTT were superimposed, there was a clear difference in the position of the C-HEAT domains relative to the N-HEAT domains. Specifically, the C-HEAT domains are rotated about 55° along the y axis and rotated about 30° along the z axis, with 35° rolling with respect to each other (Figure 4C). This was confirmed also from a cross-comparison of the whole ensemble of solutions for Q23- and Q78-HTT, where the best-ranking models were already well structurally distinct with a Cα-root-mean-square-deviation as large as 28 Å (Figure S3D). The Bridge domain seems to serve as pivot joints as observed between Q23-HTT and HTT_{HAT40}. These observations suggest that the polyQ expansion at the N terminus of HTT may induce a change in orientation of the C-HEAT domain relative to the position of the N-HEAT domain.

To further investigate the overall structural conformation and possible differences between Q23- and Q78-HTT in solution, we performed SAXS analysis using Q23- and Q78-HTT, which were crosslinked as for cryo-EM (Figure S4B–S4D and Table S2). The maximum dimension of the particles estimated from the P(r) distribution functions (D_{max} = 210 Å for Q23-HTT and 217 Å for Q78-HTT) indicate that the proteins behave as monomers in solution and that Q78-HTT is slightly more extended than Q23-HTT, which is consistent with the models obtained through cryo-EM map fitting. The variability in the extension of the polyQ region between the two constructs is reflected in the differences in the R_g and D_{max} values observed (Table S2), which is also consistent with our cryo-EM analysis on Q23- and Q78-HTT in that Q78-HTT adopts a slightly more extended conformation compared with Q23-HTT. In short, the SAXS data consistently represented a globally similar ‘‘compact’’ conformation of Q23- and Q78-HTT, as well as a slightly more extended dimension of Q78-HTT at the same time. Our comparative structural analysis between Q23- and Q78-HTT suggests that there exists a unique ensemble of similar conformations that can fit both cryo-EM densities, and the best-ranking models for Q23- and Q78-HTT show a distinct architecture with significant conformational change in the relative orientations of the three domains (Figure S3). To further assess the conformational dynamics of Q23- and Q78-HTT, we performed HDX-MS. Q23- and Q78-HTT were incubated for 10, 45, 90, and 180 min for deuteration. The samples were then subjected to protease digestion followed by mass spectrometry. We only analyzed peptides found in both Q23- and Q78-HTT, and the peptide coverage for Q23- and Q78-HTT was 42.7% in both cases (Figures S5A and S5B). The deuterium exchange rate was plotted in a butterfly plot (Figure S5C). Despite the conformational change between Q23- and Q78-HTT observed in the cryo-EM analysis, there was no significant difference in deuterium exchange rate between Q23- and Q78-HTT, which is consistent with our cryo-EM structural analysis showing that HTT is a dynamic modular structure composed of three domains, and suggests that there is no substantial structural difference within the modules. We mapped the locations of the

peptide obtained from HDX-MS on the primary sequences and the structures of Q23- and Q78-HTT (Figures S5C and S5D). Peptides showing high deuterium exchange rate are clustered at several regions. The large region around amino acids (aa) 400–600 located within the N-HEAT domain shows the highest exchange rate and is disordered in cryo-EM structures. Interestingly, many PTMs are found in this region, suggesting that the highly solvent-accessible region may serve as a functional site modulated by PTMs (Rangone et al., 2004; Ratovitski et al., 2017; Schilling et al., 2006; Warby et al., 2005). Another region showing high exchange rate is around aa 2,000–2,200 and is located at the interface between the Bridge and C-HEAT domains, which is consistent with our observation that the Bridge domain serves as pivot joints making HTT conformationally flexible. Overall these data show that HTT is a dynamic modular structure and that the polyQ expansion at the N terminus of HTT modulates the relative orientation of the C-HEAT domain, which may be implicated in mutant HTT functions.

The Conformational Flexibility of HTT

The orientation of the C-HEAT domains in Q23- and Q78-HTT relative to the N-HEAT domain significantly differs, suggesting that polyQ expansion may induce this structural change. In addition, our structural analysis revealed conformational differences between Q23-HTT in the apo state and HTT in complex with HAP40. To compare the dynamics observed among Q23-HTT, Q78-HTT, and HTT_{HAP40}, we placed the three structures with the N-HEAT domains superimposed, revealing that there is a continuum in the movement of the C-HEAT domain among HTT_{HAP40}, Q23-HTT, and Q78-HTT structures (Figures 5A and 5B). C-HEAT in HTT_{HAP40} is most closely located to the N-HEAT, resulting in the most compact structure due to the binding of HAP40. In contrast, Q78-HTT adopts the most extended conformation, in that the C-HEAT domain is far away from the N-HEAT domain. The C-HEAT domain in Q23-HTT is located in the middle of these two conformations (Figure 5A). This analysis indicates that HTT has two pivot joints located at the beginning and at the end of the Bridge domain, which leads to the dynamic nature of HTT. Specifically, at the joint at the beginning of the Bridge domain (blue circles in Figure 5B), the Bridge domain can rotate along the y axis and can roll at the longest axis. At the joint at the end of Bridge domain (pale-red circles in Figure 5B), the C-HEAT domain rotates along the y axis and rolls. It is worth noting that the C-HEAT domain of HTT_{HAP40} is located in closer proximity to the N-HEAT domain than that of Q23-HTT, which is caused by HAP40 binding, and that the relative location of the C-HEAT domain to the N-HEAT domain is altered upon the polyQ expansion. This observation suggests that the changes in the relative orientations among the domains may modulate HTT function.

Polyglutamine Tract Length Alters Phosphorylation Patterns along the Entire Protein

The 3,144 residue Q23-HTT is predicted to possess many potential phosphorylated serine (S), threonine (T), or tyrosine (Y) residues (Blom et al., 2004). In fact, more than 70 phosphorylation sites across the entire protein have been collectively reported from different studies with human cells (Hornbeck et al., 2012), implying that phosphorylation of HTT can be significantly involved in its function. Since the overall structure of HTT can be affected by the polyQ tract expansion (Figure 4) but many known phosphorylation sites were not determined in any of the HTT cryo-EM structures, we decided to investigate

how the phosphorylation patterns of HTT are affected by the polyQ expansion. To evaluate the phosphorylation status at sites consistently identified in full-length HTT, we utilized Sf9 insect cells, an excellent surrogate for the mammalian system with many kinase families conserved (Busconi and Michel, 1995), to express recombinant full-length HTT with various polyQ lengths (Seong et al., 2010). To identify phosphorylated peptides, we purified Q23-, Q46-, and Q78-HTT in the presence of phosphatase inhibitors, and performed three independent liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses. Alignment of the HTT peptides revealed high coverage (~80% of 3,144 residues) and a generally similar set of peptides with phosphorylated serine or threonine residues (Figures 6A and S6A). Sixteen phosphorylation sites were identified more than twice. These included 15 previously reported sites and one site (S2423) previously unreported (Aiken et al., 2009; Anne et al., 2007; DeGuire et al., 2018; Dephoure et al., 2008; Dong et al., 2012; Hornbeck et al., 2012; Huang et al., 2015; Humbert et al., 2002; Luo et al., 2005; Moritz et al., 2010; Phanstiel et al., 2011; Rangone et al., 2004; Ratovitski et al., 2017; Schilling et al., 2006; Warby et al., 2005), confirming the utility of insect cells for studying phosphorylation patterns in HTT. Most phosphorylation sites except S2114, S2116, S2423, and S2550 located in the C-HEAT domain are located in disordered regions in the cryo-EM structures.

To systematically examine the potential impact of polyQ tract length on the pattern of HTT phosphorylation, we first generated a panel of phosphosite-specific affinity purified antibodies for each of these 16 sites, demonstrating their site and phosphorylation specificities in a dot-blot assay using the phospho and the corresponding non-phospho peptides (Figures S6A and S6B). Using this panel of phosphosite-specific antibodies, immunoblot analysis of the purified recombinant Q2-, Q23-, Q46-, Q67-, and Q78-HTT series demonstrated that for eight reagents, the band intensities (normalized to total HTT detected by mAb2166 and relative to Q2-HTT) increased with polyQ tract length (Figures 6B, S6B, and S6C; Table S4). This finding reveals that the size of the polyQ tract affected the level of HTT phosphorylation at sites (S421, S434, S1181, S1201, S1864, S1876, S2116, and T2337) across the entire protein instead of a cluster near the polyQ tract. It is plausible that the effect of polyQ expansion on the PTM is due to the global conformational change of full-length HTT, potentially through altering the relative position of the C-HEAT domain (Figure 5).

We then investigated the phosphorylation at 16 sites in endogenous HTT from HD patient cells to understand their phosphorylation levels and patterns depending on polyQ tract length in human cells. We chose lymphoblastoid cell lines (LCLs) which most highly express *HTT* mRNA (GTEx portal). We were able to detect specific signals from LCL lysates by using our HTT phospho antibodies (data not shown) but the phospho antibody signals were too variable to obtain statistically significant patterns with a few LCLs, likely due to the dynamic nature of cellular phosphorylation. For accurate and reliable quantification in multiple lines at the same time, we utilized a parallel reaction monitoring (PRM) strategy, a highly sensitive and quantitative targeted MS/MS technique, to quantify the phosphorylation level of each site in cellular HTT. We found seven phosphopeptides suitable for the PRM assay (Figure S6D). Six of these sites (pS421, pS434, pS1181, pS1201, pS1876, and pS2116) were significantly affected by polyQ tract length in the allelic series of purified HTT while pS1872 was not. Using the PRM assay, we successfully identified

and quantified substantial amounts of these seven phosphopeptides in endogenous HTT from all 24 LCLs expressing HTT with different polyQ lengths (Figures 6C and S6E), confirming phosphorylated HTT at these seven sites in human cells. To examine the impact of the polyQ tract length on these phosphosites, we performed polyQ length correlation linear regression analysis with their relative abundances and two sites, S1201 and S2116, showed nominal p values (Table S4). The phosphorylation levels of S434 and S1181 also seemed to be associated with polyQ tract length, although their p values were slightly higher than the nominal p value. All four phosphorylation sites are significantly polyQ length dependent in purified HTT (Figures 6B and 6C), suggesting a common impact of polyQ expansion on phosphorylation, likely through changes in HTT's structure. However, cellular context also appears to be a crucial determinant of HTT phosphorylation because the direction of pS2116 correlation with CAG size of *HTT* gene in LCL was opposite in insect cells. Together, our results suggest that HTT phosphorylation and polyQ expansion are inter-related through structure and our panel of purified human HTT with different polyQ tract sizes, and provides the basis to further study the effect of polyQ expansion on HTT structure and phosphorylation.

S2114/S2116 Phosphorylation in the C-HEAT Domain Related to HTT Structure and Activity

We then determined whether phosphorylation status might influence HTT structure and/or its functional activity. We previously showed that polyQ expanded HTT (Q78) stimulates PRC2 histone methyltransferase activity (Seong et al., 2010). Notably, this assay has previously been validated as a readout for normal function of full-length HTT (and the effect of its polyQ tract) on PRC2-dependent chromatin marks and regulation as demonstrated both *in vivo* (Seong et al., 2010) and in cell culture studies (Biagioli et al., 2015; Shin et al., 2018). Using this assay system as a readout, we examined the effect of phosphorylation on HTT function. To narrow down specific phosphorylation sites that may affect the PRC2 enhancing activity of mutant HTT, we selected S434, S1181, S1201, S2114, and S2116 sites exhibiting increased levels of phosphorylation as the length of polyQ increases, then purified five alanine mutants of both Q23- and Q78-HTT at S434, S1181, S1201, S2114, and S2116 (Figure S7A). Two out of five alanine mutant Q78-HTTs, S2114A and S2116A, exhibited significantly reduced activity compared with the Q78-HTT control (Figure 7A). Moreover, when we performed the PRC2 assay with Q23 S2116A-HTT and Q78 S2116A-HTT, the alanine mutation on S2116 dampened the extra PRC2 stimulating activity of Q78-HTT, whereas the same mutation in Q23-HTT apparently did not affect it (Figure 7B). Removing the phosphorylation at the S2116 site, therefore, removes the ''extra'' activity of Q78-HTT without losing the basal activity seen in Q23-HTT. Considering that the activity gained by the polyQ expansion at the N terminus was compromised by eliminating the phosphorylation of S2116 located within the C-HEAT, there may be a functional interaction between the N-HEAT and the C-HEAT of HTT, which may be critical for interaction with PRC2 and is affected by the S2116A mutation. To test whether the mutation disrupted the interaction between PRC2 and HTT, we performed immunoprecipitation assays of reconstituted HTT-PRC2 complex with EZH2, the PRC2 catalytic subunit. The S2116A mutation mildly increased the interaction between Q23-HTT and EZH2, whereas the same mutation in Q78-HTT dramatically decreased the interaction between Q78-HTT and EZH2, while control Q78-HTT robustly immunoprecipitated (Figure 7C). These results suggest that

S2116 phosphorylation uncovers a novel property of mutant HTT related to the interaction with PRC2.

To elucidate the effect of S2116A mutation on the global structure of HTT, we probed intra-molecular interactions within Q23- and Q78-HTT and corresponding S2116A mutants by XL-MS using DSS as a crosslinking reagent. This is a suitable technique for the characterization of the architecture of proteins, especially since XL-MS can provide structural information within regions unresolvable through cryo-EM due to current technical limitations (Leitner et al., 2014a). The crosslinked peptides from all samples were identified by the xQuest search engine (Table S5 and Figure S7) and the overall numbers were comparable (Table S6). Since HTT consists of three domains (Biddy et al., 2018), we grouped the crosslinked peptide pairs into the three groups and analyzed the inter-domain (N-B, N-C, B-C) and intra-domain (N-N, C-C, B-B) changes in XL-MS analysis (Figures S8A and S8B). To quantify the differences, we assigned a fold enrichment score for each of the six types of interactions by considering the magnitude and direction of change of the crosslinked peptides (the equation is presented in STAR Methods). Consistent with previous XL-MS studies of HTT protein (Vijayvargia et al., 2016), the Bridge domain corresponding to the uncrosslinked domain still showed unusually few crosslinks (Figure S8C). Figure 8A shows a decrease in inter-domain interactions between the N-HEAT domain and the C-HEAT domain in Q78-HTT compared with those in Q23-HTT (—0.8 fold change), indicating an increase in distance between the N-HEAT and C-HEAT domains in Q78-HTT. This is surprisingly consistent with the cryo-EM results (Figure 4C) that also show a structural change resulting in increased separation of the N-HEAT and C-HEAT domains in Q78-HTT. In addition, the increased Q78-HTT N-HEAT intra-domain cross-linking (Figure 7D, 2.6 fold change) was mostly restricted to a subdomain of the N-HEAT domain that we call the NTD-III (previously called CTD-I) region (Figure S8A) (Vijayvargia et al., 2016). This region is the pivot point linking the N-HEAT and Bridge domains (blue circles in Figure 5). These results are consistent with the idea that the Bridge domain serves as the pivot joint upon the polyQ tract expansion (Figure 5B). Thus, our XL-MS data support and further specify the structural changes between Q23- and Q78-HTT.

Most importantly, our XL-MS results revealed that the S2116A mutation specifically shifts the structural changes exhibited by Q78-HTT back toward Q23-HTT. S2116A mutation of Q78-HTT dramatically increased the inter-domain interaction between the N- and C-HEAT domains (Figure 8B), whereas the same mutation in Q23-HTT showed no or minimal impact (Figure S8B), indicating its mutant HTT specificity. When comparing Q23 S2116A-HTT and Q78 S2116A-HTT, we see that the C-HEAT-Bridge and N-HEAT-C-HEAT domain interactions of Q78 S2116A-HTT have shifted toward the pattern exhibited by Q23-HTT (Figure 8C). These results strongly suggest that S2116A mutation in the C-HEAT domain of mutant HTT may reinstate the intra-molecular proximity between N-HEAT and C-HEAT domains toward normal HTT (Figure 8D) with a concomitant reduction of the functional consequences of the elongated polyQ tract on mutant HTT, implying the importance of the dynamic movement of the domains on HTT interaction.

DISCUSSION

The *HTT* CAG expansion, encoding polyQ in HTT, is a direct cause of HD. However, how the polyQ expansion in HTT protein affects its biological activity is obscure. Here, we propose a working model that the polyQ expansion alters the global structure of HTT and phosphorylation, which is related to PRC2-stimulating activity of mutant HTT. Comparative structural analysis on our Q23/Q78 structures together with HTT_{HAP40} revealed several key insights about HTT structure and differences induced by polyQ expansion and protein binding.

First, superimposing N-HEAT domains from HTT_{HAP40}, Q23-HTT, and Q78-HTT clearly manifests a large conformational change among the domains with the N and C termini of the Bridge domain acting as critical pivot joints (Figure 5A). In comparison with HTT_{HAP40}, Q23-HTT exhibited an extended structure with substantial conformational change in the C-HEAT domain relative to the N-HEAT domain (Figures 5A and 5B). Cryo-EM analysis of Q78-HTT showed a similar conformational change in the C-HEAT domain leading to an even more extended structure as compared with Q23-HTT, which is caused by polyQ expansion (Figure 5B).

Our data suggest that HTT undergoes structural changes in response to HAP40 binding and polyQ length expansion (Figure 5B), whereby modular movements of the C-HEAT domain relative to the N-HEAT domain are involved. The C-HEAT domain moves outward from the N-NEAT domain in response to polyQ expansion and moves inward to the N-HEAT domain in response to HAP40 binding (Figure 5B). While large conformational changes occur among the domains, detectable changes in structure are limited to the inter-domain movement, as shown by HDX-MS analysis (Figure S5). Hence, HTT maintains a modular structure while exhibiting dynamic movements in response to protein binding and polyQ expansion. Our comparative structural analysis of HTT proteins in different states strongly suggests that the conformational change of the C-HEAT domain relative to the N-HEAT domain may play critical roles in HTT function, and the polyQ expansion may regulate this conformational change leading to modulation of HTT function. At this moment, the mechanism by which the expansion of polyQ tract located at the N terminus alters the conformations of other parts of HTT is not clear. The previous works on the structures of HTT, including this study, indicate that polyQ tractt is unstructured (Guo et al., 2018; Kim et al., 2009). Therefore, it is unlikely the conformational changes observed in this work are due to a direct conformational/structural change within the polyQ tract. However, it is plausible that the polyQ tract located at the N terminus might be directly interacting with other parts of HTT such as C-HEAT. Therefore, the expansion of polyQ tract may alter the interaction between polyQ tract and the other parts of HTT.

Besides the polyQ tract, a large portion of N-HEAT domain (aa 409–666 in Q23-HTT) in the high-resolution cryo-EM structure is unstructured in addition to several small regions in other domains. However, the regions linking the domains are largely well structured. Interestingly, many residues in the large unstructured region in N-HEAT domain are shown to be phosphorylated, and these modifications seem to modify HTT function. Therefore, it

is possible that the unstructured region and its modifications in N-HEAT might affect the global structure of HTT.

Using a series of HTT proteins with different polyQ lengths, we showed that polyQ expansion affects the phosphorylation level of specific sites across the entire protein and that many of these sites are located in disordered and/or surfaced-exposed regions, indicating that the phosphorylation sites are largely solvent accessible and able to interact with the polyQ tract and other potential binding partners. These data indicate that the polyQ expansion not only modulates the orientation of the modular structure of HTT but also affects the biochemical characteristics of the disordered regions in HTT, including PTMs. Our data give clear evidence showing the effect of the polyQ expansion on both HTT protein structure and phosphorylation status along the entire protein. However, the physiological effects of HTT phosphorylation state in the context of HD need to be carefully investigated, as the phosphorylation state could vary depending on the cell/tissue types. Among many PTM sites, S2116 phosphorylation highlights how the interaction between the N- and C-HEAT domains can act as a critical feature for HTT function, in the context of both normal biological function and understanding the HD mechanism.

Considering that the polyQ expansion induces global structural changes by altering the conformation of the domains, this may lead to changes in PTM patterns that regulate HTT function by providing a binding platform for other proteins. Alternatively, PTMs may be directly regulated by the polyQ expansion and in turn affect the conformation of HTT, modulating HTT function. Based on our observations of S2116 and its influential role in HTT structure and function, we propose a complex, bidirectional relationship: that polyQ expansion may modulate the equilibrium between the domain movements and PTM patterns, which in turn further stabilizes and perpetuates the altered equilibrium of the domain movements. This dynamic configuration of HTT may be supported by an elastic behavior of HEAT-repeat molecules suggested by molecular dynamics simulations of PR65A (Grinthal et al., 2010). A single phosphorylation site such as S2116 may act as a key regulator of crosstalk among multiple PTMs, which enact pleiotropic effects onto the protein, such as potentially reversing the numerous structural and functional changes brought about by polyQ expansion. To further evaluate the role of the phosphorylation at S2116, the cellular consequences of its phospho mutant need to be more intensively assessed, including identification of HTT phosphorylation status in a certain cell type. It should be noted that S2116 is located at the region interacting with HAP40 and phosphorylation of S2116 would affect the interaction between HTT and HAP40 (Figure S8I), implicating the importance of S2116 for HTT function. The extent of HTT phosphorylation appears to be largely dependent on cell type and conditions; for example, S421, S1201, and S1181 sites were reported to be extensively phosphorylated in normal compared with mutant HTT (Anne et al., 2007; Warby et al., 2005). Our HTT phosphorylation PRM assays will provide a comprehensive evaluation of the phosphorylation levels of endogenous HTT with our HTT phospho antibodies. It will be also critical to determine whether there are more cellular interactors that can be affected by the phosphorylation at S2116 in a polyQ length-dependent manner. A recent study explored the role of several HTT PTM sites on mouse primary neuron toxicity caused by the polyQ expansion using nuclear condensation and mitochondria viability assays, suggesting no significant effect of S2116A, while some sites,

such as S116 and S2652, inhibited the toxicity significantly (Arbez et al., 2017). However, this may indicate that the cellular consequence of phosphorylation depends on the specific target and cell status/type.

With the integration of structural, functional, and PTM characterization, we have built a proof of concept that polyQ expansion can exhibit structural consequences on HTT that can affect its functional activity in a modulatory fashion. From the various attempts to characterize HTT's biological function, it is evident that HTT does not have a single identifiable function or a well-defined set of binding partners, but rather participates in diverse cellular processes consistent with other HEAT-repeat proteins (Kobe et al., 1999). In fact, HEAT-repeat protein PR65A can serve as a potential conceptual model for HTT: PR65A binds to two separate components, regulatory and catalytic, to build serine-threonine protein phosphatase 2A (PP2A) complex, and the interchangeable binding with the various regulatory domains influences PP2A enzymatic activity, allowing its various functions (Shi, 2009). Moreover, PR65A phosphorylation is suggested as an *in vivo* mechanism for regulating PP2A complex signaling (Kotlo et al., 2014). Similarly, HTT, through PTMs and/or conformational changes, may present slightly modified binding surfaces to other proteins, resulting in a dynamic proteomic milieu for HTT. HD therapeutics may require a more nuanced approach to ameliorating the toxicity of the mutation based on understanding various HTT functions, since the efficacy of gene therapy has not yet been confirmed and disease-modifying interventions may significantly precede the pre-symptomatic phase (Caron et al., 2018).

Overall, the work presents a structural and molecular basis of the conformational flexibility of HTT mediated by polyQ expansion and phosphorylation and provides a lead for alternative strategies to understand the HD mechanism.

STAR☆METHODS

RESOURCE AVAILABILITY

Lead Contact—Requests for resources and reagents should be contacted to the Lead Contact, Ji-Joon Song (songj@kaist.ac.kr).

Materials Availability—Materials are available upon reasonable requests.

Data and Code Availability—EMDB ID for Q23-HTT, GFP-Q23-HTT, and Q78-HTT are EMD-4937, EMD-4944, and EMD-10793, respectively. PDB IDs for Q23-HTT and Q-78-HTT models are 6RMH and 6YEJ respectively. ProteomeXchange/PRIDE. ID PXD013907 for XL-MS.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A panel of 24 EBV-transformed lymphoblastoid cell lines (LCLs) from HD subjects (17 lines) and controls (7 lines) were grown in a standard culture condition (Shin et al., 2017) for PRM-based targeted mass spectrometry analysis in order to determine the profile of HTT phosphorylation.

METHOD DETAILS

Human FLAG-Huntingtin cDNA Clones—The information about FLAG-HTT cDNA cloned into pFASTBAC1 vector (Invitrogen) was mentioned in the previous paper (Vijayvargia et al., 2016). Additionally, for making eGFP fusion HTT clones, the Kas1 restriction site which is present in the 5′ end of HTT cDNA was used. The eGFP cDNA fragment was cloned into Kas1 site with In-Fusion Cloning Kit (Clontech). The eGFP-HTT construct contains a GSGS linker.

Purification of HTT and eGFP Fusion HTT—The purification of FLAG-tag HTT has been described (Vijayvargia et al., 2016). For purification of HTT, Baculovirus protein expression system (Invitrogen) was used. The Sf9 cell lysate was obtained by centrifugation at 39,000 x g, at 4°C for 2 h and after freeze and thaw method in Buffer A (100 mM NaCl, 50 mM Tris-HCl pH 8.0, protease inhibitor cocktail (Roche), 5% Glycerol). The supernatant was incubated with M2 anti-FLAG beads (Sigma-Aldrich) for 2 h at 4°C. After washing three times with Buffer A, bound proteins were eluted with Buffer B (100 mM NaCl, 50 mM Tris-HCl pH 8.0, 5% glycerol) containing 0.4 mg/ml Flag peptide. Eluted proteins were incubated in Buffer C containing150 mM NaCl, 50 mM Tris-HCl pH 8.0, 5% glycerol, 20 mM imidazole, 1 mM 1,4-dithiothreitol (DTT) and 0.5 mM EDTA with TEV protease for cleaving the N-terminal FLAG and 6X histidine tag (16 h, 4°C). Tag cleaved HTT was further purified by Ni-NTA Agarose beads (QIAGEN) and Superdex 200 26/60 column (GE Healthcare) in running buffer (150 mM NaCl, 20 mM HEPES-HCl pH 7.5). To maintain the monomeric state of HTT, purified HTT was crosslinked with 0.5 mM DSS crosslinker for 30 min at room temperature. The monomeric fraction of HTT was separated by 10–30% sucrose density gradient fractionation. Crosslinked monomeric HTT was concentrated with Amicon Ultra 100-kDa filters (Milipore). Q78-HTT and GFP-Q23-HTT were purified by same procedure.

Cryo-EM Grid Conditioning and Image Collection—For cryo-EM, the monomeric HTT (3.5 μl of 0.08 mg/ml) was mixed with final 0.05% of octyl glucoside and loaded on R2/1 300 mesh (Quantifoil) grids covered with graphene oxide (Bokori-Brown et al., 2016). The grids were incubated with HTT sample for 30 seconds, then vitrified by plunge freezing in liquid ethane using a Vitrobot II (Thermo Fisher) with 8~10 seconds blotting time. The temperature and humidity of the vitrification chamber were maintained at 15°C and 100% respectively. The grids were transferred to liquid nitrogen and loaded into an autoloader of Titan Krios microscope with a Gatan K2 Summit direct detector. Multiple data acquisitions were carried out for this study (Table S1). The images of Q23-, Q78-, and GFP-Q23-HTT were collected at the Diamond Light Source in the UK with a Volta phase plate (VPP). The detailed imaging conditions and parameters are summarized in Table S1.

Image Processing and Analysis—Movie frame stacks were imported into Scipion (de la Rosa-Trevin et al., 2016) to set a platform for accessing various program suites for further image processing. MotionCor2 (Zheng et al., 2017) was used for beam-induced motion correction and dose weighted and un-dose weighted micrographs were generated. The CTF estimation was done by using CTFFIND4 (Rohou and Grigorieff, 2015) and GCTF (He et al., 2016), an option for detecting additional phase shift applied on VPP images. Good

images were selected based on estimated CTF resolution (<8 Å), defocus range (for VPP images, between −0.2 to 1.0 μm), and the thickness of graphene oxide. All micrographs were thoroughly inspected and the particle picking was carried out fully manually. The particles were extracted from dose weighted micrographs with a 250 × 250 pixel box size and the contrast was inverted for further image processing. The particle sets were cleaned by 4 to 8 rounds of reference free RELION 2D classification, and particles in good 2D classes were subjected to 3D initial model building using RELION 3D initial model. The initial maps (~20 Å resolution) were used as a starting model for refinement of the structure by using 3D classification and 3D auto refinement with selected particle sets in RELION (Figures S1–S3).

UCSF-Chimera (Pettersen et al., 2004) and Pymol (Alexander et al., 2011) were mainly used for further image analysis and validation. The estimated resolution was calculated from the EMDB FSC validation server (https://www.ebi.ac.uk/pdbe/emdb/validation/fsc/) using pairs of unfiltered half maps (Table S1).

Integrative Modeling of Q23 -HTT—The Q23-HTT was modeled at atomistic resolution using an integrative modeling strategy which consisted in (i) splitting the HTT_{HAP40} structure into three rigid subunits and (ii) rigidly fitting the subunits into the density map with a multi-body constrained molecular docking strategy. Details of the distinct modeling steps are described in the following sections.

Splitting HTT_{HAP40} Structure into Rigid Subunits—The HTT_{HAP40} was split into three major subunits consisting in a N-HEAT domain (residue number 91 to 1684), Bridge domain (residue number 1685 to 2091) and C-terminal HEAT domain (residue number 2092 to 3098). The missing residues corresponding to flexible regions of the HTT_{HAP40} were not directly included in the current modeling procedure.

Rigid Docking of HTT Subunits Inside cryo-EM Maps—First, the N-HEAT domain subunit density map was rigidly fit with SITUS (Wriggers et al., 1999) inside the Q23-HTT cryo-EM density map. Due to the fact the current version of SITUS could not simultaneously fit the three subunits, the fitting protocol was initiated by fitting the N-HEAT domain, which was the largest of the three partial models, into the density map using the SITUS module *colores*.

Later, the docking problem consisting in simultaneously fitting the remaining Bridge and C-HEAT domains was modeled as a constrained optimization. The aim of this approach was to optimally fit the Bridge and C-HEAT domains into the density map whilst avoiding steric clashes with the already SITUS-fitted N-HEAT domain and keeping their native inter-subunit connectivity.

To perform such a task, the mViE (memetic viability evolution) algorithm was coupled to our in-house docking protocol *pow*^{er} (parallel optimization workbench to enhance resolution). The combination of *pow*^{er} and mViE to solve integrative modeling problems has been thoroughly described in (Tamo et al., 2017) and thus only a brief explanation is given here. *Pow*^{er} is an integrative modeling platform that was designed to virtually solve any

type of optimization problems (Degiacomi and Dal Peraro, 2013). It has been successfully applied to solve a multitude of protein assembly problems (Tamo et al., 2015) by using state-of-the-art heuristic optimization algorithms to efficiently sample the large search space associated with molecular docking (Degiacomi and Dal Peraro, 2013).

Amongst the list of optimizers featured in *pow*^{er}, mViE is a recently implemented constraint optimization algorithm that evolves a population based on (1+1)-covariance matrix adaptation evolution strategy (CMA-ES) for local search of the search space, which can recombine using differential evolution for global search (Floreano, 2015). To optimize the balance between local and global search, mViE uses a dedicated task scheduler. Importantly, mViE was shown to outperform several state-of-the-art constrained optimizers on a standard constrained optimization benchmark set (Floreano, 2015). The combination of *pow*^{er} with mViE was recently shown to effectively solve the problem of fitting macromolecular assemblies inside density maps (Tamo et al., 2017), and thus was used in this work to model the Q23-HTT structure using structural information from the experimental cryo-EM density map.

The process resulting in the optimal fit of HTT subunits into the map was encoded as a contrained optimization problem. In this context, the search space sampled by *pow*^{er}-mViE consisted in the orientation and position of each HTT subunit which were encoded as [x_{Cterm}, y_{Cterm}, z_{Cterm}, α_{Cterm}, β_{Cterm}, γ_{Cterm}, x_{Bridge}, y_{Bridge}, z_{Bridge}, α_{Bridge}, β_{Bridge}, γ_{Bridge}], where x, y, and z corresponded to three translation sampled by the C-HEAT and Bridge domains respectively; and α, β *and* γ the Eulerian angles defining the spatial orientation of respective domains. These spatial parameters were simultaneously sampled in order to solve the problem formulated as:- Where $ccc = \frac{N}{N_{sol}} \frac{\sum_i^N (L \otimes ref)_i (L \otimes sol)_i}{\sqrt{\sum_k^N ref_k^2 sol_k^2}}$ corresponded to the CCC quantitatively estimating the fit between the assembled HTT subunits and the experimental cryo-EM map. In this case, L corresponds to a 3×3×3 Laplacian filter kernel used to convolute the input density maps with the aim to increase the precision of the density map fitting. N is the number of overlapping voxels between the density maps, N_{sol} the total number of voxel of the candidate solution density map, $(L \otimes ref)_i$ is the i^{th} voxel of the reference density map and $(L \otimes sol)_i$ its counterpart from the candidate monomer assembly.

The CCC was maximized upon the satisfaction of pre-defined inequality constraints. The inequality constraints were defined by lb_i and ub_i, which were the lower and upper lower bound defining the search space of X_i; ud_i which corresponded the upper Euclidian distances on Cα atoms respectively defining the spatial connectivity between residue 1713 – 1729, residue 2062 – 2092, and between two distances extracted from XL-MS experiments which were residue 943 – 1973 and residue 882 – 1882 to correctly orient the Bridge domain relative to the N-HEAT domain. These distance constraints were intentionally chosen to be < 45 Å as they were used to coarsely guide the assembly process, rather than provide accurate positions; $E_{phys} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^9 - \left(\frac{\sigma}{r}\right)^6\right]$ is a 9–6 Lennard-Jones coarse energy potential (Alemani et al., 2010) used here to prevent steric clashes, where r consists in the pairwise distance

between the Cα interfacial atoms of the subunits within a distance of 12 Å, σ = 4.7 Å and ε = 1 kcal/mol. In order to ensure a thorough sampling of the search space, a total budget of 1.8 million function evaluation was allocated to the *pow*^{er}-mViE optimization protocol. The best model ranked by constraint satisfaction and CCC value was extracted.

Integrative Modeling of Q78-HTT—The knowledge acquired from fitting the Q23-HTT domains positions enabled to detect the approximate location of N-, C-term and Bridge domain inside Q78-HTT density map. Due to the low-resolution of Q78-HTT cryo-EM map at the location of Q78-HTT Bridge domain, the Bridge structure was omitted from the final Q78-HTT model.

The rigid fit of both Q78-HTT N- and C-terminal domains was performed in successive steps. First, N-HEAT domain was fitted into Q78-HTT cryo-EM map using SITUS (Wriggers et al., 1999). Then, Q78-HTTN C-HEAT was fitted using *pow*^{er} inside its C-terminal map region. To coarsely guide the fitting of the C-HEAT domain, the Bridge domain was manually inserted close to the C-terminal region of N-HEAT domain. Precisely, the distance between residue 2062 of the Bridge and 2092 of C-HEAT domain was used during docking as a restraint (<40 Å). A total budget of 1 million function evaluations was allocated to the *pow*^{er}-mViE optimization protocol. The Bridge structure was however omitted from the final model due to the low resolution of the map at the location of Bridge domain.

Hydrogen-Deuterium Exchange Mass Spectrometry—The labeling reaction was started by mixing 3 μL of each protein batch (Q23-, or Q78-HTT) with 23 μL of deuterated buffer. For kinetics, deuteration time was 10 min, 45 min, 90 min, and 180 min (all incubations were done at room temperature). Each time point of the deuterium-labeling reaction was stopped by decreasing the pH to ~2.3 and temperature to ~4°C by means of adding 30 μL ofa solution containing 250 mM TCEP, 3 M Urea and 200 mM phosphate buffer pH 2.3.

Each labeled and quenched sample was analyzed in a semi-automated HDX-MS system (Biomotif AB, Stockholm, Sweden) in which manually injected samples were automatically digested, cleaned and separated at 2°C. Deuterated samples were digested using a Poroszyme immobilized pepsin column (Applied Biosystems 2.1 mm × 30 mm) by a 1 min at 70 μl/min flow protocol, followed by an on-line desalting step using a 2 mm I.D × 10 mm length C-18 pre-column (ACE HPLC Columns, Aberdeen, UK) using 0.05% TFA at 250 μl/min for 3 min. Peptic peptides were then separated by the following gradient (Buffer A consisted of 0.3% formic acid and 5% acetonitrile and buffer B consisted of 0.3% formic acid, 95% acetonitrile and 5% water): 5% B buffer at 0 min, 20% B buffer at 7 min, 35% B at 15 min, 90% B in 20 min and 5% B at 25 min using a Chromolith FastGradient RP-C-18, 50 mm length and 2 mm diameter analytical column (Merck) operated at 120 μl/min. An Orbitrap XL mass spectrometer (Thermo Fisher Scientific) operated at 60,000 resolution at m/z 400 was used for analysis. For peptide identification, several LC MS/MS runs were performed and searched against an in-silico database consisting of the Q23-, or Q78-HTT sequence. All obtained HDX-MS data were analyzed by the HDExaminer software (Sierra Analytics, USA).

Small-Angle X-ray Scattering—The data were collected at the European Synchrotron Radiation Facility (ESRF), beamline BM29 (Pernot et al., 2013). The samples crosslinked with DSS were concentrated to 8.8 mg/ml and were loaded onto a Superose 6 10/300 GL column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.5, 150 mM NaCl, 10% glycerol and 2 mM TCEP. The flow rate was set to 0.5 ml/min and the temperature to 18°C. As they were eluted from the column, the samples were illuminated with X-rays of wavelength 0.99 Å. A PILATUS 1M detector was used at a distance to the sample of 2.86 m covering a momentum transfer of $0.025 < q > 5$ nm-1 ($q = 4\pi\sin(\theta)/\lambda$, where θ is the total scattering angle). During the 50 min elution, 3000 scattering measurements were taken with 1s time-frames. SAXS data were analyzed using the ATSAS package and ScÅtter (Franke et al., 2017; JA, 2011). For each sample, using CHROMIXS (chromatography in-line X-ray scattering), an elution profile was generated with the integrated intensities plotted versus the recorded frame number. In CHROMIXS, ~50 buffer frames were averaged and used to (i) subtract the buffer average from each frame of the sample peak selected and (ii) calculate the corresponding Radius of Gyration (Rg). Only measurements with equal Rg values were included in the final sample peak region chosen. The subtracted sample peak region was averaged to generate the final scattering curve used for subsequent analysis. The scattering curves were initially viewed in PRIMUS (Konarev et al., 2003) where the Rg was obtained from the slope of the Guinier plot (Guinier, 1939) within the region defined by $qmin < q < qmax$ where $qmax < 1.3/Rg$ and qmin is the lowest angle data point included by the program. The P(r) function, the distribution of the intra-atomic distances (r) in the particle, was generated using the indirect transform program GNOM (Konarev et al., 2003). The maximum distance (Dmax) was selected by letting the P(r) curve decay smoothly to zero. DATPOROD, DATOW, and DATVC (Fischer et al., 2010; Petoukhov et al., 2007; Rambo and Tainer, 2013) were used to estimate the Porod Volume (Vp) and the concentration-independent estimate of the MW for the samples.

Mass Spectrometric Identification of Phosphorylation Sites—For the determination of phosphorylation sites, recombinant full-length human HTT proteins with polyglutamine tract length of 23, 46 and 78 were purified in the presence of complete protease and phosphatase inhibitor cocktails (Roche Applied Science, Branford, CT) to retain phosphorylation. 10–20 μg of purified proteins were separated by SDS-PAGE and stained with mass spectrometry compatible GelCode Blue stain reagent (Thermo Scientific, Carlsbad, CA). HTT bands were excised from gel and processed for mass spectrometry. The samples were reduced with 1 mM DTT for 30 min at 60°C and then alkylated with 5 mM iodoacetamide for 15 min in the dark at room temperature. Gel pieces were then subjected to a modified in-gel trypsin digestion procedure (Shevchenko et al., 1996). LC-MS/MS analysis of the digests was carried out on an LTQ-Orbitrap mass spectrometer (ThermoFinnigan, San Jose, CA). The eluted peptides were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching protein or translated nucleotide databases with the acquired fragmentation pattern by the software program TurboSEQUEST v.27 (ThermoFinnigan, San Jose, CA) (Eng et al., 1994). The modification of 79.9663 mass units to serine, threonine, and tyrosine was included in the database searches to determine phosphopeptides. Each phosphopeptide that was

identified by the SEQUEST program was also manually inspected to ensure confidence. Phosphopeptides obtained were manually aligned to the HTT sequence to generate the coverage map. Amino acid count in peptides was plotted to determine the sequence coverage.

Antibodies—HTT phosphorylation site specific affinity purified rabbit polyclonal antibodies against the 16 sites identified in this study were generated by New England Peptide (Gardner, MA). In short, KLH-coupled peptides bearing the specific phosphorylated residue were injected into 2 rabbits each. Booster doses were given on day 14 and 28. Production bleeds were checked for antibody titer by ELISA using purified phosphopeptides. High-titer bleeds (40–50 mL) were subjected to a double affinity purification. The first step involved negative selection wherein affinity columns with non-phosphorylated peptide backbones were used. The antibodies against the peptide backbone are retained in the column while those against the phosphorylated residue are in the flow through. The flow through is then subjected to a second cycle of affinity purification (positive selection) wherein phosphorylated peptides are used to specifically interact with phosphorylation site specific antibodies. The bound antibodies are eluted by lowering the pH. The specificity of phosphorylation site specific antibodies was validated by a dot-blot assay. Briefly, 100 ng of phosphorylated and non-phosphorylated peptides were spotted onto a nitrocellulose membrane and allowed to dry. The membrane was blocked with 5% BSA/TBST and probed overnight with 0.25 μg/mL of different phosphorylation site specific antibodies in 5% BSA/TBST at 4°C. After washing, the blots were probed with anti-Rabbit HRP secondary antibodies and developed using the ECL system. Anti-HTT mouse monoclonal antibody mAb2166 was purchased from Millipore (EMD Millipore, Darmstadt, Germany).

Immunoblot Analysis—50–100 ng of purified protein was run on a 6% or 10% Bis-Tris gel and transferred onto nitrocellulose membranes. All antibodies were blocked with 5% milk/TBST. However, phosphorylation site specific antibodies were used at 0.25 μg/mL in 5% BSA/TBST and incubated overnight at 4°C for probing. Anti-HTT antibodies were used at dilutions of 1:2000 (mAb2166). The data obtained by western blotting were quantified by subjecting the scanned images to densitometry using Image J software (NIH).

In Gel Tryptic Digestion of Endogenous HTT from Human Lymphoblastoid Cell Lines (LCL)—Gel slices were cut into 1×1 mm pieces and placed in 1.5 mL eppendorf tubes with 1 mL of water for 30 min. The water was removed and 50 μL of 250 mM ammonium bicarbonate was added. For reduction 10 μL of a 45 mM solution of DTT was added and the samples were incubated at 50°C for 30 min. The samples were cooled to room temperature and then for alkylation 10 μL of a 100 mM iodoacetamide solution was added and allowed to react for 30 min. The gel slices were washed 2 times with 1 mL water aliquots. The water was removed and 1 mL of 50:50 (50 mM ammonium bicarbonate: acetonitrile) was placed in each tube and samples were incubated at room temperature for 1 h. The solution was then removed and 200 μL of acetonitrile was added to each tube at which point the gels slices turned opaque white. The acetonitrile was removed and gel slices were further dried in a Speed Vac. Gel slices were rehydrated in 70 μL of 2 ng/μL trypsin

(Sigma) in 0.01% ProteaseMAX Surfactant (Promega): 50 mM ammonium bicarbonate. Additional bicarbonate buffer was added to ensure complete submersion of the gel slices. Samples were incubated at 37°C for 21 h. The supernatant of each sample was then removed and placed in a separate 1.5 mL eppendorf tube. Gel slices were further dehydrated with 100 μL of 80:20 (acetonitrile:1% formic acid). The extract was combined with the supernatants of each sample. The samples were then dried down in a vacuum centrifuge. Samples were dissolved in 30 μL of 5% acetonitrile in 0.1% trifluoroacetic acid prior to LC-MS/MS analysis.

HTT Phosphopeptides Quantification Using Parallel Reaction Monitoring (PRM) Assay—A 3.0 μL aliquot was directly injected onto a custom packed 2 cm × 100 μm C_{18} Magic 5 μm particle trap column. Peptides were then eluted and sprayed from a custom packed emitter (75 μm × 25 cm C_{18} Magic 3 μm particle) with a linear gradient from 95% solvent A (0.1% formic acid in water) to 35% solvent B (0.1% formic acid in Acetonitrile) in 60 min at a flow rate of 300 nL per minute on a Waters Nano Acquity UPLC system. Ions were introduced by positive electrospray ionization via liquid junction into a Q Exactive hybrid mass spectrometer (Thermo Scientific). Data-dependent acquisition (DDA) experiments for library generation were performed according to an experiment where full MS scans from 300–1750 m/z were acquired at a resolution of 70,000 followed by 10 MS/MS scans acquired under HCD fragmentation at a resolution of 17,500 with an isolation width of 1.6 Da. For parallel reaction monitoring (PRM) a total of 13 m/z's representing phosphorylated and control peptides observed in the DDA analyses were continuously monitored throughout the gradient using an isolation width of 2 Da followed by HCD fragmentation at a collision energy of 27% acquiring MS/MS spectra at a resolution of 17,500. A PRM analysis run in technical triplicates was performed on a cohort of 24 HD patient samples. Raw data files were processed with Proteome Discoverer (Thermo, version 1.4) prior to searching with Mascot Server (Matrix Science, version 2.5) against a human database which contained both variants of the HTT protein. Search parameters utilized were fully tryptic with 2 missed cleavages, parent mass tolerances of 10 ppm and fragment mass tolerances of 0.05 Da. A fixed modification of carbamidomethyl cysteine and variable modifications of acetyl (protein N-term), pyro glutamic for N-term glutamine, oxidation of methionine, and phosphorylation of Ser/Thr were considered. Quantitation of targeted peptides was accomplished through the use of the Skyline software (University of Washington, version 3.5). Spectral libraries used by Skyline were constructed from Proteome Discoverer search result MSF files generated from DDA runs. Top three fragment ion intensities were extracted from each of the 7 peptides and summed to determine the peptide abundance.

Cell-free Assay for PRC2 Methyltransferase Activity—Reconstituted *in vitro* PRC2 activity assays consisting of G5E4 DNA with 12 nucleosomes were performed essentially as previously described (Seong et al., 2010) using 20 nM of full-length HTT (Q23 and Q78 including corresponding alanine mutants) to modulate phosphorylation status with 10 nM of PRC2 complex.

Immunoprecipitation—To determine the effect of HTT phosphorylation status on its interaction with PRC2 complex, 2 μg of HTT proteins (Q23- and Q78-HTT including corresponding S2116A mutants) were mixed with 2 μg of purified PRC2 complex or buffer and incubated together for 2 h at 4°C in 100 μL of 50 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.5 mM EDTA. 1–2 μg of anti-EZH2 antibody was added subsequently and further incubated overnight at 4°C. The antigen-antibody complex was pulled down by the addition of 50 μL of protein-G agarose slurry (Roche Applied Science, Branford, CT). Non-specific binding was removed by excessive washing of the beads (6 times) with wash buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.5 mM EDTA) and the complex was eluted by adding 50 μL of Laemmli sample buffer followed by boiling at 95°C for 5 min. The eluted samples were separated by SDS-PAGE and subjected to immunoblotting with anti-HTT antibody.

Cross-linking Mass Spectrometric Analysis of HTT

Preparation of Cross-linked HTT Monomer: Cross-linking of full-length HTT was carried out as previously described (Vijayvargia et al., 2016). Briefly, HTT was cross-linked with 0.5 mM of DSS-H12/D12 (Creative Molecules) for 30 min at 25°C. The reaction was quenched by adding 20 mM of ammonium bicarbonate. Monomeric fraction of HTT was separated with 10–30% sucrose gradient by ultracentrifugation at 111,541 × g for 16 h at 4°C.

Sample Processing for MS Analysis: Cross-linked and quenched samples were dried in a vacuum centrifuge and redissolved in 50 μL of 8 M urea. Cysteines were reduced by addition of tris(2-carboxyethyl)phosphine to a final concentration of 2.5 mM and incubation for 30 min at 37°C with mild shaking. Free thiol groups were subsequently carbamidomethylated by addition of iodoacetamide to 5 mM and incubation at room temperature for 30 min in the dark. Alkylated samples were diluted to 5.5 M urea by addition of 150 mM ammonium bicarbonate and 500 ng of endoproteinase Lys-C (Wako) were added, followed by incubation for 3 h at 37°C. The samples were then further diluted to 1 M urea by addition of 50 mM ammonium bicarbonate, 1 μg of trypsin (Promega) was added and samples were incubated overnight at 37°C. Digestion was stopped by addition of formic acid to 2% (v/v) and samples were cleaned up using 50 mg Sep-Pak tC18 cartridges (Waters). Eluates were evaporated to dryness in a vacuum centrifuge and resuspended in 20 μL water/acetonitrile (70:30, v/v) for fractionation by size exclusion chromatography (SEC) (Leitner et al., 2012). SEC was performed on an Äkta micro system equipped with a Superdex Peptide PC 3.2/300 column (both GE Healthcare) and a mobile phase consisting of water/acetonitrile/trifluoroacetic acid (70:30:0.1, v/v/v), operating at a flow rate of 50 μL/min. Three 100 μL fractions corresponding to the elution window of cross-linked peptides were collected separately and evaporated to dryness.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS): Dried SEC fractions were redissolved in 20 μL of water/acetonitrile/formic acid (95:5:0.1, v/v/v) and 4 μL were used for LC-MS/MS analysis. All injections were performed in duplicate. LC-MS/MS was performed using an Easy-nLC 1000 nanoflow HPLC system coupled to an Orbitrap Elite mass spectrometer equipped with a Nanoflex source (all ThermoFisher Scientific). Peptides were separated on an Acclaim PepMap RSLC C18 column (15 cm × 75

μm, 2 μm particle size, ThermoFisher Scientific) using mobile phases A (water/acetonitrile/formic acid, 95:5:0.15, v/v/v) and B (acetonitrile/water/formic acid, 95:5:0.15, v/v/v) with a gradient of 9 to 35%B in 90 min and a flow rate of 300 nL/min. Data acquisition was configured with an MS scan in the Orbitrap analyzer (120,000 resolution), fragmentation of selected precursors in the linear ion trap (isolation with of 2.0 m/z, normalized collision energy of 35%) and acquisition of 10 dependent MS/MS scans in the linear ion trap at normal resolution. Dynamic exclusion was activated with an exclusion duration of 30 s and a repeat count of 1.

Identification of Cross-linked Peptides with xQuest: Cross-linked peptides were identified with xQuest, version 2.1.4 (Leitner et al., 2014b; Rinner et al., 2008). Searches were performed against a database consisting of the huntingtin Q23 wild-type sequence and 13 contaminants (tubulins, heat shock proteins and human keratins). Search settings included enzyme = trypsin, maximum number of missed cleavages = 2, mass tolerance = 15 ppm for MS data and 0.2/0.3 Da for MS/MS data. All candidate intra-protein cross-link identifications with a score ≥16 were further filtered with a narrower mass tolerance of ±5 ppm and a %TIC subscore ≥0.1. All remaining spectra were manually evaluated and only assignments with at least four bond cleavages overall or three consecutive bond cleavages per peptide were considered further. Because the likelihood of false positives for intra-protein cross-links at this score threshold is very low (Walzthoeni et al., 2012), we expect a false discovery rate of <1% for these datasets. For all samples combined, only five cross-links on contaminants were identified in total. The identified cross-linked pairs were grouped based upon lysine residues in the primary HTT sequence and unique cross-links number was counted. The fold enrichment score for each type of interaction was calculated based on the equation below:

$$|B - A| \times \log_2 B/A$$

A: Number of identified cross-linked pairs of Q23- or Q78-HTTB: Number of identified cross-linked pairs of Q78- or Q78 S2116A-HTT

QUANTIFICATION AND STATISTICAL ANALYSIS

The significance of the enhanced PRC2 methyltransferase activity by adding full-length HTT (Q23 and Q78 including corresponding alanine mutants) in Figure 7 was calculated and displayed using Student's *t*-test statistical analysis from three experiments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work is partially supported by grants (NRF-2016K1A1A2912057 and NRF-2020R1A2B5B03001517 to J.S.) by the National Research Foundation (NRF) of Korea, South Korea, and also supported by National Institutes of Health/National Institute of Neurological Disorders and Stroke (R01 NS079651 and R03 NS108028 to I.S.S.), United States; CHDI Foundation, United States to I.S.S.. T.J. and H.H. were supported by the Swedish Research Council (project grant 2016–03810), Sweden. The group of R.A. was supported by the European Research Council

(ERCaG grant 670821 PROTEOMICS4D), European Union. The group of M.D.P. was supported by the Swiss National Science Foundation (grant number 200021_157217), Switzerland. This work was partially supported by the Intelligent Synthetic Biology Center of Global Frontier Project funded by the Ministry of Science and ICT (MSIT) (2011–003955), South Korea. R.V. was supported by DST-SERB grant (EMR/2016/001301), India. We are grateful to Dr. Ross Tomaino (Taplin mass spectrometry facility) and Dr. Jon Leszyk (UMASS medical school mass spectrometry facility) for excellent technical support. We also thank the beamline scientists at the Beamline BM29 of the European Synchrotron Radation Facility (ESRF). The computing resource for cryo-EM analysis was supported by Global Science experimental Data hub Center, Korea Institute of Science and Technology Information (KISTI) (NRF-2018R1A6A7052113 to J.S.). For data collection, we acknowledge the staff at the UK National Electron Bio-Imaging Center (eBIC), Diamond, funded by the Wellcome Trust, MRC and BBSRC, United Kingdom and at the SciLifeLab, Swedish National cryo-EM Facility, funded by the Knut and Alice Wallenberg, Family Erling Persson, and Kempe Foundations, Sweden.

REFERENCES

Adams PD, Mustyakimov M, Afonine PV, and Langan P (2009). Generalized X-ray and neutron crystallographic analysis: more accurate and complete structures for biological macromolecules. 65, 567–573.

Aiken CT, Steffan JS, Guerrero CM, Khashwji H, Lukacsovich T, Simmons D, Purcell JM, Menhaji K, Zhu YZ, Green K, et al. (2009). Phosphorylation of threonine 3: implications for Huntingtin aggregation and neurotoxicity. J. Biol. Chem 284, 29427–29436. [PubMed: 19710014]

Alemani D, Collu F, Cascella M, and Dal Peraro M (2010). A nonradial coarse-grained potential for proteins produces naturally stable secondary structure elements. J. Chem. Theory Comput 6, 315–324. [PubMed: 26614340]

Alexander N, Woetzel N, and Meiler J (2011). bcl::Cluster: a method for clustering biological molecules coupled with visualization in the Pymol Molecular Graphics System. IEEE Int. Conf. Comput. Adv. Bio Med. Sci 2011, 13–18.

Andrade MA, and Bork P (1995). HEAT repeats in the Huntington's disease protein. Nat. Genet 11, 115–116. [PubMed: 7550332]

Anne SL, Saudou F, and Humbert S (2007). Phosphorylation of huntingtin by cyclin-dependent kinase 5 is induced by DNA damage and regulates wild-type and mutant huntingtin toxicity in neurons. J. Neurosci 27, 7318–7328. [PubMed: 17611284]

Arbez N, Ratovitski T, Roby E, Chighladze E, Stewart JC, Ren M, Wang X, Lavery DJ, and Ross CA (2017). Post-translational modifications clustering within proteolytic domains decrease mutant huntingtin toxicity. J. Biol. Chem 292, 19238–19249. [PubMed: 28972180]

Biagioli M, Ferrari F, Mendenhall EM, Zhang Y, Erdin S, Vijayvargia R, Vallabh SM, Solomos N, Manavalan P, Ragavendran A, et al. (2015). Htt CAG repeat expansion confers pleiotropic gains of mutant huntingtin function in chromatin regulation. Hum. Mol. Genet 24, 2442–2457. [PubMed: 25574027]

Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, Sun T, and Morris SA (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature 564, 219–224. [PubMed: 30518857]

Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, and Brunak S (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4, 1633–1649. [PubMed: 15174133]

Bokori-Brown M, Martin TG, Naylor CE, Basak AK, Titball RW, and Savva CG (2016). Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. Nat. Commun 7, 11293. [PubMed: 27048994]

Busconi L, and Michel T (1995). Recombinant endothelial nitric oxide synthase: post-translational modifications in a baculovirus expression system. Mol. Pharmacol 47, 655–659. [PubMed: 7536885]

Caron NS, Dorsey ER, and Hayden MR (2018). Therapeutic approaches to Huntington disease: from the bench to the clinic. Nat. Rev. Drug Discov 17, 729–750. [PubMed: 30237454]

Ciferri C, Lander GC, and Nogales E (2015). Protein domain mapping by internal labeling and single particle electron microscopy. J. Struct. Biol 192, 159–162. [PubMed: 26431894]

Cui X, Liang Q, Liang Y, Lu M, Ding Y, and Lu B (2014). TR-FRET assays of Huntingtin protein fragments reveal temperature and polyQ length-dependent conformational changes. Sci. Rep 4, 5601. [PubMed: 24998512]

Degiacomi MT, and Dal Peraro M (2013). Macromolecular symmetric assembly prediction using swarm intelligence dynamic modeling. Structure 21, 1097–1106. [PubMed: 23810695]

DeGuire SM, Ruggeri FS, Fares MB, Chiki A, Cendrowska U, Dietler G, and Lashuel HA (2018). N-terminal Huntingtin (Htt) phosphorylation is a molecular switch regulating Htt aggregation, helical conformation, internalization, and nuclear targeting. J. Biol. Chem 293, 18540–18558. [PubMed: 30185623]

Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, and Gygi SP (2008). A quantitative atlas of mitotic phosphorylation. Proc. Natl. Acad. Sci. U S A 105, 10762–10767. [PubMed: 18669648]

Dong G, Gross K, Qiao F, Ferguson J, Callegari EA, Rezvani K, Zhang D, Gloeckner CJ, Ueffing M, and Wang H (2012). Calretinin interacts with huntingtin and reduces mutant huntingtin-caused cytotoxicity. J. Neurochem 123, 437–446. [PubMed: 22891683]

Eng JK, McCormack AL, and Yates JR (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom 5, 976–989. [PubMed: 24226387]

Fischer H, de Oliveira Neto M, Napolitano HB, Polikarpov I, and Craievich AF (2010). Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. J. Appl. Crystallogr 43, 101–109.

Floreano AMGID (2015). Memetic viability evolution for constrained optimization. IEEE Trans. Evol. Comput 20, 125–144.

Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM, et al. (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J. Appl. Crystallogr 50, 1212–1225. [PubMed: 28808438]

Grant T, Rohou A, and Grigorieff N (2018). cisTEM, user-friendly software for single-particle image processing. eLife 7, e35383. [PubMed: 29513216]

Grinthal A, Adamovic I, Weiner B, Karplus M, and Kleckner N (2010). PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. Proc. Natl. Acad. Sci. U S A 107, 2467–2472. [PubMed: 20133745]

Guinier A (1939). La diffraction des rayons X aux trè s petits angles: application à l'é tude de phé nomè nes ultramicroscopiques. Ann. Phys 11, 161–237.

Guo Q, Bin H, Cheng J, Seefelder M, Engler T, Pfeifer G, Oeckl P, Otto M, Moser F, Maurer M, et al. (2018). The cryo-electron microscopy structure of huntingtin. Nature 555, 117–120. [PubMed: 29466333]

Gusella JF, and MacDonald ME (2000). Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. Nat. Rev. Neurosci 1, 109–115. [PubMed: 11252773]

Halder K, Dolker N, Van Q, Gregor I, Dickmanns A, Baade I, Kehlenbach RH, Ficner R, Enderlein J, Grubmuller H, et al. (2015). MD simulations and FRET reveal an environment-sensitive conformational plasticity of importin-beta. Biophys. J 109, 277–286. [PubMed: 26200863]

Harjes P, and Wanker EE (2003). The hunt for huntingtin function: interaction partners tell many different stories. Trends Biochem. Sci 28, 425–433. [PubMed: 12932731]

He WT, Zheng XM, Zhang YH, Gao YG, Song AX, van der Goot FG, and Hu HY (2016). Cytoplasmic ubiquitin-specific protease 19 (USP19) modulates aggregation of polyglutamine-expanded Ataxin-3 and huntingtin through the HSP90 chaperone. PLoS One 11, e0147515. [PubMed: 26808260]

Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, Latham V, and Sullivan M (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res 40, D261–D270. [PubMed: 22135298]

Huang B, Lucas T, Kueppers C, Dong X, Krause M, Bepperling A, Buchner J, Voshol H, Weiss A, Gerrits B, et al. (2015). Scalable production in human cells and biochemical characterization of full-length normal and mutant huntingtin. PLoS One 10, e0121055. [PubMed: 25799558]

Humbert S, Bryson EA, Cordelieres FP, Connors NC, Datta SR, Finkbeiner S, Greenberg ME, and Saudou F (2002). The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt. Dev. Cell 2, 831–837. [PubMed: 12062094]

Rambo RP, and Tainer JA (2011). ScÅtter https://bl1231.als.lbl.gov/scatter/.

Jeong SJ, Kim M, Chang KA, Kim HS, Park CH, and Suh YH (2006). Huntingtin is localized in the nucleus during preimplanatation embryo development in mice. Int. J. Dev. Neurosci 24, 81–85. [PubMed: 16289942]

Kim MW, Chelliah Y, Kim SW, Otwinowski Z, and Bezprozvanny I (2009). Secondary structure of Huntingtin amino-terminal region. Structure 17, 1205–1212. [PubMed: 19748341]

Kobe B, Gleichmann T, Horne J, Jennings IG, Scotney PD, and Teh T (1999). Turn up the HEAT. Structure 7, R91–R97. [PubMed: 10378263]

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, and Svergun DI (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr 36, 1277–1282.

Kotlo K, Xing Y, Lather S, Grillon JM, Johnson K, Skidgel RA, Solaro RJ, and Danziger RS (2014). PR65A phosphorylation regulates PP2A complex signaling. PLoS One 9, e85000. [PubMed: 24465463]

Lee JM, Ramos EM, Lee JH, Gillis T, Mysore JS, Hayden MR, Warby SC, Morrison P, Nance M, Ross CA, et al. (2012). CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology 78, 690–695. [PubMed: 22323755]

Leitner A, Reischl R, Walzthoeni T, Herzog F, Bohn S, Forster F, and Aebersold R (2012). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol. Cell. Proteomics 11, 10.1074/mcp.M111.014126.

Leitner A, Joachimiak LA, Unverdorben P, Walzthoeni T, Frydman J, Forster F, and Aebersold R (2014a). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. Proc. Natl. Acad. Sci. U S A 111, 9455–9460. [PubMed: 24938783]

Leitner A, Walzthoeni T, and Aebersold R (2014b). Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. Nat. Protoc 9, 120–137. [PubMed: 24356771]

Luo S, Vacher C, Davies JE, and Rubinsztein DC (2005). Cdk5 phosphorylation of huntingtin reduces its cleavage by caspases: implications for mutant huntingtin toxicity. J. Cell Biol 169, 647–656. [PubMed: 15911879]

Margolis RL, and Ross CA (2003). Diagnosis of Huntington disease. Clin. Chem 49, 1726–1732. [PubMed: 14500613]

McNeil SM, Novelletto A, Srinidhi J, Barnes G, Kornbluth I, Altherr MR, Wasmuth JJ, Gusella JF, MacDonald ME, and Myers RH (1997). Reduced penetrance of the Huntington's disease mutation. Hum. Mol. Genet 6, 775–779. [PubMed: 9158152]

Moritz A, Li Y, Guo A, Villen J, Wang Y, MacNeill J, Kornhauser J, Sprott K, Zhou J, Possemato A, et al. (2010). Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. Sci. Signal 3, 10.1126/scisignal.2000998.

Myers RH (2004). Huntington's disease genetics. NeuroRx 1, 255–262. [PubMed: 15717026]

Nucifora FC Jr., Sasaki M, Peters MF, Huang H, Cooper JK, Yamada M, Takahashi H, Tsuji S, Troncoso J, Dawson VL, et al. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science 291, 2423–2428. [PubMed: 11264541]

Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, et al. (2013). Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. J. Synchrotron Radiat 20, 660–664. [PubMed: 23765312]

Petoukhov MV, Konarev PV, Kikhney AG, and Svergun DI (2007). ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. J. Appl. Crystallogr 40, s223–s228.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612. [PubMed: 15264254]

Phanstiel DH, Brumbaugh J, Wenger CD, Tian S, Probasco MD, Bailey DJ, Swaney DL, Tervo MA, Bolin JM, Ruotti V, et al. (2011). Proteomic and phosphoproteomic comparison of human ES and iPS cells. Nat. Methods 8, 821–827. [PubMed: 21983960]

Rambo RP, and Tainer JA (2013). Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496, 477–481. [PubMed: 23619693]

Rangone H, Poizat G, Troncoso J, Ross CA, MacDonald ME, Saudou F, and Humbert S (2004). The serum- and glucocorticoid-induced kinase SGK inhibits mutant huntingtin-induced toxicity by phosphorylating serine 421 of huntingtin. Eur. J. Neurosci 19, 273–279. [PubMed: 14725621]

Ratovitski T, O'Meally RN, Jiang M, Chaerkady R, Chighladze E, Stewart JC, Wang X, Arbez N, Roby E, Alexandris A, et al. (2017). Post-translational modifications (PTMs), identified on endogenous huntingtin, cluster within proteolytic domains between HEAT repeats. J. Proteome Res 16, 2692–2708. [PubMed: 28653853]

Rinner O, Seebacher J, Walzthoeni T, Mueller LN, Beck M, Schmidt A, Mueller M, and Aebersold R (2008). Identification of cross-linked peptides from large sequence databases. Nat. Methods 5, 315–318. [PubMed: 18327264]

Rohou A, and Grigorieff N (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol 192, 216–221. [PubMed: 26278980]

de la Rosa-Trevin JM, Quintana A, Del Cano L, Zaldivar A, Foche I, Gutierrez J, Gomez-Blanco J, Burguet-Castell J, Cuenca-Alba J, Abrishami V, et al. (2016). Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. J. Struct. Biol 195, 93–99. [PubMed: 27108186]

Rubinsztein DC, Leggo J, Coles R, Almqvist E, Biancalana V, Cassiman JJ, Chotai K, Connarty M, Crauford D, Curtis A, et al. (1996). Phenotypic characterization of individuals with 30–40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats. Am. J. Hum. Genet 59, 16–22. [PubMed: 8659522]

Sapp E, Schwarz C, Chase K, Bhide PG, Young AB, Penney J, Vonsattel JP, Aronin N, and DiFiglia M (1997). Huntingtin localization in brains of normal and Huntington's disease patients. Ann. Neurol 42, 604–612. [PubMed: 9382472]

Scheres SHW (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol 180, 519–530. [PubMed: 23000701]

Schilling B, Gafni J, Torcassi C, Cong X, Row RH, LaFevre-Bernt MA, Cusack MP, Ratovitski T, Hirschhorn R, Ross CA, et al. (2006). Huntingtin phosphorylation sites mapped by mass spectrometry. Modulation of cleavage and toxicity. J. Biol. Chem 281, 23686–23697. [PubMed: 16782707]

Schneider CA, Rasband WS, and Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. [PubMed: 22930834]

Seong IS, Woda JM, Song JJ, Lloret A, Abeyrathne PD, Woo CJ, Gregory G, Lee JM, Wheeler VC, Walz T, et al. (2010). Huntingtin facilitates polycomb repressive complex 2. Hum. Mol. Genet 19, 573–583. [PubMed: 19933700]

Shevchenko A, Wilm M, Vorm O, and Mann M (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem 68, 850–858. [PubMed: 8779443]

Shi Y (2009). Serine/threonine phosphatases: mechanism through structure. Cell 139, 468–484. [PubMed: 19879837]

Shin A, Shin B, Shin JW, Kim KH, Atwal RS, Hope JM, Gillis T, Leszyk JD, Shaffer SA, Lee R, et al. (2017). Novel allele-specific quantification methods reveal no effects of adult onset CAG repeats on HTT mRNA and protein levels. Hum. Mol. Genet 26, 1258–1267. [PubMed: 28165127]

Shin B, Jung R, Oh H, Owens GE, Lee H, Kwak S, Lee R, Cotman SL, Lee JM, MacDonald ME, et al. (2018). Novel DNA aptamers that bind to mutant huntingtin and modify its activity. Mol. Ther. Nucleic Acids 11, 416–428. [PubMed: 29858077]

Tamo GE, Abriata LA, and Dal Peraro M (2015). The importance of dynamics in integrative modeling of supramolecular assemblies. Curr. Opin. Struct. Biol 31, 28–34. [PubMed: 25795087]

Tamo G, Maesani A, Trager S, Degiacomi MT, Floreano D, and Dal Peraro M (2017). Disentangling constraints using viability evolution principles in integrative modeling of macromolecular assemblies. Sci. Rep 7, 235. [PubMed: 28331186]

Vijayvargia R, Epand R, Leitner A, Jung TY, Shin B, Jung R, Lloret A, Singh Atwal R, Lee H, Lee JM, et al. (2016). Huntingtin's spherical solenoid structure enables polyglutamine tract-dependent modulation of its structure and function. eLife 5, e11184. [PubMed: 27003594]

Walker FO (2007). Huntington's disease. Lancet 369, 218–228. [PubMed: 17240289]

Walzthoeni T, Claassen M, Leitner A, Herzog F, Bohn S, Forster F, Beck M, and Aebersold R (2012). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat. Methods 9, 901–903. [PubMed: 22772729]

Warby SC, Chan EY, Metzler M, Gan L, Singaraja RR, Crocker SF, Robertson HA, and Hayden MR (2005). Huntingtin phosphorylation on serine 421 is significantly reduced in the striatum and by polyglutamine expansion in vivo. Hum. Mol. Genet 14, 1569–1577. [PubMed: 15843398]

Wriggers W, Milligan RA, and McCammon JA (1999). Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. J. Struct. Biol 125, 185–195. [PubMed: 10222274]

Yoshimura SH, and Hirano T (2016). HEAT repeats—versatile arrays of amphiphilic helices working in crowded environments? J. Cell. Sci 129, 3963–3970. [PubMed: 27802131]

Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, and Agard DA (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332. [PubMed: 28250466]

Highlights

- The cryo-EM analysis on normal (Q23) and disease (Q78) HTT are described
- The polyQ expansion alters the relative movement of HTT domains
- The polyQ expansion at N terminus alters the global phosphorylation pattern of HTT
- Ser2116 phosphorylation affects the global structure and function of HTT

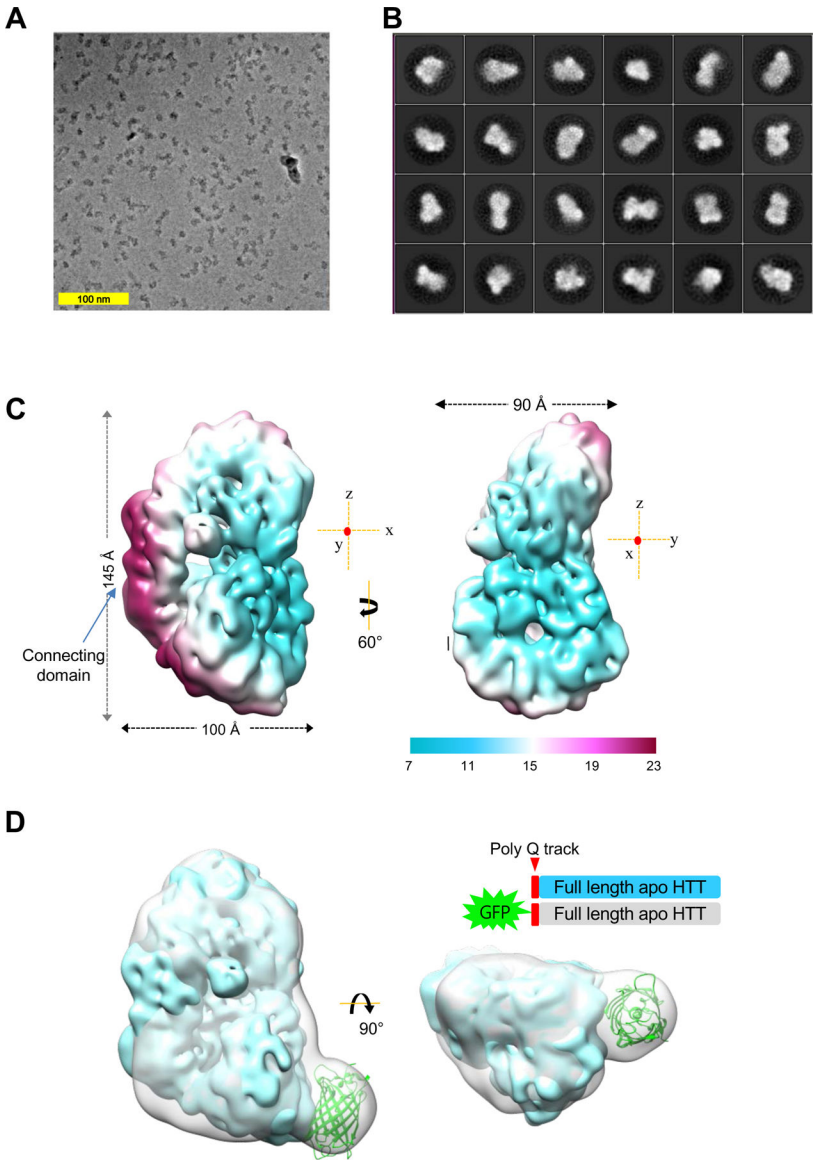


Figure 1. Cryo-EM Analysis of Full-Length Q23-HTT in Apo State

(A) A representative micrograph of Q23-HTT at a defocus of 0.7 μm recorded using a Volta phase plate.

(B) 2D class averages of Q23-HTT particles used for 3D reconstruction.

(C) Cryo-EM map of Q23-HTT with color coding by locally estimated resolution in two different view angles. The dimensions of the Q23-HTT are 100 × 90 × 145 Å along the x, y, and z axes, respectively. A three-axis plot indicates the orientation (red dots indicate an axis perpendicular to the plane).

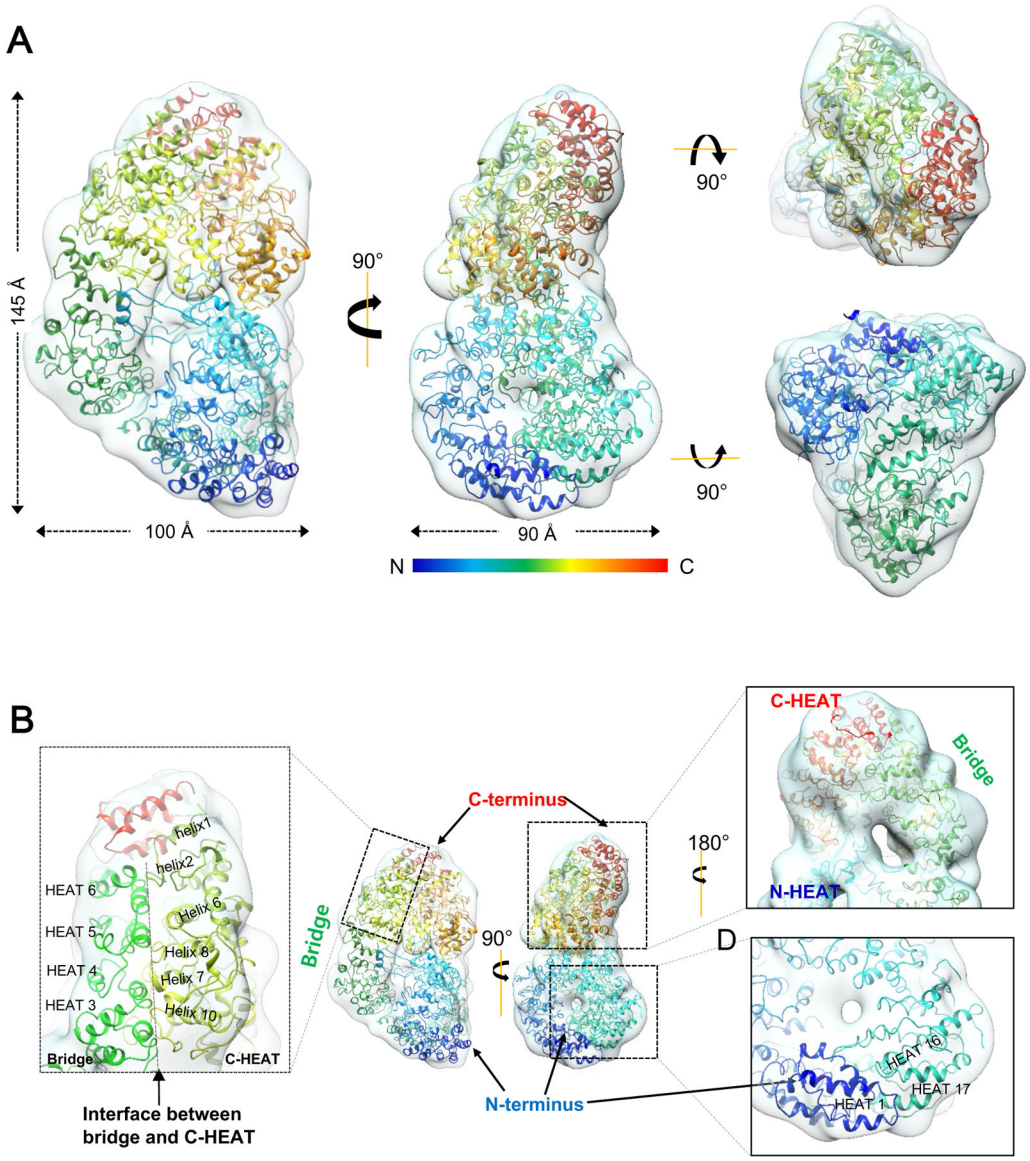


Figure 2. Modeling of the Atomic Structure of Q23-HTT Based on the HTT-HAP40 Complex

(A) Q23-HTT model superimposed on the cryo-EM map of Q23-HTT. The Q23-HTT model is shown rainbow-colored from N to C terminus in a cartoon model.

(B) Detailed view of the Q23-HTT model fitted on the cryo-EM map. The N- and C-HEAT domains are colored in rainbow, and the Bridge domain in green (see Figures S3A and S3B for additional details).

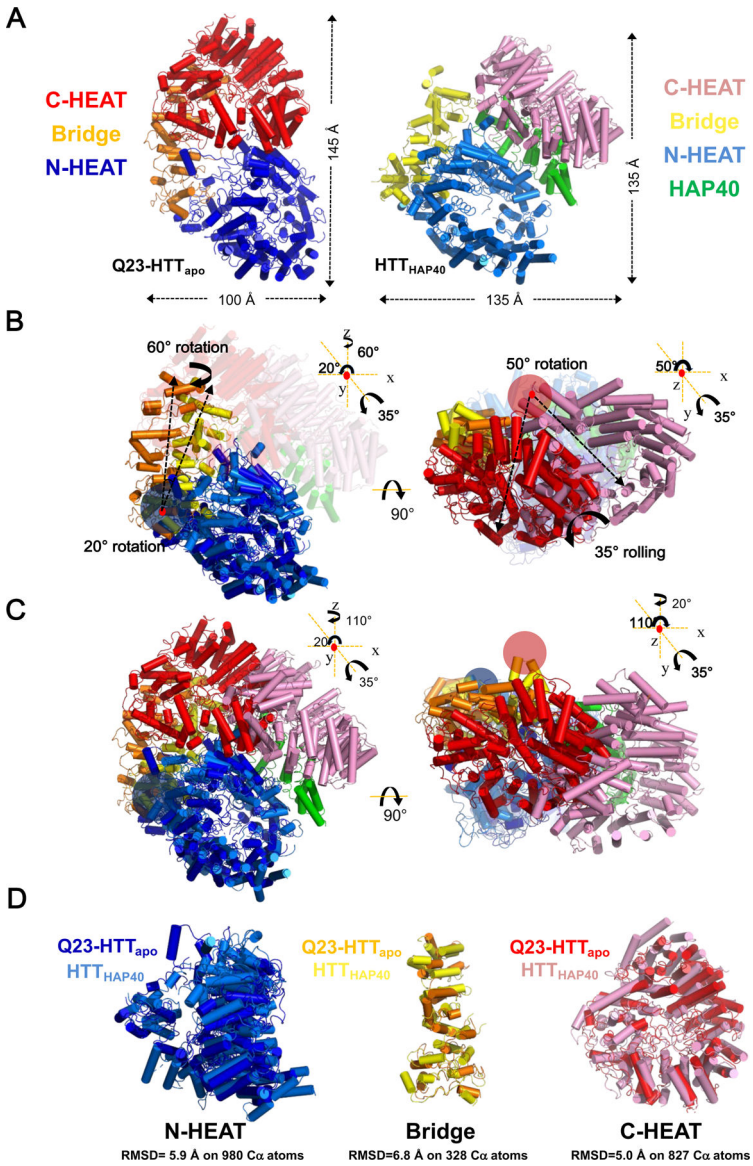


Figure 3. Conformational Change of the C-HEAT Domain in Q23-HTT Compared with HTT_{HAP40}

(A) Overall structures of Q23-HTT and HTT_{HAP40}. N-HEAT, C-HEAT, and Bridge domains are colored in blue/royal blue, orange/yellow, and red/pink, respectively.

(B) N-HEAT domains of Q23-HTT and HTT_{HAP40} were superimposed, revealing the difference in the relative positions of the Bridge domains (left panel). The Bridge domain of Q23-HTT is rotated about 20° at the y axis (red dot, perpendicular to the plane), and rotated 60° along the z axis, compared with HTT_{HAP40} (left panel). The Bridge domains were superimposed, revealing the difference in the relative positions of the C-HEAT domains (right panel). The C-HEAT domain of Q23-HTT is rotated about 50° at the z axis and rolled over about 35°. A three-axis plot with angles shows the movement of the domains. Dark-blue and pale-red circles indicate pivot joints.

(C) Overall conformational change of Q23-HTT and HTT_{HAP40} structure. A three-axis plot with angles shows the movement of the domains at the pivot joints marked with dark-blue and pale-red circles.

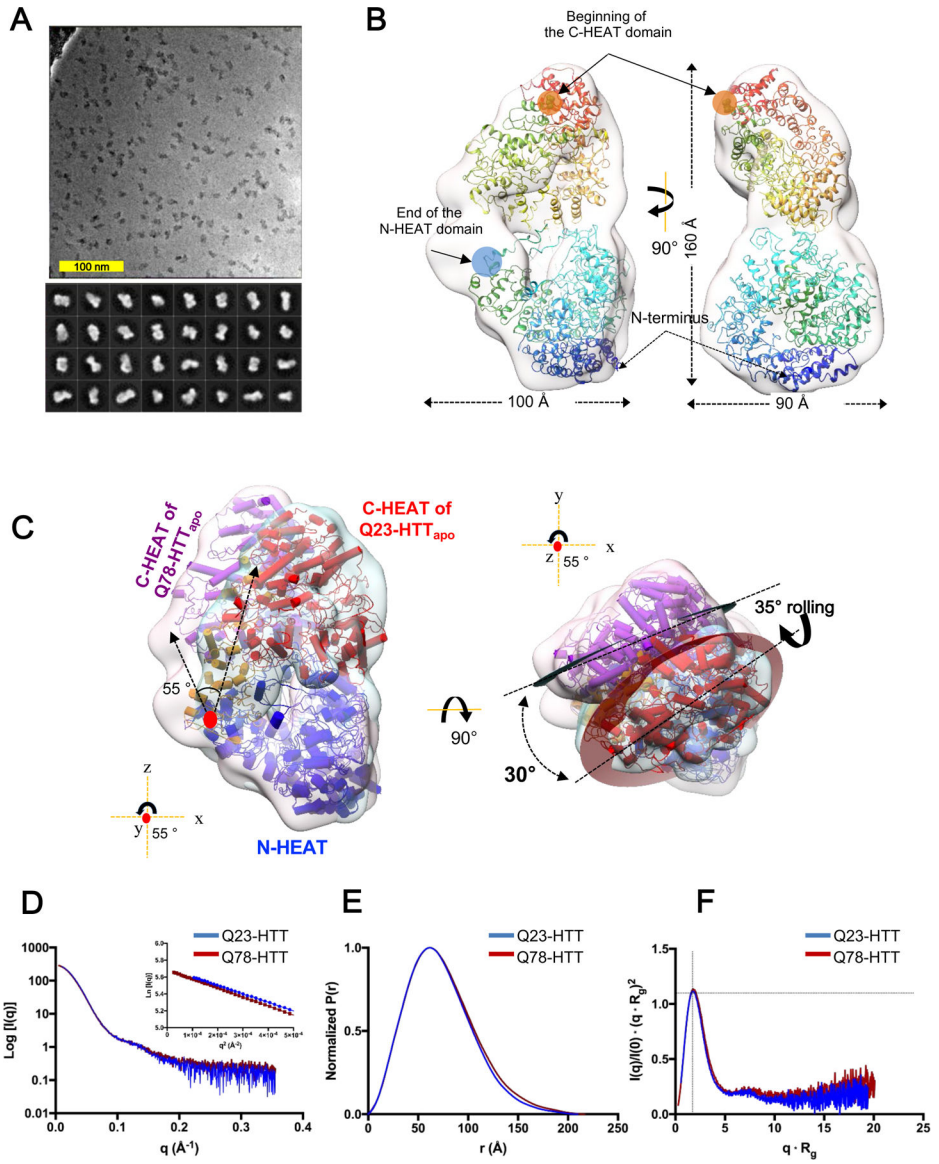


Figure 4. Cryo-EM Analysis of Full-Length Q78-HTT in Apo State and Comparison with Q23-HTT in Apo State

(A) A representative micrograph of Q78-HTT at a defocus of 0.5 mm recorded using a Volta phase plate (top). 2D class averages of selected Q78-HTT particles used for single-particle reconstruction (bottom).

(B) Cryo-EM map of Q78-HTT fitted with N- and C-HEAT domains of a rigid body fitted model. The model is depicted in rainbow colors and fit in the cryo-EM map. The cryo-EM map of Q78-HTT shows overall dimensions as 100 × 90 × 160 Å. The end of the N-HEAT domain and the pivot joints at the beginning of N-HEAT and the end of C-HEAT are indicated by blue and red circles, respectively (see Figures S3A and S3C for additional details).

(C) A superimposition of cryo-EM maps and models of Q23-HTT and Q78-HTT. N-HEAT domains of Q23-HTT and Q78-HTT are aligned, revealing structural difference between the two structures. N-HEAT is shown in blue, Bridge in orange, and C-HEAT domains are in red

for Q23-HTT and magenta for Q78-HTT. Red dots show a y axis (left panel) or z axis (right panel), which are perpendicular to the surface (see Figure S3D for additional details).

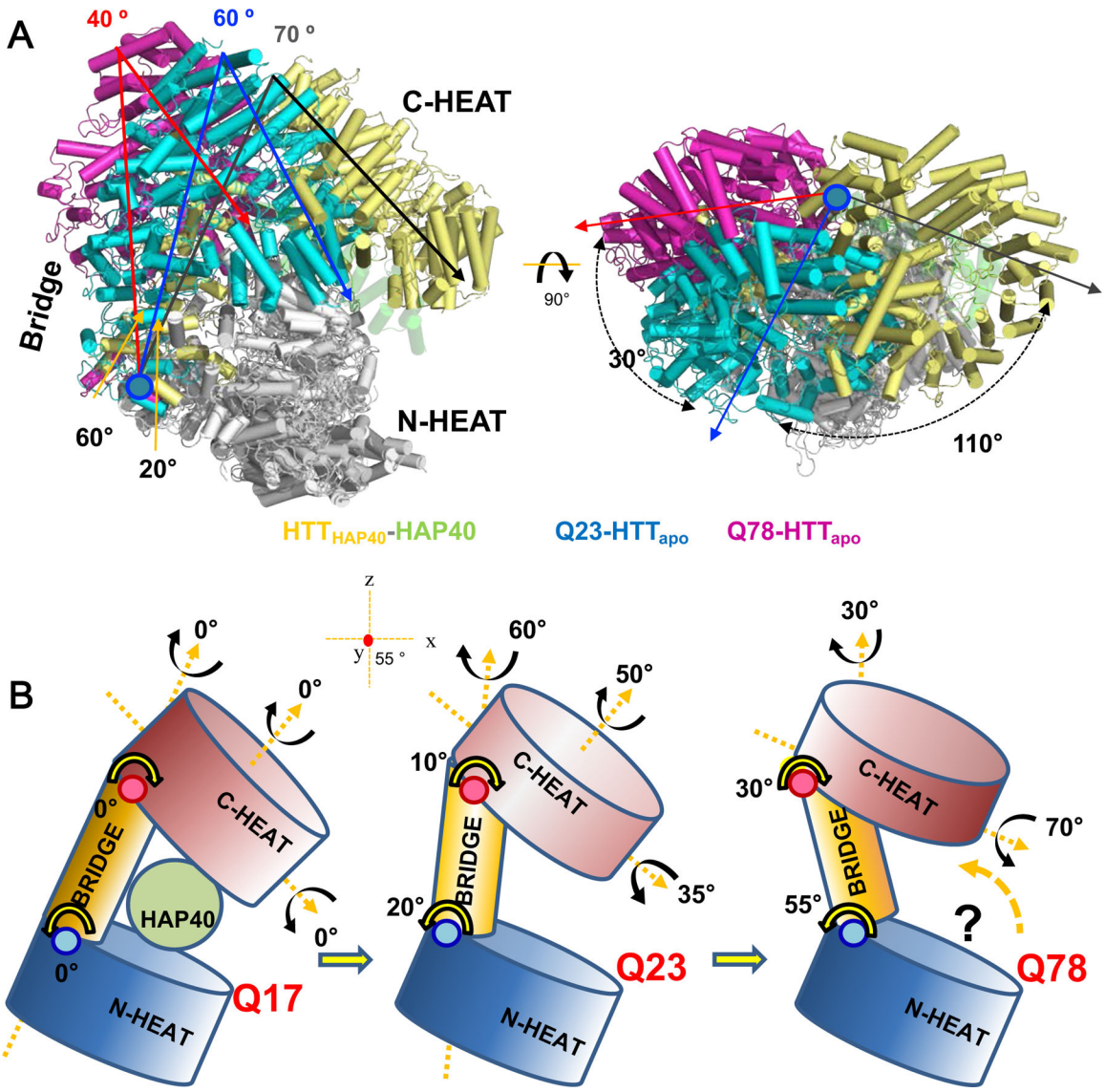


Figure 5. Conformational Flexibility of the C-HEAT Domain of HTT

(A) Superimposition of Q23-HTT, Q78-HTT, and HTT$_{HAP40}$ models aligning N-HEAT domains. Left panel: the Bridge and C-HEAT domains colored in cyan for Q23-HTT, pink for Q78-HTT, and pale yellow for HTT$_{HAP40}$. For clarity, the N-HEAT domains are colored white. The rotation angles at the pivot joint (pale-blue circle) located at the beginning of Bridge domain are indicated below, and the angles between the Bridge and C-HEAT are shown above. Right panel: 90° rotated view from the left panel. The rotation angles of C-HEAT at the pivot joint (pale-blue circle) are indicated.

(B) A schematic model for conformational flexibility of HTT. The angles in HTT$_{HAP40}$ are set 0° to compare relative positions in Q23- and Q78-HTT. N-HEAT domains are shown in blue, Bridge in yellow, and C-HEAT in red. Pale-blue and pal- red circles indicate two pivot joints.

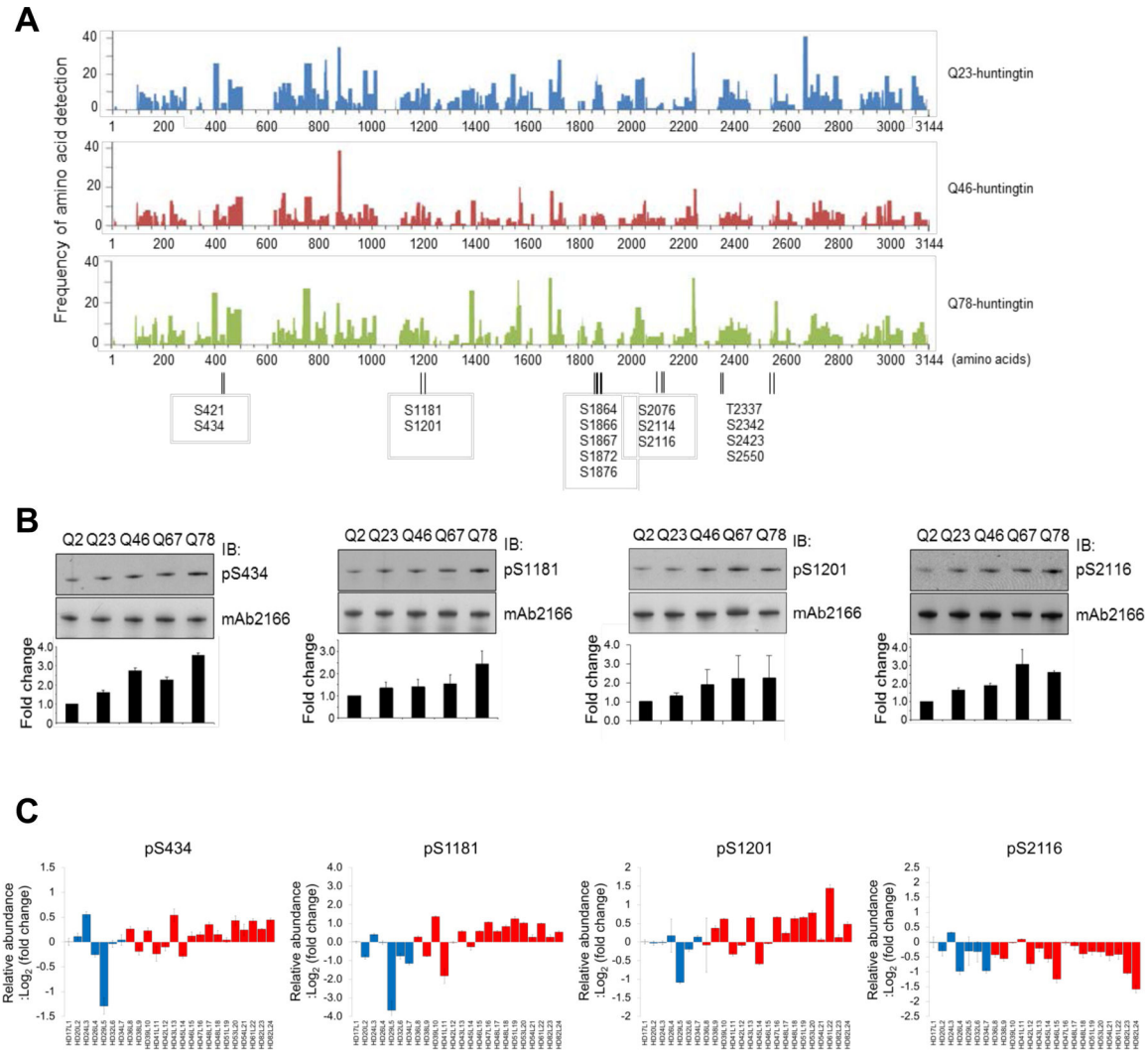


Figure 6. Polyglutamine Length-Dependent Alteration of Phosphorylation in Purified HTT Protein and Human Cells

(A) Coverage graphs of full-length HTT proteins subjected to LC-MS/MS analysis for determination of phosphorylation sites. Both normal (Q23-HTT) and mutant (Q46- and Q78-HTT) proteins show similar coverage patterns and to the same extent (~80%). Several HTT phosphorylation sites were identified by MS and are listed along the length of the HTT amino acid sequence.

(B) The purified recombinant Q2-, Q23-, Q46-, Q67-, and Q78-HTT series were subjected to immunoblot (IB) analysis and probed with a panel of 15 phosphorylation site-specific antibodies. The results of four phosphorylation sites (pS434, pS1181, pS1201, and pS2116) are shown here along with the quantification graph. The signal was normalized to total HTT levels (the signal of mAb2166). Data are averages of three experiments and error bars are SEM.

(C) A panel of 24 lymphoblastoid cell lines (7 controls, blue bar; and 17 HD subjects, red bar) was subjected to PRM-based targeted MS experiments to ascertain the profile of HTT phosphorylation in human cells. The three targeted ions derived from phosphopeptide

containing pS434, pS1181, pS1201, and pS2116 were extracted and summed to represent the relative abundance of phosphopeptides (y axis, $\log_2$ scale). Data represent mean of three experiments (technical replicates).

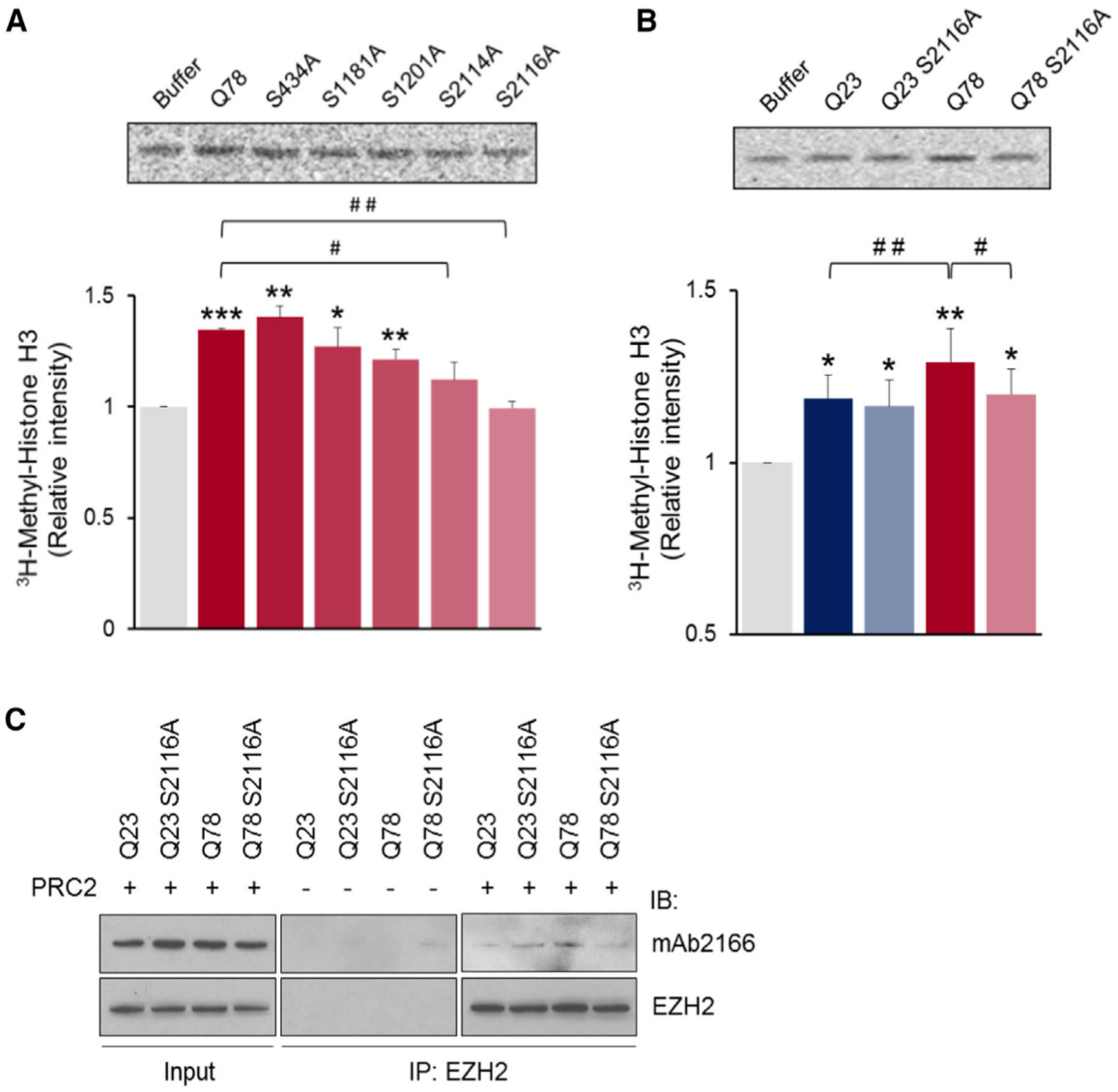


Figure 7. Phosphorylation Status Modulates HTT Activity in a Polyglutamine Length-Dependent Manner: The Critical Role of pS2116 on Mutant HTT Activity

(A) Cell-free PRC2 assay of five alanine mutants of Q78-HTT at S434, S1181, S1201, S2114, and S2116 showed that the enhanced activity of Q78-HTT was significantly decreased in both S2114A and S2116A mutants compared with wild-type Q78-HTT.

(B) Cell-free PRC2 assay of two S2116A mutants of either Q23-or Q78-HTT revealed that the S2116A mutation affects only the extra activity of mutant HTT without altering baseline activity. For both cell-free PRC2 assay (A and B), data are averages of three experiments and error bars are SEM. The significance was calculated and displayed compared with PRC2 alone (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$), or as indicated ($^{\#}p < 0.05$, $^{\#\#}p < 0.01$) using Student's t test.

(C) Effect of HTT S2116 phosphorylation status on its interaction with PRC2 was analyzed by immunoprecipitation with anti-EZH2 antibody followed by immunoblotting for HTT. The same mutation, S2116A, oppositely affected the interaction between HTT and PRC2

depending on the polyQ tract size, such as dramatically decreased in Q78-HTT while mildly increased in Q23-HTT. The immunoblotting (IB) for EZH2 showed all four different HTTs co-eluted with the same amount of PRC2. There was a weak non-specific binding of Q78 S2116A-HTT in control immunoprecipitation (IP) (fourth lane in central panel), implying even more decreased interaction between PRC2 and Q78 S2116A-HTT compared to Q78-HTT.

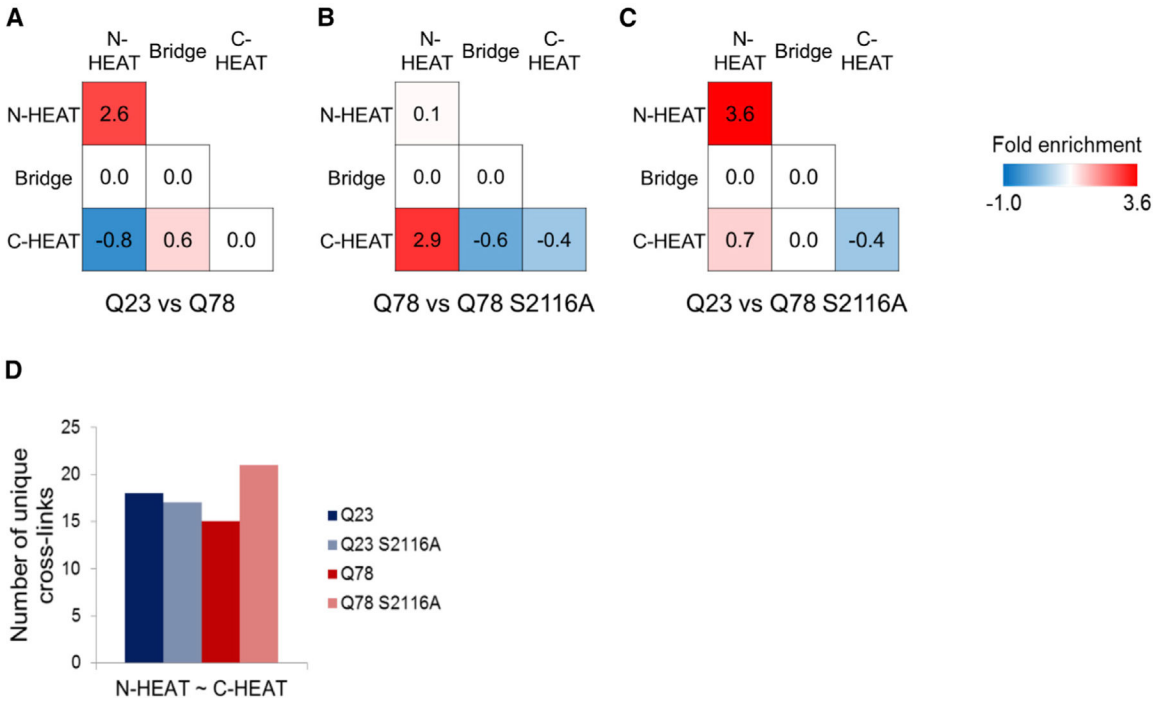


Figure 8. pS2116 Affects the Global Structure of HTT

(A–C) Q23- and Q78-HTT including corresponding S2116 alanine mutants were crosslinked by DSS and intra-domain crosslinked peptides were identified by MS. The fold enrichment score between Q23- and Q78-HTT (A), Q78- and Q78 S2116A mutant HTT (B), and Q23- and Q78 S2116A mutant HTT (C) were assigned based on the equation described in STAR Methods and depicted as a heatmap.

(D) Bar graph showing the number of unique crosslinks between N-HEAT and C-HEAT domains from Q23- and Q78-HTT including corresponding S2116 alanine mutants.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse IgG, HRP-linked	Amersham	Cat#NA931; RRID: AB_772210
Anti-rabbit IgG, HRP-linked	Amersham	Cat#NA934; RRID: AB_772206
Mouse monoclonal anti-EZH2	BD Transduction Laboratories	Cat#612667; RRID: AB_399911
Mouse monoclonal anti-HTT	EMD Millipore	Cat#MAB2166; RRID: AB_2123255
Rabbit polyclonal anti-Phospho HTT (S1181)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1201)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1864)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1866)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1867)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1872)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S1876)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2076)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2114)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2116)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2337)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2342)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2423)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S2550)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S421)	This paper	N/A
Rabbit polyclonal anti-Phospho HTT (S434)	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
Adenosyl-L-methionine, S-[methyl-^{3}H] (SAM[^{3}H])	PerkinElmer	Cat#NET155V250UC
Amicon Ultra 100-kDa filter	Millipore	ACS510024
Anti-FLAG M2 affinity gel	Sigma Aldrich	A2220-25ML
Cellfectin™ II Reagent	Thermo Scientific	Cat#10362100
DSS-H12/D12	Creative Molecules	Cat#001S
Disuccinimidyl suberate (DSS)	Thermo Scientific	Cat#21555
DL-Dithiothreitol (DTT)	Soltec Ventures	M112_100g
Ethylenediaminetetraacetic acid (EDTA)	Sigma Aldrich	E9884-1KG
FLAG-peptide	Apeptide	APG1474
Glycerol	LPS solution	GLYC01
Graphene oxide	Sigma Aldrich	763705-100ML
HEPES	Sigma Aldrich	H3375-550G
Hydrochloric Acid (HCl)	SAMCHUN	H0255
Imidazole	Bio basic	IB0277
Nickel-nitrilotriacetic acid (Ni-NTA) resin	Qiagen	Cat#30230
Octyl β-D-glucopyranoside	Sigma Aldrich	O8001-1G

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Phosphopeptides used in this work are listed in Figure S6	This paper	N/A
Polycomb repressive complex 2 (PRC2)	This paper	N/A
Q23 S2116A-HTT	This paper	N/A
Q23-HTT	This paper	N/A
Q2-HTT	This paper	N/A
Q46-HTT	This paper	N/A
Q67-HTT	This paper	N/A
Q78 S1181A-HTT	This paper	N/A
Q78 S1201A-HTT	This paper	N/A
Q78 S2114A-HTT	This paper	N/A
Q78 S2116A-HTT	This paper	N/A
Q78 S434A-HTT	This paper	N/A
Q78-HTT	This paper	N/A
Sodium Chloride (NaCl)	LPS solution	NACL05
Superdex 200 26/60	GE healthcare	Cat#28-9893-36
Tris	Amresco	0497-5kg
Adenosyl-L-methionine, S-[methyl-^{3}H] (SAM[^{3}H])	PerkinElmer	Cat#NET155V250UC
Deposited Data		
HTT-HAP complex	Guo et al., 2018	PDB: 6EZ8
GFP-Q23-HTT, EM map	This paper	EMDB: EMD-4944
Q23-HTT, coordinates	This paper	PDB: 6RMH
Q23-HTT, EM map	This paper	EMDB: EMD-4937
Q78-HTT, coordinates	This paper	PDB: 6YEJ
Q78-HTT, EM map	This paper	EMDB: EMD-10793
XL-MS data	This paper	ProteomeXchange/PRIDE: PXD013907
Experimental Models: Cell Lines		
Escherichia Coli DH5α	Enzynomics	CPO11
Sf9 cells in Sf-900™ II SFM	Thermo Scientific	Cat#11496015
Software and Algorithms		
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/; RRID:SCR_003070
IBM SPSS Statistics was used for the regression analysis	IBM Corporation	https://www.ibm.com/us-en/marketplace/spss-statistics; RRID:SCR_002865
cisTEM 1.0	Grant et al., 2018	https://cistem.org/; RRID:SCR_016502
CTFFind4	Rohou and Grigorieff, 2015	http://grigoriefflab.janelia.org/ctffind4 ; RRID: SCR_016732
EPU	FEI	https://www.fei.com/software/epu-automatedsingle-particles-software-for-life-sciences/
HDExaminer	Sierra Analytics	http://massspec.com/hdexaminer/

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MotionCor2	Zheng et al., 2017	http://msg.ucsf.edu/em/software/motioncor2.html ; RRID:SCR_016499
PHENIX	Adams et al., 2010	https://www.phenix-online.org/ ; RRID: SCR_014224
pow^{er}	Degiacomi and Dal Peraro, 2013	https://www.epfl.ch/labs/lbm/resources/
PyMOL	Schrödinger, LLC.	https://pymol.org/2/ ; RRID: SCR_000305
RELION 2.0	Scheres, 2012	https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
SITUS	Wriggers et al., 1999	http://situs.biomachina.org/fdownmain.html
UCSF Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/ ; RRID: SCR_004097
xQuest	Rinner et al., 2008; Walzthoeni et al., 2012	
Other		
Quantifoil R2/1	ELECTRON MICROSCOPY SCIENCES	Q250CR1, Q350CR1