dGAE(297-391) tau fragment promotes formation of CTE-like full-length tau filaments

Kristine Kitoka[1], Alons Lends[1], Gytis Kučinskas[2,3], Anna Lina Bula[1], Lukas Krasauskas[4], Vytautas Smirnovas[4], Rostislav Skrabana[5], Jozef Hritz[2,6], and Kristaps Jaudzems[1]*

1) Latvian Institute of Organic Synthesis, Riga, Latvia
2) Central European Institute of Technology, Masaryk University, Brno, Czech Republic
3) Faculty of Science, National Centre for Biomolecular Research, Masaryk University, Brno, Czech Republic
4) Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania
5) Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia
6) Department of Chemistry, Faculty of Science, Masaryk University, Brno, Czech Republic

Abstract

The microtubule associated protein tau forms disease-specific filamentous aggregates in several different neurodegenerative diseases. In order to understand how tau undergoes misfolding into a specific filament type and to be able to control this process for drug development purposes, it is crucial to study *in vitro* tau aggregation methods and investigate the structures of the obtained filaments at atomic level. Here, we used the tau fragment dGAE(297-391), which aggregates spontaneously, to seed formation of full-length tau filaments. The structures of dGAE and full-length tau filaments were investigated by solid-state MAS NMR showing that dGAE allows propagating a chronic traumatic encephalopathy (CTE)-like fold to full-length tau. This work demonstrates that *in vitro* preparation of disease-specific types of full-length tau filaments is feasible.

Introduction

Protein misfolding into insoluble amyloid deposits is a hallmark of many neurodegenerative diseases [1]. Usually, the amyloid-forming proteins are poorly soluble and aggregation-prone under physiological conditions. However, the microtubule-associated protein tau is highly soluble [2,3]. Despite this, insoluble tau filaments are formed in several neurodegenerative diseases such as Alzheimer's disease (AD) [4], corticobasal degeneration (CBD) [5], chronic traumatic encephalopathy (CTE) [6], and Pick's disease (PiD) [7]. Remarkably, the adopted folds of tau filaments are disease specific [4-7].

In the brain, tau is present as six different isoforms comprising 352-441 residues produced through alternative splicing. The isoforms differ by the number of N-terminal inserts (0N, 1N or 2N) and C-terminal repeats (3R or 4R isoforms) [8]. An enormous amount of work utilizing various biophysical techniques has been done to investigate the filamentous tau structures and identify the parts of the protein which are incorporated into the rigid cores of the filaments [9,10]. Typically, the disease-associated tau filament cores do not exceed one-quarter of the protein sequence. In electron micrographs, the filaments

appear to be surrounded by a fuzzy outer coat [11-13]. A recent breakthrough discovery using cryo-EM on patient-derived material revealed that the rigid core of AD paired helical and straight filaments is formed by the V306-F378 fragment [4]. This region largely overlaps with the I297-E391 (dGAE) fragment (Figure 1A), previously found to be the main component of AD patient-derived material [11]. dGAE covers the end of R2, whole R3 and R4 repeats, and the major part of R' repeat of the full-length (2N4R) tau protein. Although the dGAE fragment is slightly longer than the AD tau filament core, Lovestam et al. showed that the positions of its N- and C-termini are essential for filament formation *in vitro* because several shorter tau constructs were ineffective in forming filaments. Additionally, high-throughput cryo-EM structure determination has proven that dGAE can form filaments *in vitro*, resembling those found in the CTE and AD patient brains [14].

In contrast to full-length tau, dGAE aggregates in the absence of any aggregation inducers, allowing preparation of filaments *in vitro* without heparin [15-17]. Heparin screens electrostatic interactions, thereby inducing conformational rearrangement of the tau protein that leads to its self-assembly [18,19]. Until recently, heparin was the most commonly used cofactor for preparing tau filaments, as it ensures relatively high aggregation yields [20-25]. However, cryo-EM has shown that heparin-induced tau filaments are structurally heterogeneous, questioning the relevance of such aggregation protocols [26]. Thus, dGAE provides a biologically more relevant route to generate tau filaments for *in vitro* studies, necessary for developing new therapies against tau. However, it remains elusive whether the findings from dGAE studies can be directly translated into the context of full-length tau.

Here, we use solid-state NMR (ssNMR) and other complementary biophysical techniques to characterize the structures of both the truncated tau fragment dGAE and the full-length tau filaments obtained by seeding with dGAE. By comparing the spectra of dGAE and seeded full-length 2N4R filaments, we demonstrate that this approach allows preparation of CTE-like full-length tau filaments *in vitro*.

Results

dGAE filaments were formed by incubating a high concentration (400 μM) protein solution without cofactors in PBS pH 7.4. Continuous orbital agitation at 800 rpm and 37 °C was used for 48 h under reducing conditions (5 mM DTT) to prevent unwanted disulfide bond formation by C322 [4,27]. Thioflavin T (ThT) fluorescence showed a typical sigmoidal curve indicative of amyloid formation via primary nucleation, elongation, and secondary nucleation (Figure 1B) [28]. The average lag time for filament formation was approximately 8 h, after which an exponential increase in ThT fluorescence was observed with a half-time of 12 h reaching a plateau at 20 h. The small standard deviation values indicate repeatable aggregation kinetics of dGAE under these conditions. SDS-PAGE analysis (Figure 1C) confirms that the vast majority of dGAE was aggregated at the end of the incubation period. The morphology of the obtained filaments was characterized by negative-stain electron microscopy (Figure 1D). The micrographs show the formation of

slightly twisted and unbranched filaments without significant macroscopic heterogeneity or amorphous aggregates.

Solid-state NMR was performed on uniformly ^{13}C,^{15}N-labeled dGAE filaments using ^{13}C -detected experiments at 12 kHz MAS. 2D ^{13}C -^{13}C correlation spectrum with 20 ms mixing DARR displayed line widths from 0.4 to 1.0 ppm for well isolated peaks, suggesting that the filament preparation yielded a single conformation. Examination of several residue-type-specific regions (cysteine, serine, threonine, proline, isoleucine, valine, and alanine regions) indicated the expected number of correlations (Figure S1A), corroborating high homogeneity of the sample.

Sequence-specific assignment of the backbone ^{15}N and ^{13}C chemical shifts were obtained for residues V306-D314, V318-F346, R349-L357, D358-G365, H388-A390 from the 3D NCACX, NCOCX and CANCO experiments (Figure S2). Side chain assignments were completed using the 2D DARR spectra with longer mixing times. To simplify the spectra and confirm the assignment of lysine, leucine, isoleucine, and valine residues, two selectively unlabeled samples were prepared and analyzed by suppressing either lysine or leucine, isoleucine, valine and lysine labeling (Figure S3). The assigned 2D NCA spectrum is shown in (Figure 2A). The obtained assignment suggests that at least 50 residues spanning the region V306-G365 are incorporated into the rigid core of our dGAE filaments. The chemical shifts of C322 (Cα at 60.4 ppm and Cβ at 24.9 ppm) correspond to a reduced cysteine, in agreement with the reducing conditions used. Residues H388-A390 were assigned tentatively, implying that the C-terminal part folds back and attaches to the filament core.

To probe the mobile parts outside the rigid filament core, we recorded INEPT-based experiments that allow the detection of highly dynamic residues. In the INEPT-based 2D ^{1}H-^{13}C correlation spectrum (Figure 2B), the few cross peaks exhibiting random coil chemical shifts were assigned to methionine, isoleucine, lysine, histidine, valine, glycine, alanine, and threonine residues. Based on the uniqueness of some amino acids outside the assigned core, these mobile residues most likely belong to fragments I297-G302 and A382-T386. Altogether the MAS ssNMR data indicate that the dGAE rigid core is composed of residues V306-G365, which interacts with the very C-terminus (residues H388-A390), the N-terminus (residues I297-G302) and residues A382-T386 near the C-terminus are mobile, whereas residues G366-N381 are semi-rigid as they are not observed in either cross-polarization or INEPT-based experiments.

To identify the locations of secondary structures and make a comparison with other tau structures, deviations of backbone Cα and Cβ chemical shifts from random-coil values, known as the secondary chemical shifts (SCS), were analyzed. Most of the assigned residues were found to adopt a β-sheet conformation, whereas unassigned fragments coincide with loops in the cryo-EM structures (Figure 2C). Strand β1 covers the R3 hexapeptide motif V306–Y310, β2 is formed by V313-D314, β3 by V318-K321, β4 by S324-H330. Strand β5 is formed by V337-F346 and covers the R4 hexapeptide motif V337-E342, V350-G355 forms β6, and L357-H362 forms β7. As expected, both hexapeptides are involved in dGAE filament formation. Compared to the cryo-EM

structures of patient-derived tau, our assignment covers residues that belong up to the β7 strand of the cryo-EM structures. The locations of β1-β7 identified in our dGAE filaments fit the best to the CTE I cryo-EM data (Figure 2C). We also compared the assigned β-strand locations to other dGAE in vitro studies, cryo-EM with the closest possible conditions (PDB: 7QJW, 7QLS, 7QK2) [14] and ssNMR [17]. The comparison shows that overall β-strand positions differ, especially in the region of R4 hexapeptide.

dGAE filaments were used as a template for seeded aggregation of full-length tau, using identical conditions as for dGAE alone. The sample was incubated for three days. The efficiency of aggregation was assessed by SDS-PAGE (Figure 3A) which showed that a major part of the full-length tau protein remained in the solution (Figure 3A. Lane S), and only a minor part was in the pellet (Figure 3A. Lane P). This suggests that full-length tau reaches an equilibrium of aggregation/dissolution processes even in the presence of dGAE seeds.

The obtained full-length tau filaments were long and formed a net-like arrangement on EM grids (Figure 3B). Compared with dGAE filaments, the body of the 2N4R filaments seems to be covered by additional densities, most likely the fuzzy coat. In the AFM micrographs, the 2N4R filaments appear to have left-handed twists (Figure 3C).

Structural analysis of dGAE-seeded full-length tau filaments was performed by comparison of 2D solid-state MAS NMR fingerprints with dGAE filaments. Figure 4A shows an overlay of the ^{1}H-^{13}C INEPT spectra of dGAE and seeded full-length tau. Full-length tau filaments exhibit many more intense signals in the INEPT spectrum, confirming the formation of full-length tau filaments with mobile flanking regions. Several resonances, such as isoleucine and alanine, overlap in these spectra, and some become more intense in the full-length filaments.

Notable differences are also observed between the DARR spectra (Figure 4B). Alanine resonances, which in the DARR spectrum of dGAE exhibit random coil chemical shifts, are shifted toward beta-sheet chemical shifts. One of these resonances most likely belongs to A390. The increased number of alanine resonances suggests that additional β-strands outside V306-F378 are present in the full-length tau filaments. In addition, the $C322_{dGAE}$ peak is significantly shifted/missing in the full-length DARR. The changes to the C322 peak indicate that the C322 conformation and/or dynamics are different in full-length tau filaments.

The overlay of 2D NCA (Figure 4C) shows the disappearance of several resonances assigned to dGAE in the full-length tau spectrum. These include the C322-S324 fragment and several glycines. Surprisingly, both K340 and S341 also disappear, although they are part of the R4 hexapeptide V337-E342. The INEPT spectrum of 2N4R (Figure 4A) does not exhibit any new resonances that could belong to serine. Taken together, these changes most likely indicate that the R4 hexapeptide, when incorporated into full-length filaments, is influenced by the mobile regions of full-length tau filaments.

Overall, the DARR spectra of dGAE and full-length tau filaments look quite similar. However, we observed that the R4 repeat of full-length tau filaments is most likely affected by mobile regions outside the rigid core. Moreover, a short immobilized fragment H388-A390 lost its residual mobility in full-length tau filaments. This leads us to think the C-terminal part also contributes to filament rigid core structure. In particular, this is an interesting observation because these parts were not detected in cryo-EM studies of tau filaments so far.

Discussion

Recently, several groups have studied filaments formed by dGAE(297-391) fragment [14,27,29,30]. Lovestam et al. showed that the use of NaCl or PBS gives rise to either CTE-type filaments or new morphologies [14]. Although similar solution conditions were used for aggregation in our case, the secondary structures (Figure 2C) are different, which may be because of different shaking speeds and aggregation volumes (we used 1.5 mL microcentrifuge tubes containing 1 mL protein solution instead of microplates). Another ssNMR study of dGAE filaments determined secondary structures that differed because of different aggregation conditions. Interestingly, the rigid core identified in this study also lacked the C-terminal beta strands from the cryo-EM structures [17].

To date, the major bottleneck in the structural studies of disease-relevant full-length tau filaments is the high solubility of tau. Spontaneous tau aggregation without cofactors is usually less efficient than cofactor-induced aggregation. In addition, aggregation becomes even less efficient by increasing the length of tau from truncated constructs to full-length isoforms [31,32]. In this study, we showed that full-length tau protein can be efficiently aggregated by template-driven aggregation using the truncated tau fragment dGAE(297-391), which may pave the way for disease-specific full-length tau filament studies. However, seeded full-length tau aggregation time was approximately 3-fold longer compared to dGAE alone.

In summary, we used MAS NMR spectroscopy to investigate the features of tau filaments in both dGAE and full-length. The rigid core of the obtained single-conformation dGAE filaments spans residues V306-G365 based on the assigned NMR chemical shifts. We were able to assign residues up to β7 seen in cryo-EM. In addition, we assigned an immobilized short fragment with no β-sheet propensity T387-A390 at the C-terminus of dGAE. We expect that this short fragment in R' forms cross-β hydrogen bonds and joins the rigid core. Taken together, this work sets the basis for future solid-state NMR spectroscopy to investigate the role of flanking mobile regions of disease-specific tau filaments. A better understanding of the interactions between the rigid core and mobile regions could help to identify important drug-binding sites, allowing the design of more effective therapeutic candidates against the progression of AD and other tauopathies.

Acknowledgements

We acknowledge funding from the Latvian Council of Science grant no. lzp-2019/1-0244 and European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 873127. K.K. is supported by "Mikrotīkls doctoral scholarship in the field of exact and medical sciences", administered by the University of Latvia Foundation. A.L. is supported by Marie Curie widening fellowship nr. 101038074—"Oligomers-MAS-NMR". G.K. and J.H. acknowledge the Czech Science Foundation (no. GF20-05789L) and Nanobiotechnology and Cryo-electron microscopy and tomography core facilities CEITEC MU of CIISB, Instruct-CZ Centre supported by MEYS CR (LM2018127). R.S. is supported by the Slovak Research and Development Agency grant APVV-21-0479.

References

[1] F. Chiti, C. M. Dobson, Annual Review of Biochemistry 2006, 75, 333–366.
[2] Y. Wang, E. Mandelkow, Nature Reviews Neuroscience 2015, 17, 22–35.
[3] N. Sibille, A. Sillen, A. Leroy, J.-M. Wieruszeski, B. Mulloy, I. Landrieu, G. Lippens, Biochemistry 2006, 45, 12560–12572.
[4] A. W. P. Fitzpatrick, B. Falcon, S. He, A. G. Murzin, G. Murshudov, H. J. Garringer, R. A. Crowther, B. Ghetti, M. Goedert, S. H. W. Scheres, Nature 2017, 547, 185–190.
[5] W. Zhang, A. Tarutani, K. L. Newell, A. G. Murzin, T. Matsubara, B. Falcon, R. Vidal, H. J. Garringer, Y. Shi, T. Ikeuchi, S. Murayama, B. Ghetti, M. Hasegawa, M. Goedert, S. H. W. Scheres, Nature 2020, 580, 283–287.
[6] B. Falcon, J. Zivanov, W. Zhang, A. G. Murzin, H. J. Garringer, R. Vidal, R. A. Crowther, K. L. Newell, B. Ghetti, M. Goedert, S. H. W. Scheres, Nature 2019, 568, 420–423.
[7] B. Falcon, W. Zhang, A. G. Murzin, G. Murshudov, H. J. Garringer, R. Vidal, R. A. Crowther, B. Ghetti, S. H. W. Scheres, M. Goedert, Nature 2018, 561, 137–140.
[8] M. Goedert, M. Masuda-Suzukake, B. Falcon, Brain 2016, 140, 266–278.
[9] K. Kitoka, R. Skrabana, N. Gasparik, J. Hritz, K. Jaudzems, Frontiers in Molecular Biosciences 2021, 8, 761227.
[10] S. H. Scheres, W. Zhang, B. Falcon, M. Goedert, Current Opinion in Structural Biology 2020, 64, 17–25.
[11] C. M. Wischik, M. Novak, P. C. Edwards, A. Klug, W. Tichelaar, R. A. Crowther, Proceedings of the National Academy of Sciences 1988, 85, 4884–4888.
[12] A. Sillen, J.-M. Wieruszeski, A. Leroy, A. B. Younes, I. Landrieu, G. Lippens, Journal of the American Chemical Society 2005, 127, 10138–10139.
[13] S. Wegmann, I. D. Medalsy, E. Mandelkow, D. J. Müller, Proceedings of the National Academy of Sciences 2012, 110, DOI 10.1073/pnas.1212100110.
[14] S. Lövestam, F. A. Koh, B. van Knippenberg, A. Kotecha, A. G. Murzin, M. Goedert, S. H. Scheres, eLife 2022, 11, 76494.
[15] P. Jayan, A. A. Vahid, S. T. Kizhakkeduth, S. O. H. Muhammed, A. B. Shibina, V. Vijayan, ChemBioChem 2021, 22, 2093–2097.
[16] P. Chakraborty, G. Rivière, S. Liu, A. I. de Opakua, R. Dervişoğlu, A. Hebestreit, L. B. Andreas, I. M. Vorberg, M. Zweckstetter, Nature Communications 2021, 12, 4231.
[17] Y. K. Al-Hilaly, C. Hurt, J. E. Rickard, C. R. Harrington, J. M. D. Storey, C. M. Wischik, L. C. Serpell, A. B. Siemer, *Frontiers in Neuroscience* 2022, *16*, 988074.
[18] M. Goedert, R. Jakes, M. G. Spillantini, M. Hasegawa, M. J. Smith, R. A. Crowther, *Nature* 1996, *383*, 550–553.

[19] Y. Fichou, Y. Lin, J. N. Rauch, M. Vigers, Z. Zeng, M. Srivastava, T. J. Keller, J. H. Freed, K. S. Kosik, S. Han, *Proceedings of the National Academy of Sciences* 2018, *115*, 13234–13239.
[20] O. C. Andronesi, M. von Bergen, J. Biernat, K. Seidel, C. Griesinger, E. Mandelkow, M. Baldus, *Journal of the American Chemical Society* 2008, *130*, 5922–5928.
[21] V. Daebel, S. Chinnathambi, J. Biernat, M. Schwalbe, B. Habenstein, A. Loquet, E. Akoury, K. Tepper, H. Müller, M. Baldus, C. Griesinger, M. Zweckstetter, E. Mandelkow, V. Vijayan, A. Lange, *Journal of the American Chemical Society* 2012, *134*, 13982–13989.
[22] S. Xiang, N. Kulminskaya, B. Habenstein, J. Biernat, K. Tepper, M. Paulat, C. Griesinger, S. Becker, A. Lange, E. Mandelkow, R. Linser, *Journal of the American Chemical Society* 2017, *139*, 2639–2646.
[23] A. J. Dregni, V. S. Mandala, H. Wu, M. R. Elkins, H. K. Wang, I. Hung, W. F. DeGrado, M. Hong, *Proceedings of the National Academy of Sciences* 2019, *116*, 16357–16366.
[24] A. Savastano, G. Jaipuria, L. Andreas, E. Mandelkow, M. Zweckstetter, *Scientific Reports* 2020, *10*, 21210.
[25] A. J. Dregni, H. K. Wang, H. Wu, P. Duan, J. Jin, W. F. DeGrado, M. Hong, *Journal of the American Chemical Society* 2021, *143*, 7839–7851.
[26] W. Zhang, B. Falcon, A. G. Murzin, J. Fan, R. A. Crowther, M. Goedert, S. H. Scheres, *eLife* 2019, *8*, DOI 10.7554/elife.43584.
[27] Y. K. Al-Hilaly, S. J. Pollack, D. M. Vadukul, F. Citossi, J. E. Rickard, M. Simpson, J. M. D. Storey, C. R. Harrington, C. M. Wischik, L. C. Serpell, *Journal of Molecular Biology* 2017, *429*, 3650–3665.
[28] S. I. A. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, T. P. J. Knowles, *Proceedings of the National Academy of Sciences* 2013, *110*, 9758–9763.
[29] Y. K. Al-Hilaly, B. E. Foster, L. Biasetti, L. Lutter, S. J. Pollack, J. E. Rickard, J. M. D. Storey, C. R. Harrington, W. Xue, C. M. Wischik, L. C. Serpell, *FEBS Letters* 2019, *594*, 944–950.
[30] L. Lutter, Y. K. Al-Hilaly, C. J. Serpell, M. F. Tuite, C. M. Wischik, L. C. Serpell, W.-F. Xue, *Journal of Molecular Biology* 2022, *434*, 167466.
[31] O. Schweers, E. M. Mandelkow, J. Biernat, E. Mandelkow, *Proceedings of the National Academy of Sciences* 1995, *92*, 8463–8467.
[32] J. Stöhr, H. Wu, M. Nick, Y. Wu, M. Bhate, C. Condello, N. Johnson, J. Rodgers, T. Lemmin, S. Acharya, J. Becker, K. Robinson, M. J. S. Kelly, F. Gai, G. Stubbs, S. B. Prusiner, W. F. DeGrado, *Nature Chemistry* 2017, *9*, 874–881.

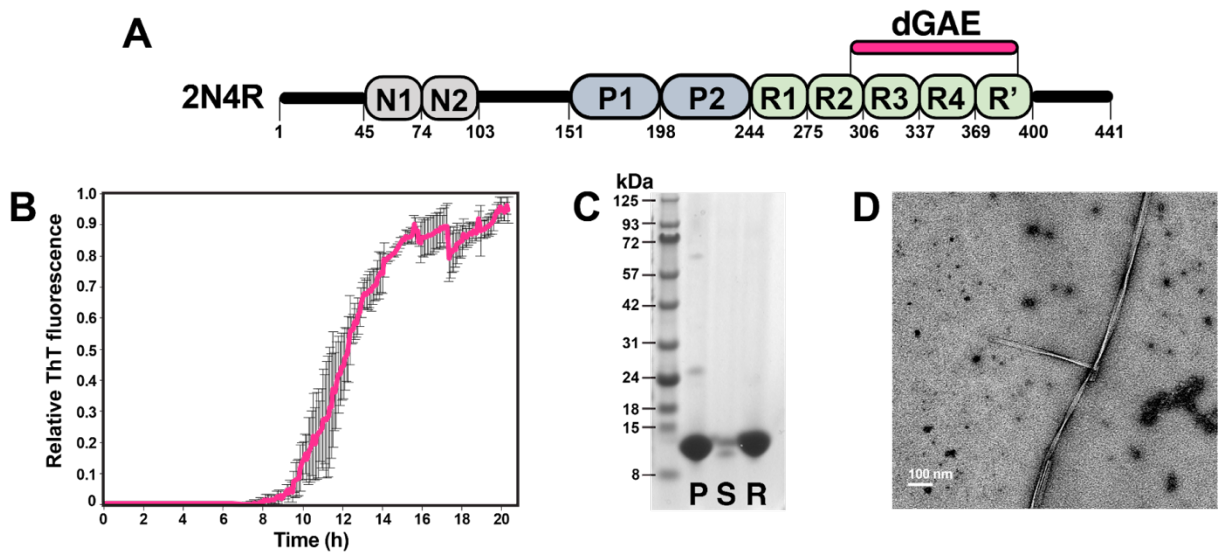

Figure 1. Preparation of homogeneous single-morphology dGAE filaments. (a) Full-length tau 2N4R domain architecture. N1 and N2 are two exon-coded inserts, P1 and P2 indicate the proline-rich regions, and R1 to R' indicate repeat domains. dGAE includes residues I297-E391 from the repeat domains of the 2N4R full-length tau. (b) ThT fluorescence curve of dGAE aggregation kinetics (average of three replicates with standard deviation values). (c) SDS-PAGE gel of dGAE before and after aggregation in PBS pH 7.4 5 mM DTT. Lane P corresponds to the pellet fraction after aggregation, lane S – the supernatant fraction after aggregation shows two distinct bands corresponding to dGAE and a truncated product left after the purification process, and lane R – a reference sample before dGAE aggregation. (d) Representative negative-stain EM micrograph of dGAE(297-391) tau filaments.

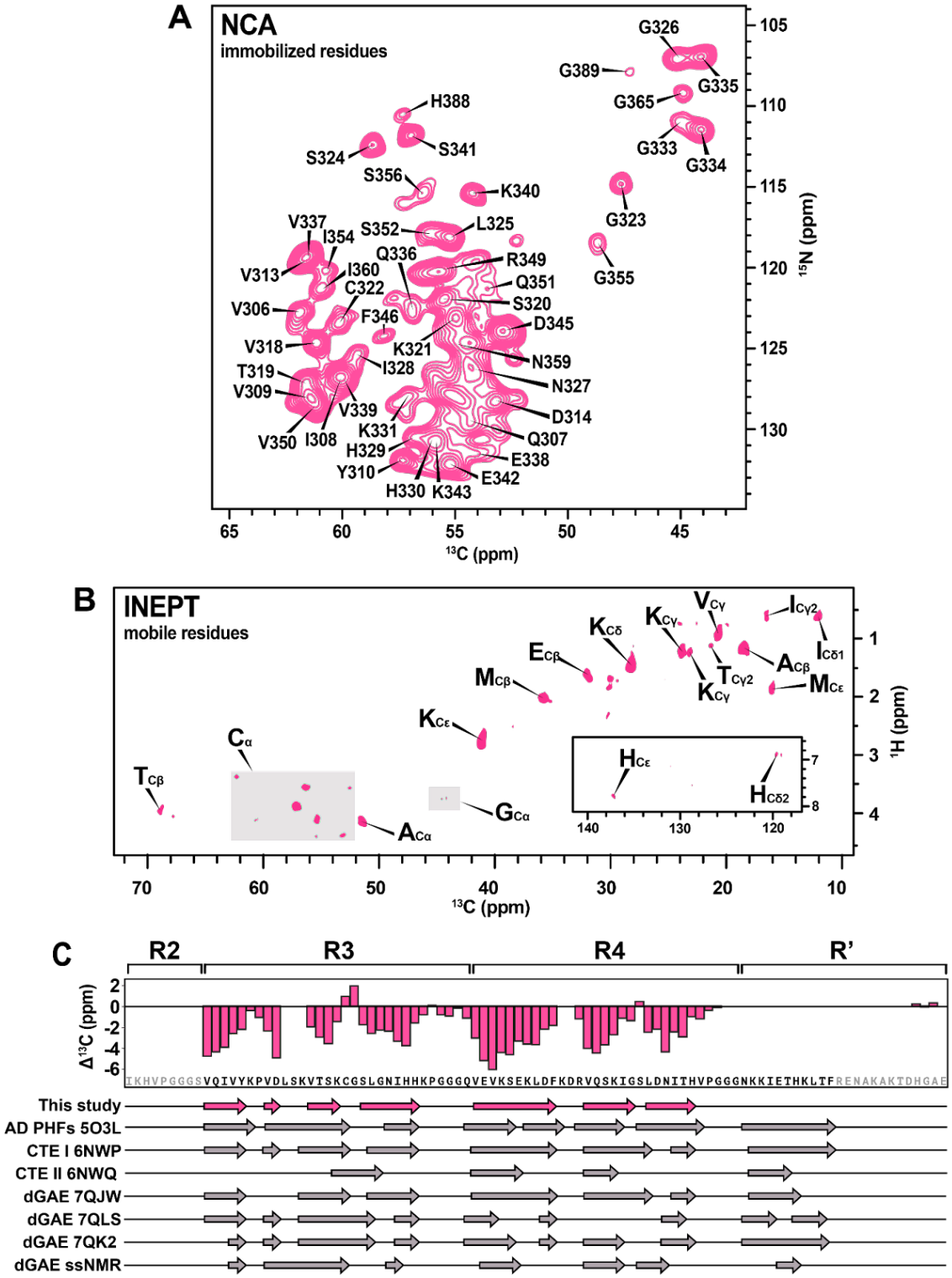

Figure 2. MAS solid-state NMR of dGAE filaments. (a) Assigned 2D NCA spectrum of dGAE filaments. (b) 2D ^{1}H-^{13}C INEPT spectrum of dGAE filaments. (c) Comparison of secondary structures between dGAE filaments presented here, patient-derived tau filaments[4,6], and other *in vitro* dGAE studies[14,27].

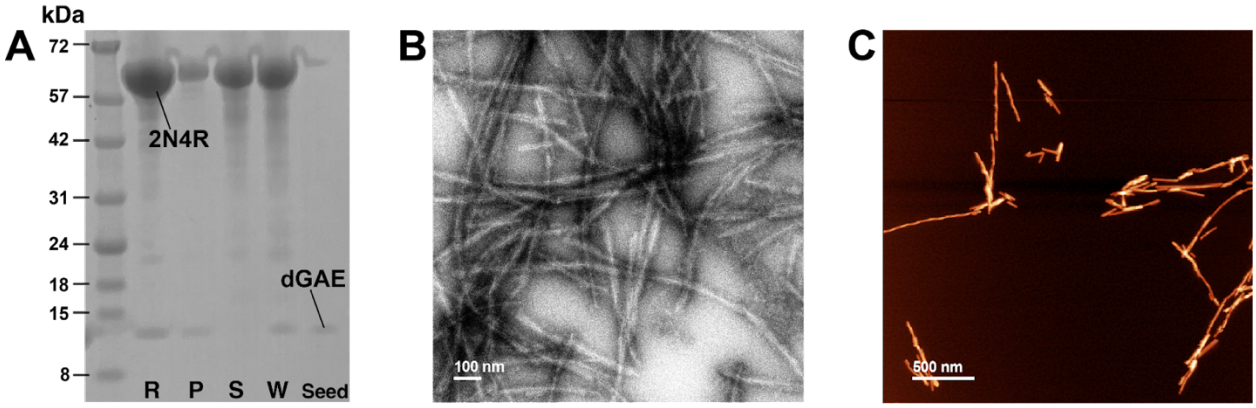

Figure 3. Preparation of dGAE-seeded full-length tau filaments. (a) SDS-PAGE gel of the full-length tau sample with dGAE seeds before and after aggregation in PBS pH 7.4 5 mM DTT. Lane R is a reference sample at the beginning of aggregation, lane P corresponds to the pellet fraction after aggregation, lane S – the supernatant fraction after aggregation, lane W - whole reaction mixture after aggregation, and lane Seed is the dGAE seed. (b) Representative negative-stain EM micrograph of the seeded full-length tau filaments. (c) Representative AFM image of the seeded full-length tau filaments.

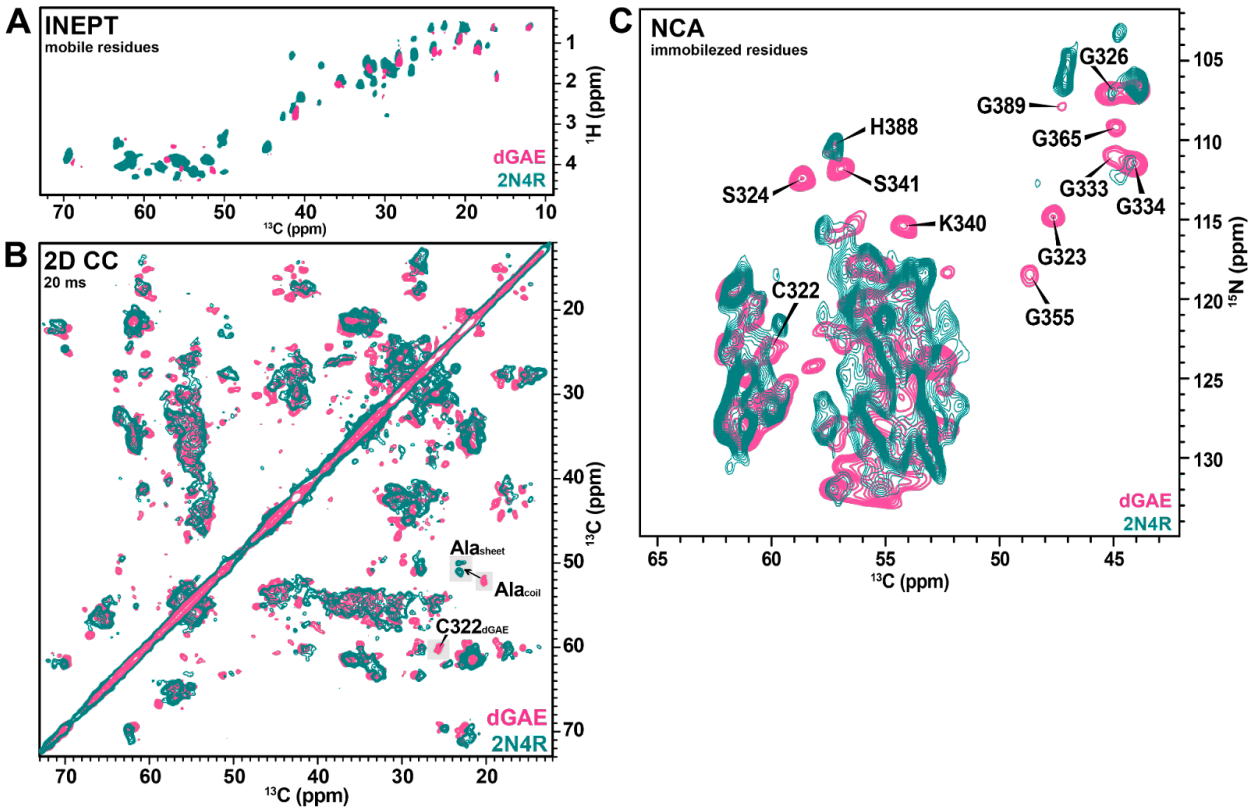

Figure 4. MAS solid-state NMR of seeded full length tau filaments. (a) Overlay of 2D ^{1}H-^{13}C INEPT spectra of dGAE and seeded full-length tau filaments. (b) Overlay of 2D ^{13}C-^{13}C DARR spectra of dGAE and seeded full-length tau filaments. (c) Overlay of 2D NCA spectra of dGAE and seeded full-length tau filaments. The perturbed resonances are identified.