Site-specific C-terminal fluorescent labeling of Tau protein

Louise Bryan¹, Saurabh Awasthi¹S, Yuanjie Li¹, Peter Niraj Nirmalraj², Sandor Balog¹, Jerry Yang³, Michael Mayer¹S

¹ Adolphe Merkle Institute, University of Fribourg, Chemin des Verdiers 4, CH-1700 Fribourg, Switzerland
² Transport at Nanoscale Interfaces Laboratory, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf CH-8600, Switzerland.
³ Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA, 92093-0358 USA

$ Corresponding Author 1: michael.mayer@unifr.ch
$ Corresponding Author 2: saurabh.awasthi@unifr.ch

Formation of Tau protein aggregates in neurons is a pathological hallmark of several neurodegenerative diseases, including Alzheimer's disease. Fluorescently-labeled Tau protein is therefore useful to study the aggregation of these pathological proteins and to identify potential therapeutic targets. Conventionally, cysteine residues are used for labeling Tau proteins, however, the full-length Tau isoform contains two cysteine residues in the microtubule-binding region, which are implicated in Tau aggregation by forming intermolecular disulfide bonds. In order to prevent the fluorescent label from disturbing the microtubule binding region, we developed a strategy to fluorescently label Tau at its C-terminus while leaving cysteine residues unperturbed. We took advantage of a Sortase A-mediated transpeptidation approach to bind a short peptide (GGGH6-Alexa647) with a His-tag and a covalently attached Alexa 647 fluorophore to the C-terminus of Tau. This reaction relies on the presence of a Sortase recognition motif (LPXTG), which we attached to the C-terminus of recombinantly expressed Tau. We determined the possible effects of the resulting C-terminal modification on the secondary structure of Tau protein, its aggregation kinetics, and fibril morphology compared to the unlabeled native Tau protein in vitro. The results showed no significant differences between the native and C-terminally labeled Tau monomer with regard to aggregation kinetics, secondary structure, and morphology.
Keywords: Tau protein, C-terminal labelling, Sortase A, aggregation, Alzheimer’s disease, fluorophore Alexa

Introduction

Tauopathies are a group of diseases affecting the brain caused by the aggregation of microtubule-associated protein Tau.\textsuperscript{1, 2} Alzheimer’s disease (AD) is one of the most common neurodegenerative disorders, characterized by progressive memory loss and cognitive dysfunction and Tau proteins are increasingly implicated in AD as well as other neurodegenerative diseases.\textsuperscript{3, 4} Tau proteins are predominantly found in neuronal cells and are essential for the assembly and stabilization of microtubules.\textsuperscript{5, 6} Human Tau protein exists in six different isoforms ranging in length (from 352 to 441 amino acids) that vary in the number of N-terminal inserts (N1, N2) and contain 3 or 4 imperfect repeats (R1-R4). Repeats R1-R4 correspond to the microtubule-binding region (Figure 1A).\textsuperscript{7, 8} The soluble monomeric form of Tau is thought to be intrinsically disordered and can exist in different conformations.\textsuperscript{9, 10} Numerous post-translational modifications of Tau protein, such as phosphorylation, acetylation, or methylation, regulate its interactions with microtubules or other proteins.\textsuperscript{11, 12} Phosphorylation is one of the most prevalent modifications with up to 85 potential sites on serine, threonine, or tyrosine residues.\textsuperscript{13} Hyper-phosphorylation of Tau results in impaired microtubule binding and leads to misfolding and aggregation, affecting neuronal stability and causing neuronal loss.\textsuperscript{14, 15} The exact aggregation mechanism of Tau and hyper-phosphorylated Tau is still under investigation; however, it is thought to undergo a nucleation-dependent pathway.\textsuperscript{16} Studies have suggested that Tau also undergoes liquid-liquid phase separation, a property of intrinsically disordered proteins (IDPs).\textsuperscript{17} The most toxic species are early aggregates of Tau protein, i.e., soluble or low-n oligomers.\textsuperscript{18, 19} These small aggregates self-assemble into fibrils and finally into neurofibrillary tangles.\textsuperscript{2} High-resolution cryo-electron microscopy (cryo-EM) images of purified Tau aggregates from the brains of AD patients display an ordered core of pairs of protofilaments
comprising regions R3 and R4 (Figure 1B), with the disordered N- and C-termini forming so called fuzzy coat.\textsuperscript{20, 21}

**Figure 1.** Full length Tau protein and strategy for C-terminal labeling. A) Schematic representation of full-length Tau protein (2N4R) colored by domain. N1, N2: N-terminal inserts, PRR: proline rich region, R1-4: imperfect regions, indicates cysteine residues. B) Schematic structure of tau fibril consisting of a core and fuzzy coat. Reprinted by permission from Macmillan Publishers Ltd: copyright 2017.\textsuperscript{22} C) Schematic representation of C-terminal labeling of Tau protein with the enzyme Sortase A. The recombinantly expressed full-length Tau protein is extended on its C-terminus by addition of an LPETGG peptide. Sortase A catalyzes the attachment of a his-tag with a covalently attached Alexa647 fluorophore. D) Top: Coomassie blue staining of an SDS-PAGE gel of Tau protein after a reaction at time 0, 1h, 2h and 3h upon addition of Sortase A and labelled peptide; bottom: fluorescence image of the Tau band. E) Size exclusion chromatography trace (FPLC) of Tau-Alexa647 purification (blue: absorbance at 214 nm, black: absorbance at 647 nm) and SDS-PAGE gel with top: coomassie stain and bottom: fluorescence of Tau-Alexa647.

Single-molecule fluorescence microscopy has the potential to improve our understanding of Tau aggregation. Labeling of Tau protein with a high-quality fluorophore such as the Alexa647 dye used here is essential for such studies. It is, however, important to consider two key factors when designing a method to label Tau protein: First, the presence of two hydrophobic hexapeptide sequences; \textsuperscript{275}VQIINK\textsuperscript{280} (known as PHF6* in R2 and especially \textsuperscript{306}VQIVYK\textsuperscript{311} (PHF6) in R3,\textsuperscript{23} which
are strongly associated with Tau protein aggregation and should not be perturbed.\textsuperscript{24, 25} Second, the presence of two cysteine residues (at position 291 and 322) (\textbf{Figure 1A}), which are involved in intermolecular disulfide bond formation should be preserved. In particular Cys-322 is considered to drive the initial dimerization of Tau protein monomers and aggregation into paired helical filaments.\textsuperscript{26-28} Since modification of these cysteine residues with a fluorophore may influence the aggregation mechanism of Tau protein, we sought a strategy to label the protein on either the N- or C-terminus. Both termini are part of the fuzzy coat and are located away from the core region of Tau fibrils (\textbf{Figure 1B}). To our knowledge, labeled Tau has so far only been prepared using K\textsubscript{18} or shorter Tau isoforms by modification of amine groups or by site-specific mutagenesis of one of the two cysteine residues in the core region of full-length Tau.\textsuperscript{29-33}

Here, we employed a site-specific approach to label the C-terminal end of full-length Tau protein and investigated the effects of modifications on its secondary structure and aggregation kinetics compared to those of native Tau protein \textit{in vitro}. We monitored the aggregation of native and fluorescently labeled Tau over time in the presence of polyanionic inducer heparin\textsuperscript{34} using a ThT assay.\textsuperscript{35} In addition, we compared the morphology of the resulting Tau fibrils using transmission electron microscopy (TEM) and atomic force microscopy (AFM).\textsuperscript{36-39} We observed that the native and C-terminally labeled Tau protein exhibit similar properties with respect to secondary structure, fibril morphology and aggregation kinetics based on analysis by CD Spectroscopy and imaging by TEM and AFM.

\textbf{Results and Discussion}

\textbf{C-terminal labeling of Tau protein}

We explored a Sortase A-mediated labeling strategy to fluorescently label Tau protein while causing minimal perturbation at the core region of the protein in the fibrillar state.\textsuperscript{40-42} The method relies on the addition of a specific recognition motif LPXTG (X being any amino acid) to the C terminus of
recombinantly expressed Tau protein. Sortase A from *Staphylococcus aureus* cleaves off the terminal glycine and forms a thioester with threonine. This step is followed by conjugation to a (G)$_n$ peptide/protein. We thereby used full-length Tau with the addition of the short sequence LPETGG at its C-terminus and enzymatically ligated it to a short peptide bearing a His-tag and a covalently attached fluorescent molecule: GGGH$_6$C-Alexa647 (*Figure 1C*). We chose Alexa647 since it is a high-quality, photostable, far-red fluorescent dye and to prevent spectral overlap with Thioflavin T whose fluorescence can be excited from 385 nm to 450 nm and whose emission maximum ranges from 445 nm to 482 nm.$^{43}$ We added the hexahistidine sequence to the peptide to provide an additional tag for efficient separation of the labeled from the unlabelled protein and to concentrate labeled Tau species since the Sortase A labeling reaction typically provides labeling efficiencies of 20 to 90%; largely dependent on the Sortase A variant.$^{44}$ We used the commercially available Sortase pentamutant (Sortase 5A, with a His-tag), which increases the rate and efficiency of Sortase labeling compared to the wild type Sortase A.$^{45}$ The labeling reaction was carried out at 10°C to prevent degradation of the Tau protein or the formation of small aggregates. The reaction mixture was then purified on a Ni-NTA column where the labeled protein, GGH$_6$-Alexa647 and Sortase A, all containing His-tags were retained on the column. This affinity chromatography was followed by size exclusion chromatography to purify the labeled protein from the remaining molecules in the mixture (*Figure 1D and 1E*). This procedure purified C-terminal labeled Tau protein for further structural characterization of the monomeric form of Tau and its aggregates.

*Secondary structure and aggregation kinetics of labeled and native full-length Tau*

We determined the secondary structure of monomeric Tau-A647 and native full-length tau protein by CD spectroscopy. *Figure 2A* shows a minimum at 200 nm for both native Tau and Tau-A647, implicating that they are unstructured in solution and the addition of a short peptide sequence with a fluorescent dye on the C-terminus did not change the secondary structure. Furthermore, we
analyzed the CD spectra and determine the propensity of different secondary structural elements such as α-helix, β-sheet, and irregular or disordered state for native and labeled Tau protein monomers. Table 1 reveals a similar secondary structure in both the variants (i.e. native and labeled), suggesting a minimal effect of C-terminal labelling on Tau protein structure in solution (Figure 2A).

![CD spectra](image)

**Figure 2.** Comparison of the monomer secondary structure of native and labeled Tau in solution as well as their aggregation kinetics. A) CD spectra of native (black), and labelled Tau (red) in their predominantly monomeric form. The inset shows the percentage of different secondary structures for native and labeled Tau as determined using the online server CAPITO. B) ThT fluorescence intensity measurements to compare the aggregation kinetics of native and labeled Tau protein. The data points are the average of three independent measurements. The error bars show the standard deviation from three independent measurements.
We also carried out a comparative analysis of aggregation kinetics of labeled and native full-length Tau. In order to follow their aggregation kinetics, each Tau variant was incubated with the inducer, i.e., heparin (a Tau:heparin ratio of 4:1) and ThT at 37°C in Tau aggregation buffer and their fluorescence spectrum was measured at regular time points (Figure 2B). Overall, we observed a similar trend in the case of native and labelled Tau protein with the ThT fluorescence increasing overtime corresponding to the formation of aggregated species. We estimate the lag time ($t_{lag}$) of aggregation using Equation 1:

$$ Y = y_i + m_i t + \frac{y_f + m_f t}{1 + e^{-[(t-t_{1/2})/\tau]}} $$

where, $Y$ is the fluorescence intensity as a function of time $t$, $y_i$ and $y_f$ are the intercept of the initial and final fluorescence values with the y-axis, $m_i$ and $m_f$ are slopes of the initial and final baselines, $t_{1/2}$ is the time needed to reach halfway through the elongation phase and $\tau$ is the elongation time constant. The apparent rate constant, $k_{app}$, for the growth of fibrils is given by 1/s, and the lag time is usually defined as $t_{lag} = t_{1/2} - 2\tau$. This analysis revealed that $t_{lag}$ was 16.6 h for native Tau and 13.9 h for labeled Tau.

<table>
<thead>
<tr>
<th>Protein variant</th>
<th>α-Helix (%)</th>
<th>β-Sheet (%)</th>
<th>Irregular (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Tau</td>
<td>3.4 ±2.5</td>
<td>44.0 ±3.6</td>
<td>60.0 ±3.6</td>
</tr>
<tr>
<td>Labeled Tau</td>
<td>4.7 ±4.2</td>
<td>42.7 ±7.3</td>
<td>57.0 ±2.6</td>
</tr>
</tbody>
</table>

**Labeled and native Tau exhibit similar fibril morphology**

In order to compare the fibril morphology, we carried out TEM and AFM imaging of aggregates formed by native and labeled full-length Tau protein. Figure 3 shows that TEM imaging of Tau fibrils
after 7 and 14 days of aggregation *in vitro* induced by heparin revealed similar fibril morphology. From these images, we determined the width of fibrillar aggregates. Both the native full length-Tau protein and labeled Tau-A647 exhibit a similar fibril width estimates of $14.55 \pm 1.8 \text{ nm}$ and $14.02 \pm 1.7 \text{ nm}$, respectively (*Figure 3C and 3F*). Earlier studies have shown that Tau fibrils extracted from brain samples exhibit a width of $19 \text{ nm}$, whereas in the case of heparin-induced Tau protein fibrils they were around $16.9 \text{ nm}$. More recently, Zhang *et al* showed heparin induced Tau fibrils are in fact highly heterogeneous with a width ranging from 4 to 25 nm. Our estimates of fibrillar width for both the native and labeled Tau protein are in good agreement with these earlier reports.

![Figure 3. Transmission electron microscopy (TEM) imaging and characterization of fibril width. A) and B) TEM images of native Tau protein aggregates incubated for 7 days and 14 days at 37°C under shaking conditions, in the presence heparin. D) and E) TEM images of labelled Tau protein incubated for 7 days and 14 days respectively. (Scale bar: 100 nm). C) and F) Histograms of the fibril width after 14 days of aggregation.](image)

Furthermore, we characterized the fibrils formed after 7 days of incubation by labeled and native Tau protein using liquid-based AFM to estimate their height (*Figure 4 A-F*). Histograms of the height of individual fibrils revealed a height range of $5.0 \pm 1.1 \text{ nm}$ for native Tau and $5.6 \pm 1.2 \text{ nm}$ for Tau-
A647 (Figure 4F). In general, the high-resolution AFM images do not show any drastic differences in topography, morphological variations along the fibril, height profile or alignment between the native and labelled Tau fibrils.

Figure 4. Atomic force microscopy (AFM) imaging and characterization of the height of Tau fibrils. (A) AFM image of native Tau fibrils. (B) AFM image showing the topography of labeled Tau fibrils. Height distribution (based on individual cross-sectional analyses of the AFM data obtained on native (C) and labelled (D) topographic images) of native Tau fibrils (black histogram) and labelled Tau fibrils (red histogram). A mean fibril height of 5.0±1.1 nm was obtained for the native Tau fibrils and a mean fibril height of 5.6 ± 1.2 nm was calculated for the labelled Tau fibrils from the height distribution plots.

Conclusion

Using Sortase A-mediated labeling this work demonstrates that: 1) C-terminal labeling of Tau protein has minimal effect on secondary structure in solution, 2) both the labelled (i.e. Tau-647) and native-full length Tau protein exhibit similar aggregation kinetics, and 3) both the labelled and native Tau protein exhibit similar fibrillar morphology. We propose that the C-terminal labeling strategy of Tau protein presented here may therefore be useful for studies of Tau aggregation using (single-
molecule) fluorescence methods with minimal effects on the structure of the native protein conformation.

Acknowledgement

SA acknowledges financial support by the Swiss National Science Foundation’s “SPARK” funding (CRSK-3_195960). We acknowledge support from the National Institute on Aging of the National Institutes of Health under Award Numbers RF1AG062632 (JY) and RF1AG077802 (JY and MM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. MM acknowledges financial support from the Swiss National Science Foundation (Grant number: 200020_197239) and from the Adolphe Merkle Foundation, Switzerland.

Materials and Methods

Tau labeling

The pET30a plasmid containing GST-Tau-LPETGG-H6 was synthesized by GenScript. The protein was expressed by Oscar Vadas at the Protein and Peptides platform, University of Geneva, in E. Coli followed by purification. For labeling, 50 µM of purified Tau-LPETGG was mixed with 150 µM of the peptide GGGH_6-Alexa647 (Bio-Synthesis Inc.) and 3 µM Sortase A5 (Active Motif) in 20 mM Tris, 150 mM NaCl, 10 mM CaCl_2, 1 mM DTT pH 7.4. The mixture was incubated at 10°C for 3h and then loaded onto a His Spintrap column (GE Healthcare). After 1-hour incubation, the column was washed with 10 mM Tris, pH 7.8 100 mM NaCl, 10 mM imidazole followed by elution with 10 mM Tris, pH 7.8, 100 mM NaCl, 250 mM imidazole. The elute was concentrated using a vivaspin 500, 3000 MWCO centrifugal concentrator (Sartorius, Germany) and then purified on a Superdex200 increase 10/300 column (GE Healthcare) in 10 mM Tris, 100 mM NaCl, 1 mM DTT; pH 7.4 using a FPLC AKTA system (GE Healthcare). The fractions containing the purified labelled Tau protein were pooled,
concentrated with a vivaspin 500, 3000 MWCO centrifugal concentrator, aliquoted, flash frozen in liquid N\textsubscript{2} and stored at -80°C. At each step, samples were taken, mixed with SDS-PAGE loading buffer and loaded onto an SDS-polyacrylamide gel, for 40 min run at 150 V in order to verify the presence of the protein.

**ThT assays**

The aggregation of native full-length Tau procured from Eurogentec and Tau-LPETGGGH\textsubscript{6}C-Alexa\textsubscript{647} henceforth called Tau-A\textsubscript{647} was monitored by measuring fluorescence spectra of ThT over time using a fluorometer from Horiba Scientific. For the Tau aggregation assay, a solution of 2 µM Tau was mixed with 0.5 µM heparin sodium salt (MW: 6’000-3’0000 g/mol), 1 µM of ThT in Tau aggregation buffer (25 mM sodium phosphate, 25 mM NaCl, 2.5 mM EDTA, 0.33 mM DTT, pH 6.8). The aggregation mix was incubated at 37°C, under constant agitation with 300 rpm in protein Lobind tubes (Eppendorf). Fluorescence emission spectra were measured at different time points (0h, 1h, 24h, 48h, 72h, 96h, 192h and 216h) by using an excitation wavelength of 412 nm and a slit of 5 nm in black quartz cuvettes (Hellma).

**CD spectroscopy**

CD spectra were acquired on a Jasco J-810 CD spectropolarimeter using 0.1 cm quartz cuvettes (Hellma), operated at 20°C. To minimize buffer absorption, Tau samples were dialysed overnight against 10 mM sodium phosphate buffer, pH 7.8 in Slide-A-Lyzer MINI dialysis devices (100 ul). CD spectra were recorded from 190-260 nm at a scan speed of 20 nm/min and an increment of 1 nm. Four scans were recorded for each sample. The buffer was used as a blank for background subtraction.
AFM

A total of 5 μl solution containing fibrils of native Tau or Tau-LPETGGGH$_6$C-Alexa$_{647}$ both at an equivalent monomer concentration of 100 nM in a buffer containing HEPES pH 7.4 was pipetted onto Si (111), n-type substrate. The sample was left to settle for 30 min followed by injecting 10 μl of pure water to rinse the buffer salt solution. Atomic force microscope images were acquired in tapping mode in pure water with a Bruker Multimode 8, E-scanner. The tip used was a Scanasyst-AIR, 0.4 N/m, 70 kHz (Bruker AFM probes). Images were recorded at a resolution of 1024 by 1024 pixels at a scan rate of 1 Hz. Analysis of AFM data was conducted using Bruker Nanoscope Analysis software.

Electron Microscopy

Transmission electron microscopy images were recorded with a FEI Tecnai Spirit operating at 120 k. Carbon coated 300-mesh copper grids (Electron Microscopy Sciences, Hatfield) were plasma cleaned for 5 s using an oxygen plasma cleaner (Zepto RIE, Dienner), 5 μl of Tau protein sample in aggregation buffer (diluted 100 times) was pipetted on top of the grids and incubated for 2 min, the grids were then washed in a water droplet. Uranyl acetate (2% w/v) was added (3ul) and incubated for 2 min. Excess stain was blotted off with a filter paper and dried. We used ImageJ to estimate the width of Tau protein fibrils.$^{50}$
References


(35) Ramachandran, G.; Udgaonkar, J. B. Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by Tau protein. *The Journal of biological chemistry 2011*, *286* (45), 38948-38959. DOI: 10.1074/jbc.M111.271874


