J|A|C|S COMMUNICATIONS

Published on Web 01/07/2010

Induction of Amyloid Fibrils by the C-Terminal Fragments of TDP-43 in Amyotrophic Lateral Sclerosis

Allan K.-H. Chen,[†] Ryan Y.-Y. Lin,[†] Eva Z.-J. Hsieh,[†] Pang-Hsien Tu,[‡] Rita P.-Y. Chen,[§] Tai-Yan Liao,[§] Wenlung Chen,[||] Chih-Hsien Wang,[||] and Joseph J.-T. Huang*,[†]

Institute of Chemistry, Biomedical Sciences, and Biological Chemistry, Academia Sinica, No. 128, Sec 2, Academia Road, Nankang, Taiwan, R.O.C., and Department of Applied Chemistry, National Chiayi University, 300 University Road, Chiayi, Taiwan, R.O.C.

Received August 5, 2009; E-mail: jthuang@chem.sinica.edu.tw

In year 2006, the TAR DNA-binding protein (TDP-43) was identified as the main sediments in the histopathological inclusion bodies of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD-U).[1] ALS is a progressive neurodegenerative disease affecting the degeneration of motor neurons. As yet, there is no treatment of this fatal disease.[2] The discovery of ubiquitinated TDP-43 in ALS patients has opened up a new window to explore these motor neuron degenerative diseases. Although TDP-43 has been directly implicated as a culprit, why this nuclear protein becomes N-terminally truncated and forms aggregates in the cytoplasmic environment remains unclear.

TDP-43 contains 414 residues, with two central RNA recognition motifs flanked by N-terminal and C-terminal domains (Figure 1A).[1] At this juncture, conformation information is only available for the RNA recognition motifs[3] (RRM1: 2cqg and RMM2: 1wf0, 3d2w). There is no structural information of any kind on the C-terminal domain. Recent literature has indicated that the major component in neuronal inclusion bodies in ALS and FTLD patients are C-terminal fragments that have been hyper-phosphorylated and ubiquitinated.[1,4−6]

The importance of the TDP-43 C-terminus in ALS has provided a rational to analyze the sequence in this region. Based on the predictor of naturally disordered regions (PONDR PROGRAM),[7] it appears that 80% of the C-terminus is structurally disordered, with the scores all higher than 0.5, starting from 287 to 414 (Figure 1A). Intrinsically disordered proteins or structurally disordered regions of proteins could be involved in a myriad of functions, including DNA recognition, modulation of protein binding, molecular threading, activation of cleavage, and control of protein lifetimes. Although these intrinsic disordered regions lack secondary or tertiary structures in their native states, they may undergo disorder-to-order transitions upon binding to their partners.[8]

In an effort to clarify the conformational properties and other physical characteristics of the disordered regions (sequence: 287−414) of the C-terminal domain, we have synthesized four peptide fragments including D1 (287−322), D2 (321−353), D3 (351−383), and D4 (381−414) (see Supporting Information (SI)). When 50 μM of D1−4 peptides were incubated in phosphate buffer (70 mM KCl, 20 mM sodium phosphate, pH 7.0) at 37 °C for 2 weeks, electron microscopy (EM) showed that only D1 formed fibers, among the four fragments, either straight or gently curved, with a cross section of ∼11 nm in width (Figure 1B). Different strings of fibrils could further twist around each other and associate to form thicker filaments. In contrast, the D2−4 peptides formed amorphous aggregates, showing different aggregation propensities (see SI). From analysis of the sequences of D1−4, we found that ∼70% of D1 belonged to the glycine rich domain (GRD). Moreover, with the aid of the protein aggregation predictor PROGRAM PASTA, we found that a large peak (309−318) was located at the end of the D1 and extended a little to D2 (see SI), showing an extraordinary aggregation propensity here. Since fibrillar aggregates of truncated TDP-43 are found in ALS patients,[9] it is very likely that this sequence in D1 plays an important role in aggregation as well as in promoting the fiber formation.

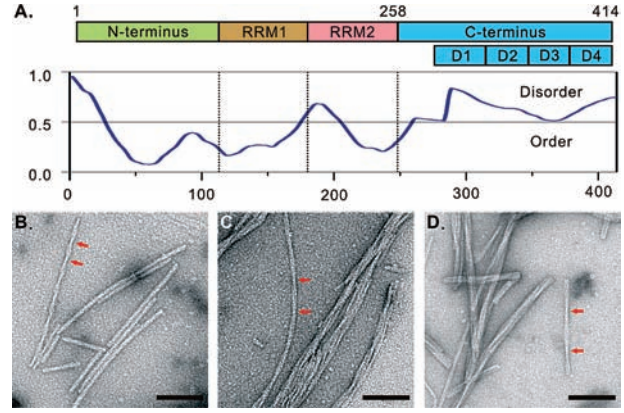

Figure 1. (A) PONDR analysis of the full-length TDP-43 (1−414). Calculations use VL3 version of PONDR. EM images of (B) D1, (C) A315T, and (D) G294A incubated in pH 7.0 phosphate buffer at 37 °C for 2 weeks. The scale bars represent 100 nm.

In 2007, Cairns et al. identified A315T as a novel missense mutation in Familial Motor Neuron Disease patients.[10] In TDP43, this alanine residue is highly conserved from *Homo sapiens* to *Xenopus tropicalis*, and this A315T mutation is absent from a large number (1505) of ethnically matched elderly control subjects.[10] Thus, it is very likely that this alanine residue is involved in the function of TDP-43 and the replacement of this alanine may cause protein aggregation and disease. Within the same year, Christopher et al. found another mutation G294A in sporadic ALS cases.[11]

To explore whether or not pathological mutations can alter the aggregation propensity of D1 as well as the conformational properties of the protein in the D1 region, we have synthesized the A315T and G294A mutants (see SI). Incubation of A315T and G294A in phosphate buffer for 2 weeks reveals formation of fibrils similar to those of D1 (Figure 1C and D). The arrows in Figure 1B−D indicate intertwined fibrils with periodicities of approximately 43, 68, and 85 nm in D1, A315T, and G294A fibrils, respectively. According to the immunoreactive TEM studies on the sporadic and familial cases of FTLD with TDP-43 proteinopathy,

[†] Institute of Chemistry Academia Sinica.
[‡] Institute of Biomedical Sciences Academia Sinica.
[§] Institute of Biological Chemistry Academia Sinica.
[||] Department of Applied Chemistry, National Chiayi University.

10.1021/ja9066207

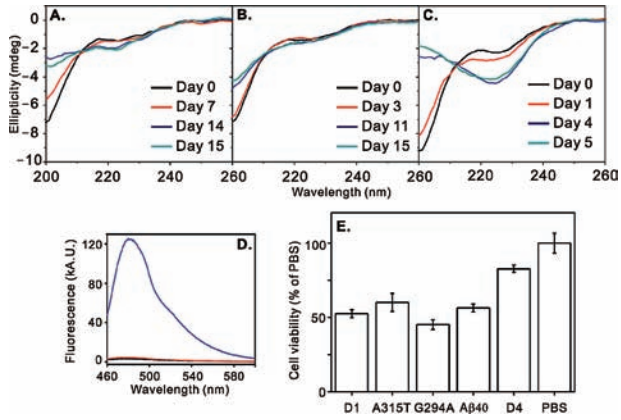

Figure 2. Time course CD measurements of (A) D1, (B) A315T, and (C) G294A peptides. The results of ThT staining fluorescence experiments for 50 μM of D1 (black), A315T (red), and G294A (blue) are shown in (D). Comparison of the cytotoxicities of D1, A315T, G294A, Aβ40, and D4 with PBS measured by MTT assay is shown in (E) with error bars.

predominantly fibrillary cytoplasmic inclusions (NCIs) are also found with intense TDP-43 positivity.[9] While the average filament width (range) of these fibrils is 9 nm (4−16 nm), the mean widths of our fibrils including D1, A315T, and G294A are all ∼11 nm.

Despite homology in the fibril's morphology, peptides of D1, A315T, and G294A show marked differences in their propensity toward secondary structure transformation during the fibrillization process as monitored by circular dichroism (CD) over time. The results show that the initial conformations of D1, A315T, and G294A are all random coil structures at day 0 when dissolved in phosphate buffer (Figure 2A−C, black line). The time course of the CD spectra indicates that the content of random coil structures of D1 and A315T decreases with time according to the decrease in the ellipticity at 200 nm (Figure 2A and B), and there is little change in CD spectra after 15 days. On the other hand, G294A exhibits a fast transformation from random coil to β-sheet content within 5 days in light of the significant increase at 218 nm (Figure 2C).

The Thioflavin T (ThT) assay is commonly employed for amyloid fibril identification, though the detailed mechanism and the specificity of binding remain poorly understood.[12,13] Binding of ThT to amyloid fibrils can exhibit enhancement of emission intensity at ∼482 nm upon excitation at 442 nm. Our results show that the fluorescence of D1 and A315T is not induced by the addition of ThT after incubating in phosphate buffer for 2 weeks. This makes D1 and A315T fibrils special examples without spectroscopic characteristics of amyloid. Only the G294A showed extraordinary strong emission around 482 nm (Figure 2D). However, to our surprise, the FT-Raman spectra of all three fibril samples (D1, A315T, and G294A) at the amide I region (1590−1720 cm^{-1}) have shown that all fibrils are β-sheet rich structures (see SI). From the deconvolution and numerical curve fitting, the β-sheet components (%) of D1, A315T, and G294A are 78%, 64%, and 73%, respectively. We presume that this Gly-rich segment is in the core of the amyloid fibril. Because Gly is achiral and devoid of a steric side chain, the fibrils formed from D1 and A315T did not give a clear negative ellipticity at 218 nm in CD[14] and ThT cannot anchor into these fibrils properly and emit strong fluorescence.

Despite the structural similarities among the three peptides, D1, G294A, and A315T are toxic to mouse N2a neuroblastoma cells compared with D4 (forming nonfibrillar aggregates) and the positive control (Aβ40). The cell viability associated with these three peptides ranges from 45% to 60% (Figure 2E; details in SI). Furthermore, it has been long recognized that aggregation kinetics could reveal the amyloid formation kinetics by plotting the time course of monomer concentration decrease.[15] Our result showed that the kinetics of fibril formation of A315T and G294A were much faster than D1 (data not shown).

In summary, we have shown that the D1 peptide within D1−4 can form twisted fibrils under the incubation of phosphate buffer. This Gly-rich segment on D1 might prohibit the amorphous aggregation as ground state conformers in D2−4. Similar Gly-rich segments are also found in amyloid β peptides and prion proteins. In addition to D1, the mutants A315T and G294A also form fibrils that share a similar shape and morphology with NCIs. Since all these peptides contain numerous glycine repeats (13 glycines in D1 and A315T; 12 glycines in G294A) and are predicted to be natively disordered, we propose that the residues in this region (287−322) may contribute significantly to the fiber formation as well as aggregation propensity. From the EM and Raman spectroscopic properties, we found that D1, A315T, and G294A all formed an amyloid structure though only G294A showed noticeable ThT fluorescence, providing insights into the nature of their aggregation *vis a vis* the other fragments in the C-terminus of TDP-43. Moreover, the pathological mutations have accelerated the amyloidogenesis in these fragments.

Acknowledgment. This work was supported by the Academia Sinica and the National Science Council, Taiwan (NSC 97-2113-M-001-003-MY2). We thank Prof. Sunney I. Chan for helpful discussions. We also thank Shu-Ping Li (EM), Yao-Yin Chung (EM), and Dr. Mei-Chun Tseng (mass spectrometry).

Supporting Information Available: Materials and methods; complete refs 1, 10, 11, and 15; sequences of peptides; EM of D2−4; PASTA aggregation prediction of C-terminus; and FT-Raman spectra are summarized. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Neumann, M.; et al. *Science* **2006**, *314*, 130–3.
(2) Banks, G. T.; Kuta, A.; Isaacs, A. M.; Fisher, E. M. *Mamm. Genome* **2008**, *19*, 299–305.
(3) Kuo, P. H.; Doudeva, L. G.; Wang, Y. T.; Shen, C. K.; Yuan, H. S. *Nucleic Acids Res.* **2009**, *37*, 1799–808.
(4) Igaz, L. M.; Kwong, L. K.; Chen-Plotkin, A.; Winton, M. J.; Unger, T. L.; Xu, Y.; Neumann, M.; Trojanowski, J. Q.; Lee, V. M. *J. Biol. Chem.* **2009**, *284*, 8516–24.
(5) Zhang, Y. J.; Xu, Y. F.; Cook, C.; Gendron, T. F.; Roettges, P.; Link, C. D.; Lin, W. L.; Tong, J.; Castanedes-Casey, M.; Ash, P.; Gass, J.; Rangachari, V.; Buratti, E.; Baralle, F.; Golde, T. E.; Dickson, D. W.; Petrucelli, L. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 7607–12.
(6) Hasegawa, M.; Arai, T.; Nonaka, T.; Kametani, F.; Yoshida, M.; Hashizume, Y.; Beach, T. G.; Buratti, E.; Baralle, F.; Morita, M.; Nakano, I.; Oda, T.; Tsuchiya, K.; Akiyama, H. *Ann. Neurol.* **2008**, *64*, 60–70.
(7) Radivojac, P.; Obradovic, Z.; Brown, C. J.; Dunker, A. K. *Pac. Symp. Biocomput.* **2003**, 216–27.
(8) Dunker, A. K.; Garner, E.; Guilliot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J. E. *Pac. Symp. Biocomput.* **1998**, 473–84.
(9) Thorpe, J. R.; Tang, H.; Atherton, J.; Cairns, N. J. *J. Neural. Transm.* **2008**, *115*, 1661–71.
(10) Gitcho, M. A.; et al. *Ann. Neurol.* **2008**, *63*, 535–8.
(11) Sreedharan, J.; et al. *Science* **2008**, *319*, 1668–72.
(12) Nilsson, M. R. *Methods* **2004**, *34*, 151–60.
(13) Fandrich, M. *Cell. Mol. Life Sci.* **2007**, *64*, 2066–78.
(14) Darnell, G.; Orgel, J. P.; Pahl, R.; Meredith, S. C. *J. Mol. Biol.* **2007**, *374*, 688–704.
(15) Thakur, A. K.; et al. *Nat. Struct. Mol. Biol.* **2009**, *16*, 380–9.

JA9066207