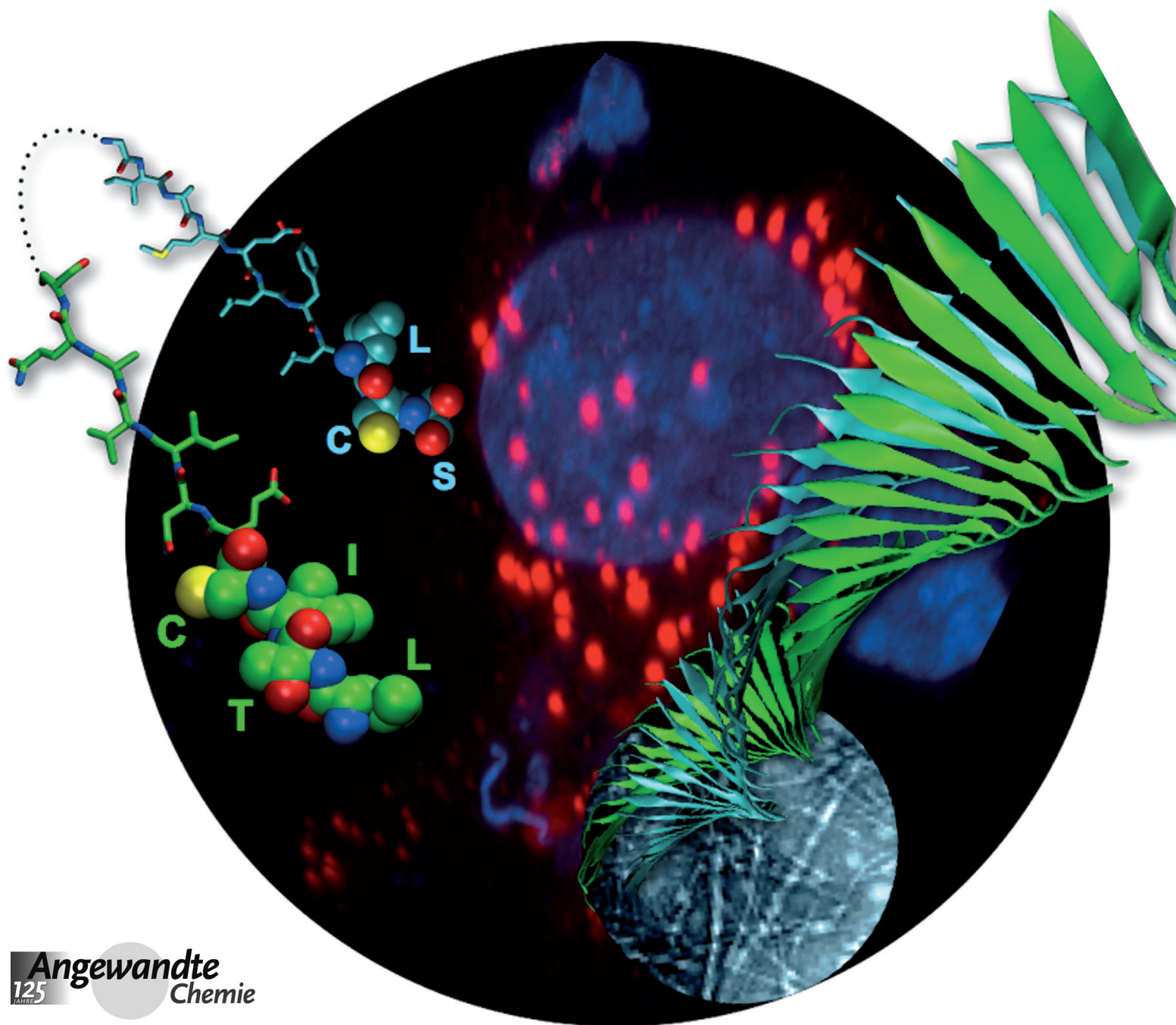


Discovery of a Novel Aggregation Domain in the Huntingtin Protein: Implications for the Mechanisms of Htt Aggregation and Toxicity**

Zhe-Ming Wang and Hilal A. Lashuel*



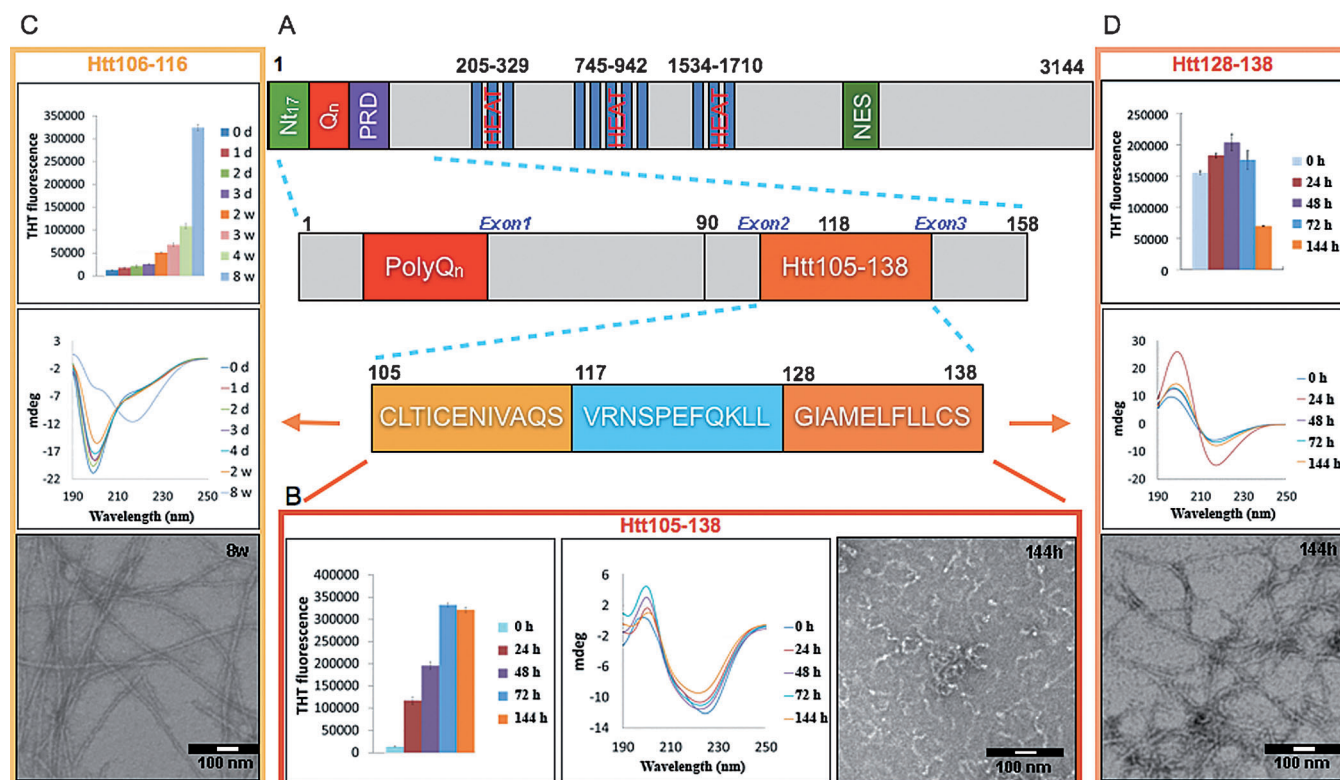


Figure 1. A) Schematic depiction of the Htt105–138 domain and the three peptides generated to determine the sequence determinants of its aggregation, B,C,D) Time course aggregation of each peptide using ThT binding assay, CD, and TEM: Htt105–138 and Htt128–138 were prepared as 40 μM solutions in 10 mM Tris pH 7.4 buffer, whereas Htt106–116 was prepared as 200 μM solutions in 10 mM Tris pH 7.4 buffer. All samples were incubated at 37°C; the scale bar in the TEM images represents 100 nm.

Huntington's Disease (HD) is caused by CAG trinucleotide expansion, which translates to polyQ repeats, within exon1 of the Htt gene, which encodes for the Huntingtin (Htt), a 350 kDa protein (Figure 1 A) that is highly expressed in the brain and different organs.^[1] In HD-affected people, the length of the polyQ repeat varies from 36–120 residues, with a length of 42–50 residues being associated with adult onset, and a length of over 60 residues associated with juvenile onset.^[2] Although the mechanisms by which the expanded polyQ contribute to Htt-induced toxicity remain unknown, the formation of nuclear inclusions of Htt in the brain of HD patients is an invariant feature and thought to contribute to disease pathogenesis by a gain-of-toxic-function and partial loss of normal function linked to the polyQ mediated Htt aggregation.^[3] The toxic fragment hypothesis suggests that

proteolysis plays a crucial role in the pathogenesis of HD and proposes that N-terminal Htt fragments (NtHtts) containing the polyQ domain may constitute the primary toxic form of Htt by virtue of the ability to aggregate rapidly, nucleate, and accelerate the aggregation process.^[4] The inverse correlation between the length of the polyQ and the age at onset and severity^[5] initially led to the suggestion the polyQ is the primary determinant of Htt aggregation and toxicity. Moreover, several studies have demonstrated a direct relationship between the length and the aggregation propensity of polyQ repeats in peptides,^[6] NtHtts,^[7] and full-length Htt (Fl-Htt) in vitro, in cultured cells and animal models.^[8] This hypothesis was further supported by the identification of other polyQ-containing proteins as key players in other neurodegenerative diseases (e.g. SBMA and the SCAs).

However, several lines of evidence suggest that Htt oligomerization, inclusion formation, and toxicity are strongly influenced by sequences outside the polyQ domain. This hypothesis is supported by several findings that indicate extensive cross-talk and interactions between the different domains in Htt: 1) Mutations, deletions, or post-translational modifications (PTMs) within the flanking sequences of the polyQ repeats region in Httex1 (N-terminal 17 amino acids [Nt17],^[7,10] polyP,^[11] and (P/Q) repeat domains) strongly influence the aggregation, cellular behavior, and toxicity of mutant Htt.^[9] 2) NtHtt fragments of different lengths (e.g., 67, 117, 171, and 586) containing similar numbers of polyQ repeats exhibit different aggregation and toxicity properties

[*] Dr. Z. M. Wang, Prof. Dr. H. A. Lashuel
Laboratory of Molecular and
Chemical Biology of Neurodegeneration, Brain Mind Institute
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne (Switzerland)
E-mail: hilal.lashuel@epfl.ch

[**] We would like to express our thanks and deep appreciation to Dr. Ruth Luthi-Carter for insightful discussions and support. We thank Annalisa Ansaloni for insightful discussions and John Perrin for assistance with TEM measurements.

Supporting information (experimental conditions and synthetic methods) for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201206561>.

in cellular and animal models of HD.^[4d,12] 3) The longer NtHtt fragments showed enhanced aggregation (e.g. Htt1–117) and/or toxicity (Htt1–208^[13]) compared with the shorter exon1 fragment. Altogether, these findings indicate that the sequence context is an important determinant of polyQ-mediated Htt aggregation and toxicity.

While working on the chemical synthesis of the NtHtt fragment comprising residues 1–140 (Figure 1A), we observed that the residues 105–138 exhibited poor solubility and a high propensity to form amyloid-like fibrils *in vitro*. This direct observation combined with the fact that this peptide contains sequences shared by several well-characterized NtHtt fragments prompted us to further investigate the structural and biophysical properties of this peptide. To identify the sequences that are responsible for its enhanced aggregation properties, we generated a series of synthetic peptide fragments spanning different region of Htt105–138 and assessed their structural and aggregation properties using circular dichroism (CD), transmission electron microscopy (TEM), and the Thioflavin T (ThT) fluorescence assay. Our studies led to the discovery and identification of novel sequences within exons2 and 3 of Htt that exhibit a high propensity to aggregate and form amyloid-like fibrils.

To solubilize the peptides, we followed the procedure developed by Chen and Wetzel.^[14] The Htt105–138 was subjected to this method and then prepared as a 40 μ M solution in 10 mM Tris pH 7.4 buffer. The far-UV circular dichroism (CD) spectra obtained immediately after dissolving the peptide in the Tris buffer showed a negative minimum centered at approximately 220 nm and a positive maximum at approximately 200 nm (Figure 1B). These results demonstrate that Htt105–138 readily self-assembles to form a β -sheet structure. TEM analysis of the same sample (Figure 1B) revealed predominately short curvilinear structures with an average length and diameter of 200 nm and 10 nm, respectively. Interestingly, these structures do not convert with time into the typical long filamentous structures that are characteristic of amyloid fibrils formed by other amyloidogenic proteins *in vitro* or isolated post-mortem from human brains. To determine if the β -sheet aggregates formed by Htt105–138 have amyloid-like characteristics, we monitored and quantified its aggregation using the amyloid-specific dye ThT. Figure 1B shows a time-dependent increase in ThT fluorescence, thereby demonstrating that the Htt105–138 aggregates bind to ThT.

To better understand the molecular and sequence determinants that govern the aggregation of the Htt105–138, we synthesized three peptides corresponding to the N-terminal (106–116), middle (117–127), and C-terminal (128–138) regions of Htt105–138 and assessed the propensity of each peptide to form a β -sheet structure and to aggregate by using CD, TEM, and ThT. Unlike Htt106–116 and Htt128–138, Htt117–127 is readily soluble in buffers, and its CD spectrum indicates a random coil structure. Under all incubation conditions tested, Htt117–127 remained unstructured and did not show any tendency to aggregate (Figure S1A Supporting Information).

To examine the aggregation propensity of Htt106–116, the peptide was prepared as a 200 μ M solution in 10 mM Tris

pH 7.4 buffer after disaggregation treatment. Under these conditions, the CD spectra showed a negative minimum centered at approximately 195 nm and a second minimum at 218 nm consistent with a mixed structure of random coil and β -sheet structures (Figure 1C). Upon further incubation at 37°C, we did not observe any change in the CD spectra or ThT signal for up to two weeks. However, after two weeks, we observed a gradual increase in the ThT signal and a shift in the CD spectra from a predominantly random coil to β -sheet structure as evidenced by the negative minimum centered at approximately 218 nm. The TEM image shown in Figure 1C demonstrates that this peptide forms long fibrillar assemblies with an average length and diameter of 2 μ m and 11 nm, respectively. Interestingly, at pH values under 5, especially at pH 3.4 and 4.4, the Htt106–116 peptide (pI = 4.0) exhibited a higher tendency to form a β -sheet-rich structure and aggregates. Under these conditions, the CD spectra showed a mixture of β -sheet and random coil structures with the latter being the dominant structure. However, after 24 h, the peptide exhibited double minima of equal intensity suggesting an equal mixture of β -sheet and random coil structures (Figure S1C). In contrast to Htt106–116 and Htt117–127, the Htt128–138 peptide exhibited a very high propensity to form β -sheet-rich fibrillar aggregates (Figure 1D). At 144 h, we observed fibrils with an average length and diameter of 400 nm and 10 nm, respectively. The CD spectrum at time 0 is typical for a β -sheet structure with a minimum at 218 nm and a maximum at 195–200 nm. Upon further incubation at 37°C for a period of up to 6 days, the CD spectra remained unchanged. The rapid aggregation of Htt128–138 was evident by the high ThT signal at time 0, which remained unchanged up to 72 h and by the detection of extensive fibrils in all samples. These results suggest that the majority of the Htt128–138 were already assembled into fibrils upon dissolution in aqueous buffer, Figure S1D.

Having identified the main regions responsible for driving the aggregation of Htt105–138, we sought to further characterize these regions and determine the key residues that mediate the aggregation of each region. Towards this goal, we synthesized a series of short peptides corresponding to different fragments of the two amyloid-forming sequences, Htt106–116 and Htt128–138, and assessed the structure and the propensity for amyloid formation of each peptide by CD, ThT, and TEM (Figure 2). When the aggregation was assessed at pH 7.4, only Htt105–138 and Htt128–138 showed an increase in the ThT signal with time and formed amyloid fibrils (Figure 2B). Interestingly, even at a higher concentration (200 μ M), the N-terminal-derived peptides did not aggregate.

Consistent with the results in Figure 1C, Htt106–116 showed only a slight increase in ThT signal, and rare fibrils were detected by TEM. However, similar to Htt106–116, Htt108–116 (pI = 4.0) also demonstrated a pH-dependent aggregation behavior (Figure S1E) and exhibited a higher aggregation propensity at pH values under 5, Htt108–116 exhibited a CD spectrum that is consistent with the presence of a predominantly β -sheet-rich structure (Figure 2C). Taken together, these results suggest that the first four residues, LTIC, are important for the aggregation of the Htt106–116

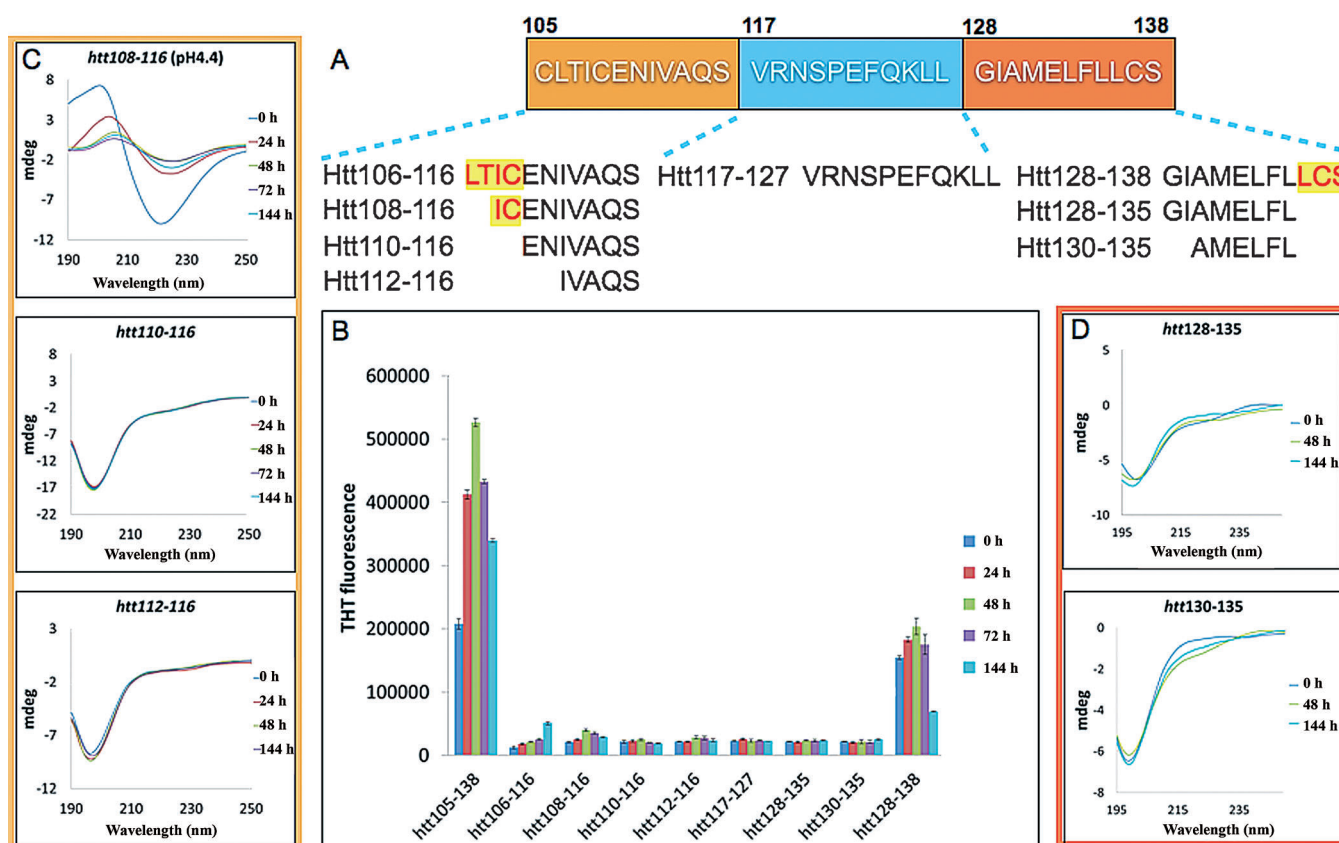


Figure 2. Identification of key amino acids that are required for the aggregation of the two amyloidogenic regions: A) Schematic depiction of the various peptides that were made and characterized. B) Comparison of the ThT binding time course of the 9 peptides: samples were prepared as 40 μ M solutions in 10 mM Tris pH 7.4 buffer except for Htt106–116 series (200 μ M). C, D) CD spectra of the Htt106–116 and Htt128–138 series: samples were prepared as 200 μ M (C) and 40 μ M (D) solutions in 10 mM Tris pH 7.4 buffer or 10 mM NaAc pH 4.4, and incubated at 37 °C.

peptide. Similarly, the removal of the last three C-terminal residues, LCS (Htt128–135), resulted in the significant inhibition of β -sheet formation (Figure 2D) and abolished the aggregation of Htt128–138, suggesting that these three residues are essential for the aggregation of this peptide. The aggregation of both peptides (Htt106–116 and Htt128–138) was not influenced by the presence of reducing agent, thus ruling out the possibility of disulfide-mediated assembly.

To further validate our findings and determine the critical role of the N-terminal region (Htt106–116) in the aggregation of the longer peptide Htt105–138, the following peptides were synthesized: Htt110–138 and Htt117–138. Figure 3C illustrates that upon dissolution, both peptides possess a transient helical structure but undergo rapid transformation to a β -sheet-rich structure. Interestingly, the removal of the Htt106–116 but not the first four amino acids (106–110) resulted in a significant reduction of amyloid formation (Figure 3B). Both peptides formed classical amyloid-like fibrillar structures (Figure 3D) but differed in their average length and diameter: 2 μ m and 12 nm for Htt110–138 and 830 nm and 15 nm for Htt117–138. Similarly, to determine which C-terminal residues are critical for aggregation, we synthesized two peptides in which either the last 3 or 11 amino acids were removed: Htt106–135 and Htt106–127, respectively. The removal of the C-terminal residues 128–138 resulted in enhanced solubility and a significant reduction in the β -

sheet structure and amyloid formation (Figure 3C). Upon dissolving the peptide in aqueous buffer, its CD spectrum showed a minima at 200 nm and 222 nm indicating the presence of a mixed random coil and β -sheet structure. Consistent with this interpretation was the presence of short fibrillar structures with an average length and diameter of 400 nm and 12 nm, respectively. The corresponding ThT fluorescence signal is consistent with a lower propensity to form amyloid structure compared to Htt106–135 (Figure 3B). Although the removal of the last three C-terminal amino acids (LCS) was sufficient to abolish the aggregation of the short peptide Htt128–138, in the context of the full-length peptide, it reduced but did not block β -sheet and fibril formation by Htt106–135. The fibrillar structure of Htt106–135 has an average length and diameter of 700 nm and 13 nm, respectively. These findings combined with the increased aggregation properties of Htt128–138 compared with Htt106–116 (Figure 1) or Htt108–116 (Figure 2) suggest that both of these regions contribute to the aggregation of Htt105–138 and NtHtt, with the Htt128–138 being more amyloidogenic. To what extent these two amyloidogenic regions contribute to the aggregation and toxicity of the full-length protein remains unknown. However, given that either or both of these regions could become highly exposed upon proteolysis of Htt, they could potentially contribute significantly to enhancing the aggregation properties of the NtHtt fragments.

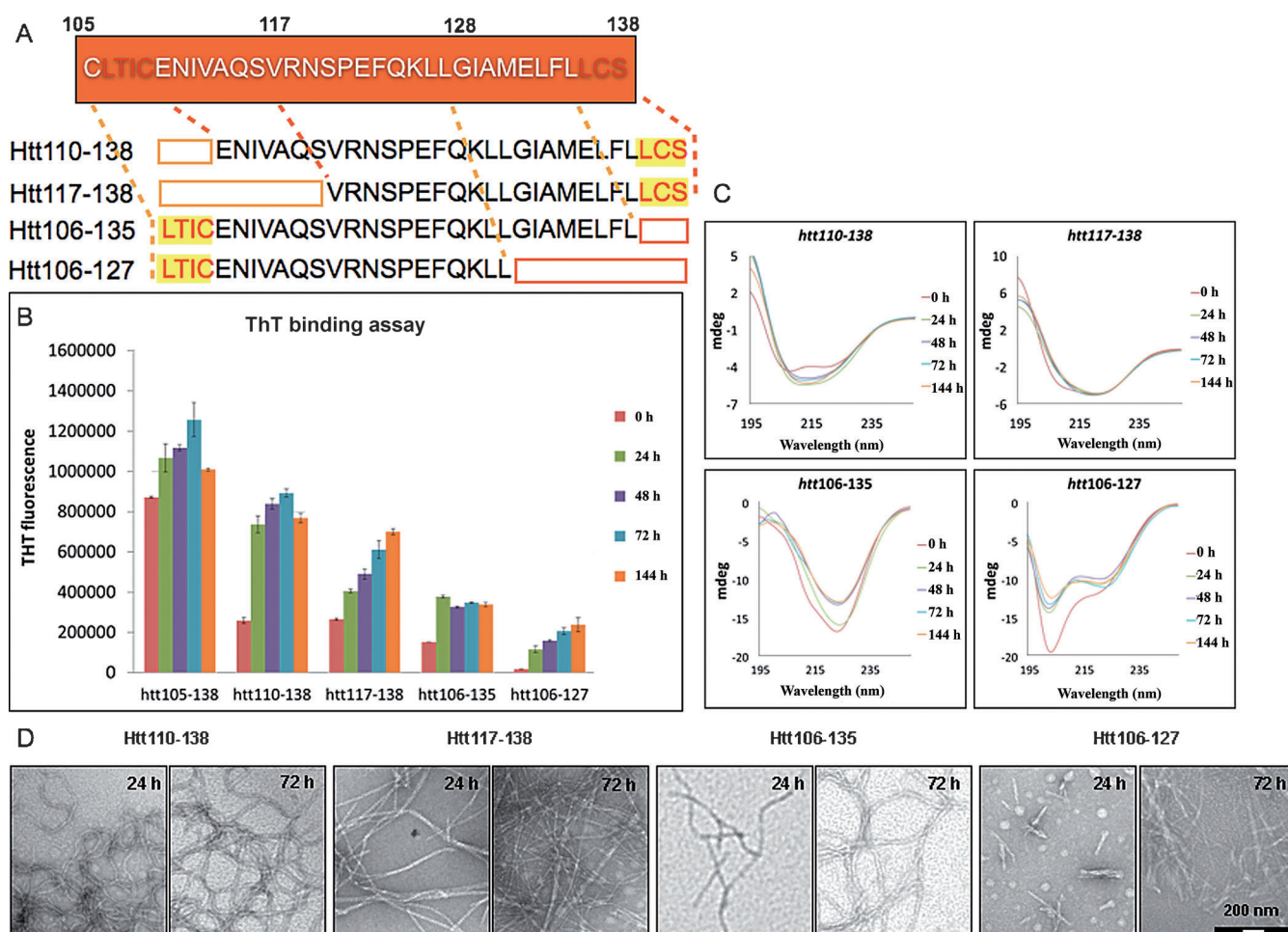


Figure 3. A) Schematic depiction of the various peptides that were made to validate the relative contribution of the N- and C-terminal regions (yellow background) to the aggregation of 105–138, B) Comparison of the extent of fibril formation of the 5 peptides shown in (A) by using the ThT assay CD spectroscopy (C) and TEM (D): samples were prepared as 40 μ M solutions in 10 mM Tris pH 7.4 buffer; the scale bar in (D) represents 100 nm.

Htt105–138 is located within the NtHtt fragments and its residues span exon2 (Htt91–118) and exon3 (Htt119–158), Figure 1A. Interestingly, two NtHtt fragments used to generate animal models of HD (Htt1–117 and Htt1–171)^[12a,b] were shown to give rise to different phenotypes. For the Htt1–117 fragment, one of the amyloidogenic regions that we identified (106–116) is present and highly exposed, whereas in the Htt1–171 NtHtt fragment, the entire amyloidogenic domain, 105–138, is present. Based on our results, we predict that both NtHtt fragments would exhibit increased aggregation propensity relative to exon1 (1–90), with similar polyQ repeats. The overexpression of an expanded variant of Htt1–117 (“shortstop” mice) was reported to result in significantly increased and widespread inclusions compared with an exon1 variant containing the same polyQ repeats. Despite this increased aggregation, the shortstop mice did not show any signs of neurodegeneration or behavioral deficits.^[12a] However, these findings were recently challenged by another study demonstrating that overexpression of Htt1–118–82Q with a different promoter resulted in phenotypes similar to the R6/2 and N171–82Q models. Despite the differences in terms of pathological phenotypes, overexpres-

sion of mutant forms of Htt1–117/118 result in enhanced aggregation or the formation of larger inclusions.^[15]

Our findings that the peptide (LTICENIVAQS) corresponding to the C-terminal residues of Htt1–117 and 1–118(82Q) exhibits a high propensity to aggregate may explain the increased aggregation of the shortstop protein and the formation of larger inclusions in the transgenic mice expressing Htt1–118(82Q). Similarly, the expanded (82Q) Htt1–171 formed more neuronal pathology and showed greater toxicity than the exon1 and the expanded FL–Htt.^[13] This fragment contains the entire amyloidogenic domain (105–138) and thus would be expected to exhibit increased aggregation and inclusion formation compared with the shorter N-terminal fragments (exon1 and shortstop). Altogether, these findings are consistent with increasing evidence indicating that although the expanded polyQ in Htt and other proteins plays a central role in the pathogenesis of HD, it is not the sole determinant of Htt aggregation and toxicity. The flanking sequences within exon1 and the length of the NtHtt fragment influence the protein’s structure, aggregation, subcellular localization and toxicity.

Herein, we describe for the first time the discovery of a novel amyloidogenic domain (105–138) outside exon1 of the Htt protein. A systematic analysis of peptides spanning different regions within this domain led to the identification of two distinct sequence motifs (108–116 and 128–138) that are responsible for the aggregation of this domain. Our results suggest that the aggregation propensity of this novel amyloidogenic domain may influence the rate and aggregation pathway of the NtHtt fragments and their toxicity to an extent that is dependent on whether either or both amyloid-forming motifs are present in these fragments. Given that multiple Htt fragments, extending beyond residue 138 have been shown to exist in vivo, conducting a similar and more expanded systematic analysis of N-terminal residues beyond residue 138 would be interesting to determine whether these regions may contain other amyloidogenic or protective sequences. The interplay between these different domains may account for the differences in aggregation, toxicity and disease pathology exhibited by the various N-terminal Htt fragments.

Taken together, our results provide novel insight into the possible molecular bases underlying the differences in aggregation, toxicity and neurologic phenotypes observed for the expanded N-terminal fragments spanning exons1–3 of Htt in HD model systems. Further studies are required to investigate potential cross-talk between the novel aggregation motifs identified in this work and the polyQ repeat region and elucidate their potential roles in regulating the physiological and pathogenic properties of the full-length protein and disease-associated N-terminal fragments in cell culture and animal models of HD.

Received: August 14, 2012

Published online: November 12, 2012

Keywords: aggregation · amyloid fibrils · Huntington's disease · N-terminal fragment

- [1] a) J. M. van der Burg, M. Bjorkqvist, P. Brundin, *Lancet Neurol.* **2009**, 8, 765.
- [2] R. R. Brinkman, M. M. Mezei, J. Theilmann, E. Almqvist, M. R. Hayden, *Am. J. Hum. Genet.* **1997**, 60, 1202.
- [3] a) J. M. Gil, A. C. Rego, *Eur. J. Neurosci.* **2008**, 27, 2803; b) M. Borrell-Pagès, D. Zala, S. Humbert, F. Saudou, *Cell. Mol. Life Sci.* **2006**, 63, 2642.
- [4] a) C. L. Wellington, L. M. Ellerby, A. S. Hackam, R. L. Margolis, M. A. Trifiro, R. Singaraja, K. McCutcheon, G. S. Salvesen, S. S. Propp, M. Bromm, K. J. Rowland, T. Zhang, D. Rasper, S. Roy, N. Thornberry, L. Pinsky, A. Kakizuka, C. A. Ross, D. W. Nicholson, D. E. Bredesen, M. R. Hayden, *J. Biol. Chem.* **1998**, 273, 9158; b) C. L. Wellington, M. R. Hayden, *Curr. Opin. Neurol.* **1997**, 10, 291; c) G. Bates, *Lancet* **2003**, 361, 1642; d) L. Mangiarini, K. Sathasivam, A. Mahal, R. Mott, M. Seller, G. P. Bates, *Nat. Genet.* **1997**, 15, 197.
- [5] a) M. Weigell-Weber, W. Schmid, R. Spiegel, *Am. J. Med. Genet.* **1996**, 67, 53; b) Y. Trottier, V. Biancalana, J. L. Mandel, *J. Med. Genet.* **1994**, 31, 377.
- [6] a) S. Chen, V. Berthelie, W. Yang, R. Wetzel, *J. Mol. Biol.* **2001**, 311, 173; b) C. C. Lee, R. H. Walters, R. M. Murphy, *Biochemistry* **2007**, 46, 12810.
- [7] A. K. Thakur, M. Jayaraman, R. Mishra, M. Thakur, V. M. Chellgren, I. J. Byeon, D. H. Anjum, R. Kodali, T. P. Creamer, J. F. Conway, A. M. Gronenborn, R. Wetzel, *Nat. Struct. Mol. Biol.* **2009**, 16, 380.
- [8] a) J. G. Hodgson, D. J. Smith, K. McCutcheon, H. B. Koide, K. Nishiyama, M. B. Dinulos, M. E. Stevens, N. Bissada, J. Nasir, I. Kanazawa, C. M. Disteche, E. M. Rubin, M. R. Hayden, *Hum. Mol. Genet.* **1996**, 5, 1875; b) J. G. Hodgson, N. Agopyan, C. A. Gutekunst, B. R. Leavitt, F. LePiane, R. Singaraja, D. J. Smith, N. Bissada, K. McCutcheon, J. Nasir, L. Jamot, X. J. Li, M. E. Stevens, E. Rosemond, J. C. Roder, A. G. Phillips, E. M. Rubin, S. M. Hersch, M. R. Hayden, *Neuron* **1999**, 23, 181.
- [9] a) E. Rockabrand, N. Slepko, A. Pantalone, V. N. Nukala, A. Kazantsev, J. L. Marsh, P. G. Sullivan, J. S. Steffan, S. L. Sensi, L. M. Thompson, *Hum. Mol. Genet.* **2007**, 16, 61; b) Z. H. Qin, Y. Wang, E. Sapp, B. Cuiffo, E. Wanker, M. R. Hayden, K. B. Kegel, N. Aronin, M. DiFiglia, *J. Neurosci.* **2004**, 24, 269; c) P. W. Faber, C. Voisine, D. C. King, E. A. Bates, A. C. Hart, *Proc. Natl. Acad. Sci. USA* **2002**, 99, 17131.
- [10] S. Tam, C. Spiess, W. Auyeung, L. Joachimiak, B. Chen, M. A. Poirier, J. Frydman, *Nat. Struct. Mol. Biol.* **2009**, 16, 1279.
- [11] A. Bhattacharyya, A. K. Thakur, V. M. Chellgren, G. Thiagarajan, A. D. Williams, B. W. Chellgren, T. P. Creamer, R. Wetzel, *J. Mol. Biol.* **2006**, 355, 524.
- [12] a) E. J. Slow, R. K. Graham, A. P. Osmand, R. S. Devon, G. Lu, Y. Deng, J. Pearson, K. Vaid, N. Bissada, R. Wetzel, B. R. Leavitt, M. R. Hayden, *Proc. Natl. Acad. Sci. USA* **2005**, 102, 11402; b) G. Schilling, M. W. Becher, A. H. Sharp, H. A. Jinnah, K. Duan, J. A. Kotzuk, H. H. Slunt, T. Ratovitski, J. K. Cooper, N. A. Jenkins, N. G. Copeland, D. L. Price, C. A. Ross, D. R. Borchelt, *Hum. Mol. Genet.* **1999**, 8, 397; c) E. Waldron-Roby, T. Ratovitski, X. Wang, M. Jiang, E. Watkin, N. Arbez, R. K. Graham, M. R. Hayden, Z. Hou, S. Mori, D. Swing, M. Pletnikov, W. Duan, L. Tessarollo, C. A. Ross, *J. Neurosci.* **2012**, 32, 183.
- [13] Z. X. Yu, S. H. Li, J. Evans, A. Pillarisetti, H. Li, X. J. Li, *J. Neurosci.* **2003**, 23, 2193.
- [14] S. Chen, R. Wetzel, *Protein Sci.* **2001**, 10, 887.
- [15] A. T. Tebbenkamp, D. Swing, L. Tessarollo, D. R. Borchelt, *Hum. Mol. Genet.* **2011**, 20, 1633.