# Structural Evaluations of tau Protein Conformation: Methodologies and Approaches

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Biochemistry and Cell Biology</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>bcb-2016-0227.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Mini Review</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>10-Feb-2017</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Zabik, Nicole; Oakland University, Chemistry Imhof, Matthew; Oakland University, Chemistry Martic, Sanela; Oakland University, Chemistry</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue?:</td>
<td></td>
</tr>
<tr>
<td>Keyword:</td>
<td>tau protein, intrinsically disordered proteins, spectroscopy, computational methods, post-translational modifications</td>
</tr>
</tbody>
</table>

[https://mc06.manuscriptcentral.com/bcb-pubs](https://mc06.manuscriptcentral.com/bcb-pubs)
Structural Evaluations of tau Protein Conformation: Methodologies and Approaches

Nicole L. Zabik, Matthew M. Imhof, and Sanela Martic-Milne*
Department of Chemistry, Oakland University, Rochester, MI 48309

ABSTRACT

Protein-misfolding diseases are based on a common principle of aggregation initiated by intra- and intermolecular contacts. The structural and conformational changes, induced by biochemical transformations, such as post-translational modifications (PTMs), often lead to protein unfolding and misfolding. Thus, these order-to-disorder or disorder-to-order transitions may regulate cellular function. Tau, a neuronal protein, regulates microtubule (MT) structure and overall cellular integrity. However, misfolded tau modified by PTMs results in MT destabilization, toxic tau aggregate formation, and ultimately cell death, leading to neurodegeneration. Currently, the lack of structural information surrounding tau severely limits understanding of neurodegeneration. This mini-review focuses on the current methodologies and approaches aimed at probing tau conformation and its role in various aspects of tau biochemistry. The recent applications of nuclear magnetic resonance, mass spectrometry, Förster resonance electron transfer, and molecular dynamics simulation toward structural analysis of conformational landscapes of tau will be described. The strategies developed for structural evaluation of tau may significantly improve our understanding of misfolding diseases.

KEYWORDS: tau protein, hyperphosphorylation, NMR, FRET, MS, molecular dynamics, intrinsically disordered proteins
1. INTRODUCTION

Intrinsically disordered proteins (IDPs) lack a well-defined three-dimensional structure, yet retain a flexible conformation that is important to their role in cellular processes (Wright et al. 2015). IDPs undergo order-to-disorder or disorder-to-order transitions and are commonly present in humans and bacteria. Their structure and function may be modulated by protein chaperones, post-translational modifications (PTMs), and degradation processes (Martin et al. 2016; Uversky et al. 2014). Misregulation of modulators leads to dysfunction of the protein, most commonly resulting in unfolding, misfolding, aggregation, or loss of function (Goldberg et al. 2003). Tau protein, a microtubule-associated protein linked to neurodegeneration, is recognized as an IDP. Tau is a highly disordered soluble protein and remains a challenging protein target for structural analysis. Therefore, the lack of structural information about tau limits the progress in neurodegeneration research and development of effective therapeutic strategies.

Six isoforms of tau exist and differ in primary sequence: variation in N-terminal exons (N) and microtubule binding repeat regions (R). Figure 1 depicts a schematic of various non-phosphorylated tau (nTau) isoforms and illustrates N-terminal exons (N), Pro-rich domain (PRD), and R repeats in the primary sequence. In vivo, all 6 tau isoforms coexist and variation in their relative amounts has pathological consequences (Mukrasch et al. 2005). In neurons, tau promotes and stabilizes microtubules (MT), and facilitates the cellular function. The PTMs of tau modulate tau-MT interactions leading to MT destabilization and cell death. In turn, tau aggregates into oligomers, proto-fibrils, fibrils, tangles, paired helical filaments (PHFs), and neurofibrillary tangles (NFTs).
Figure 1. Schematic illustration of 6 tau isoforms.

Soluble oligomers, rather than insoluble tangles, are thought to be tau’s most neurotoxic conformations (Walsh et al. 2002). However, the insoluble aggregates may serve a protective role by sequestering soluble oligomers (Santacruz et al. 2005).

The longest isoform of tau (2N4R) is hydrophilic with an overall basic character, a predominantly acidic N-terminus (including exons), and neutral C-terminus (~40 residues). The middle region, PRD (150-240), contains numerous Pro residues, such as SerPro or ThrPro motifs, which are targets for Pro-directed kinases (Figure 2). In addition, PRD contains seven Pro-X-X-Pro motifs which are binding sites for proteins with SH3 domains. The entire tau molecule may be considered natively unfolded or intrinsically disordered, but it has a low content of secondary structures which include α-helices (114-123 and 428-437), β-strands (86-92, 161-166, 224-230, 274-284, 305-315, and 336-345), and poly-proline II helices (175-184, 216-223, and 232-239). These hints of secondary structure are largely transient. However, the
electrostatic charge compensation of basic residues and high propensity for β-sheet formation are essential for tau assembly. The PTMs of tau modulate electrostatic charges and accessibility of peptide sequences prone to β-sheet formation.

PTM is a covalent functionalization of amino acid residue which regulates cell signaling and may dramatically modulate protein conformation. The normal structure and function of tau requires a delicate balance between all PTMs (Bah et al. 2015; Metskas et al. 2015). Various PTMs of tau protein have been identified and include phosphorylation, acetylation, glycosylation, glycation, prolyl-isomerization, truncation, nitration, polyamination, ubiquitination, sumoylation, and oxidation (Martin et al. 2011). Acetylation at Lys residues regulates tau aggregation \textit{in vitro}. Tau glycosylation induces tangle formation, and reduces tau’s accessibility to MT. Tau glycosylation also decreases its phosphorylation. Glycation of Lys residues promotes tau polymerization and stabilizes aggregated tau protein. Prolyl-isomerization has a dramatic impact on tau conformation, specifically, pThr231 regulates \textit{cis} to \textit{trans} isomerization of tau. The efficient prolyl-isomerization of phosphorylated tau (pTau), when mediated by Pin1, promotes dephosphorylation by phosphatases. Truncation of tau by caspases at Glu391 and Asp421 promotes tau aggregation. Nitration of tau occurs at Tyr residues and decreases tau’s ability to promote tubulin assembly, further leading to aggregation. Polyamination of tau cross-links Gln and Lys residues and may be involved in stabilization and formation of NFTs. Ubiquitination of tau at C-terminus promotes proteolytic degradation and occurs after a hyperphosphorylation event. Sumoylation at Lys340 counteracts the ubiquitination. Tau oxidation may occur at Cys322, within R3 domain, leading to PHF formation. Among all of PTMs, hyperphosphorylation is the most biologically relevant PTM of tau and has been correlated with neurodegenerative diseases. \textit{In vivo}, tau is partially
phosphorylated (~5 phospho-sites) by protein kinases. Three types of phosphorylation of tau exist: 1) SerPro/ThrPro motif in the PRD phosphorylated by proline-directed protein kinases, such as GSK-3β, 2) LysXGlySer motif in the R repeat region phosphorylated by non-proline-directed kinases such as Microtubule-affinity regulating kinase (MARK), and 3) Tyr residues targeted by sarcoma-related kinases, such as Fyn (Drewes et al. 1997; Iqbal et al. 2016; Lee et al. 2004). Pathological, phosphorylated tau can be phosphorylated at over 30 sites by all three types of kinases (Mair et al. 2016). Figure 2 illustrates some of the phosphorylation sites of tau introduced by GSK-3β, MARK and Fyn protein kinases, as representative examples.

**Figure 2.** Schematic illustration of specific peptide sequences and key phosphorylation sites introduced by GSK-3β, MARK and Fyn protein kinases.

The introduction of negative charges through phosphorylation modulates inter- and intramolecular electrostatic interactions between amino acid residues, such as salt-bridges.
charge imbalance of pTau may lead to significant conformational changes, unfolding, or misfolding. These structural transitions have detrimental effects on tau-MT interactions, tau’s aggregation propensities, and cytotoxicity of tau aggregates.

The basic residues of tau within the microtubule binding region (MTBR) (Figure 2) are involved in electrostatic interactions with Glu residues of MTs. Figure 2 shows that the MTBR domain is composed of R1-R4 repeats as well as several key phosphorylation sites. When tau becomes post-translationally modified, such as phosphorylated (pTau), the key salt-bridges between tau and MT are compromised. In addition, the electrostatic imbalance may cause structural changes in tau, reducing accessibility of its binding domain to MTs. The outcome of tau being hyperphosphorylated is detachment from MT, as depicted in Figure 3. For example, the phosphorylation of residues in the R repeats and PRD of tau can modulate MT binding (Brandt et al. 1994; Lin et al. 2007; Mandelkow et al. 2004).

Figure 3. Schematic of α/β-tubulin subunits within MT and tau prior to and post PTM. PTMs lead to tau modifications and MT depolymerization.
If the polyproline II helix is stabilized by phosphorylation within the PRD, then this specific conformation may inhibit binding to MTs as well. As a consequence, MTs are destabilized, leading to loss of neuronal cell function and cell death, i.e. neurodegeneration.

Following PTM of tau and its detachment from MT, tau is prone to misfolding and aggregation. A current hypothesis is that hyperphosphorylation leads to aggregation (Köpke et al. 1993). Partially phosphorylated normal tau is not prone to aggregation, unlike hyperphosphorylated tau. How hyperphosphorylation modulates electrostatic interactions and affects tau conformation, unfolding or misfolding, is not well characterized. Such transient structures may have hydrophobic domains prone to β-sheet formation and aggregation. The repeat domains of tau (R1-R4) (Figure 2) are mostly prone to aggregation via β-sheet formation and make up the self-assembled PHFs of tau (Akoury et al. 2016). The in vitro aggregation of nTau, induced by anionic agents, has been well characterized (Chirita et al. 2003). However, the aggregation mechanism of pTau remains unclear.

Recent tau conformation models have used pseudo-phosphorylated mutants, containing Asp or Glu residues in lieu of phosphosites, as mimics to probe aggregation of pTau (Combs et al. 2011; Leger et al. 1997; Sun et al. 2009). This approach is most often used to elucidate the role of negative charges upon binding to other proteins, like tubulin. The advantage of pseudophosphorylation mimics is having precise control over type and number of phospho-sites, as opposed to normal kinase-catalyzed phosphorylations. However, pseudophosphorylation has propensity to change native protein conformation by a simple amino acid substitution, therefore structural contributions stemming from the site-mutagenesis and pseudophosphorylation must be carefully considered. Ideally, site-mutagenesis should have minimal effect on native conformation, which will be dependent on amino acid residue chosen.
In addition, the conformations of nTau, pTau, and aggregated tau may be probed by conformation-specific antibodies (Combs et al. 2016; Lasagna-Reeves et al. 2012). The T22 tau antibody (epitope including Thr231) has previously shown binding specificity for oligomeric, over monomeric, tau and may be used to identify oligomers generated by PTMs. Additionally, by using TNT1 antibody specific to N-terminal domain, the conformation of oligomers and filaments of pathological tau was evaluated. The N-terminal region was highly accessible to TNT1 antibody (7-12) in pathological oligomers but inaccessible in filaments. This type of biochemical analysis relies on the use of antibodies which may have some cross-reactivity with other proteins.

While it is clear that PTMs trigger tau aggregation, the relationship between hyperphosphorylation and conformational outcome of tau remains elusive. The identification of phosphorylation patterns and unique conformers will allow for mechanistic insight into protein unfolding, misfolding, and aggregation. Currently, the state of tau research is challenged by a lack of methodological platforms which would allow for structural characterization of PTM-induced transformations. In this mini-review, we will discuss how current groundbreaking experimental and computational methodologies and approaches have been used towards structural evaluation of tau protein and its PTMs, specifically phosphorylation, moving towards a better understanding of tau biochemistry. Figure 4 is the conceptual preview of methods and approaches, described herein, used for evaluation of tau protein phosphorylation, conformation, and aggregation.
2. NUCLEAR MAGNETIC RESONANCE

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical technique which utilizes magnetic properties of particular atomic nuclei to obtain structural information about molecules. Not only can structural information be ascertained by use of NMR, but its non-destructive nature allows for further sample analysis by other methods. Given its ability to characterize protein ensembles, including secondary structures and long range intramolecular interactions, NMR has proven useful for studying IDPs (Dyson et al. 2004; Haba et al. 2013; Jensen et al. 2013). NMR has previously been used to characterize IDPs, such as nonphosphorylated and phosphorylated Sic1 found in yeast, resulting in a greater understanding of Sic1 structure and function (Mittag et al. 2010). Recently, NMR studies identified structural importance of Pro residues in hyperphosphorylated IDPs, as well. The presence of Pro residues
has an expanding effect on IDP conformational propensity, allowing for ease of access by modifying enzymes (Martin et al. 2016). NMR capabilities were also applied to tau protein and its peptides in an effort to understand tau structure.

Solution- and Solid-state NMR have been used with both tau peptide fragments and full-length tau protein. The peptide model studies allow for investigation of specific domains and identification of aggregation prone regions of nTau. For example, solid-state NMR studies were used with K18 (243-372) and K19 (243-372, missing R2) (Figure 2) peptide fragments in order to determine the role of R2 repeat in aggregation. With K19, NMR data revealed the presence of three β-strands in the MTBR region of tau, showing similar features to that of PHFs found in Alzheimer’s diseased brains (Daebel et al. 2012; Mukrasch et al. 2005). Solid-state NMR studies of K18 fragment also demonstrated a presence of aggregation prone β-structures in all four repeats (Ozenne et al. 2012). These findings indicate that the presence or absence of R2 has negligible effect on tau aggregation, and that tau peptide models like K18 and K19 are good mimics of biologically relevant tau aggregates. Further studies on the K18 fragment were performed using solution-state NMR in order to determine structural changes as a result of anion binding to the positively charged tau R domain (Akoury et al. 2016). Anionic cofactors, such as polyglutamic acid, heparin, RNA, and arachidonic acid, have been associated with PHFs and NFTs, and are commonly used for inducing tau aggregation in vitro. Anion binds to positively charged motifs (Lys-Lys, Lys-His, and His-His-Lys) at the end of each R repeat (Mukrasch et al. 2005). Each repeat contains a binding site for MT as well. The binding sites are located at equivalent positions in each repeat, specifically at the N-terminus to the PGGG motif. The use of polyglutamic acid over heparin is favorable because it has shown the ability to delay tau conversion into oligomers. This delay in aggregation allows for analysis of structural changes
that occur upon binding of polyglutamic acid to R domains. Data from this study suggest intramolecular compaction of aggregation-prone hexapeptide sequences \( ^{275}\text{VQIINK}^{280} \) and \( ^{306}\text{VQIVYK}^{311} \) of tau (Figure 2) (Akoury et al. 2016). Further, delay of tau aggregation occurs as a result of intramolecular crosslinking between native Cys291 and Cys322 residues. While tau fragments are used for conformational studies, they do not reflect the complexity of tau protein.

The structure of 2N4R was determined by solution-state NMR (Andronesi et al. 2008; Mukrasch et al. 2009). Figure 5 displays the ensemble of structures populated by 2N4R in compact conformations. The C- and N-termini were shown to be in proximity to one another over the repeat domains of tau, forming a ‘paperclip’ conformation, which is in contrast to the previous idea that tau conformation in solution is an extended state. The ‘paperclip’ conformation of tau in solution shows many long range interactions. The major points of contact include: 1) C- and N-termini interacting with the R3 repeat, 2) transient helical structure within C-terminal sequence, and 3) interaction of a helical structured region of the N-terminus with R1 and R2 repeats. Proximity of PRD P2 region to R1, R2, R4 and C-terminal helix may suggest that phosphorylation at PRD P2 may modify tau conformation and its propensity to aggregate into PHF.

In addition to nTau, NMR evaluations of pTau were also reported. NMR may be used for detection of phosphorylation, identification of phosposites, and screening of phosphorylation inhibitors (Landrieu et al. 2006).
Figure 5. Proposed solution structure of 2N4R determined by solution-state NMR. α-Helical structures H1 (114-123) and H2 (428-437) and β-structures B2 (274-284), B3 (305-315), and B4 (336-345) are displayed in gray and light green, respectively. PPP (175-184), PP (216-223) and P (232-239) are polyproline II structures shown in purple. N and C represent the respective termini of tau (Mukrasch et al. 2009).

By comparing NMR spectra for nTau and pTau, the structural differences induced by phosphorylation may be determined. Typically, phosphorylating Ser and Thr residues results in their deshielding, as well as deshielding of residues in close proximity to phosphosites (Bienkiewicz et al. 1999). However, phosphorylation at Thr231 has demonstrated a shielding effect of up to 10 neighboring amides; this is due to stabilization of an α-helical structure present in this domain (Sibille et al. 2012). Additionally, phosphorylation of Ser235, Ser237 and Ser238 further stabilized a transient helix between Ser238 and Arg242 (Schwalbe et al. 2015). Phosphorylation at Thr231 also results in salt-bridge formation with positively charged amino
acids, such as Arg230. The salt-bridges are critical intermolecular point of contacts between tau and tubulin, therefore the pThr231-Arg230 salt-bridge may compete with electrostatic interactions between tau and tubulin. Data indicate that phosphorylation events may modulate tau-MT interactions and MT stability. Thus, NMR may also be used to probe conformational outcomes stemming from site-specific phosphorylations.

NMR of the pseudophosphorylated tau mutants Ser199Glu/Ser202Glu/Thr205Glu, Thr212Glu/Ser214Glu, and Ser396Glu/Ser404Glu have shown that pseudophosphorylation weakens transient tau folding, due to reduced paramagnetic effects from the flanking regions (Bibow et al. 2011). Pseudophosphorylated mutants are characterized by the reduced proximity of PRD to N-terminus. In addition, there is a reduction in compaction of the C-terminus. Specifically, pseudophosphorylations in the C-terminus induce the most dynamic fluidity compared to other domains.

Biological relevance of tau interaction with the MT building block, tubulin, may also be monitored using solution-state NMR. It has been shown that tau interactions with MT induce chemical shift changes for residues in the PRD and R1-R4 domains of 2N4R (Mukrasch et al. 2009). In addition, NMR was used to evaluate tubulin dimer binding to TauF4 (208-324) fragment (Figure 2). This biomolecular interaction produced minimal change in tau peptide conformation (Figure 6), except upon dimerization of tubulin subunits (Gigant et al. 2014). The U-turn formation at R1 repeat of TauF4 fragment promoted vertical tubulin dimer aggregation, as depicted in Figure 6. Upon tubulin dimerization, the C-terminus of TauF4 fragment assumes an extended conformation, allowing for an additional MT growth. Given the success of NMR studies between tubulin and tau, similar strategies may be applied to tubulin interactions with pTau.
**Figure 6.** Binding model based on NMR studies showing interactions between TauF4 fragment with tubulin heterodimers. In U-turn conformation, TauF4 promotes tubulin dimerization. Further MT growth requires extended TauF4 conformation.

Recent studies have identified other structural factors, such as Pro isomerization, as contributors to conformational change of tau. For example, predominance of the *cis* conformation of pThr231-Pro232 in biological samples is proposed to be a biomarker of early onset neurodegeneration (Nakamura et al. 2012). However, solution-state NMR studies of TauF4 peptide fragment have shown the presence of predominantly *trans* conformation associated with most prolyl bonds, including the pThr231-Pro232 bond (Ahuja et al. 2016). The discrepancy observed in these two studies may be explained by differing experimental models and will thus require further probing in order to determine consequences of *cis/trans* isomerization of pathogenic tau prolyl bonds.

Clearly, NMR is a powerful tool for 1) identification of PTMs of tau, 2) evaluation of structural role PTMs play in tau conformation, and 3) binding studies of tau to biologically
relevant ligands. In combination, 2D NMR, NMR-derived distance constraints, residual dipolar couplings, and chemical shift restraints may be used in molecular ensemble calculations of various pTau isoforms.

3. MASS SPECTROMETRY

Mass spectrometry (MS) methods have been used to identify pTau sites, but have rarely been applied to conformational studies.

Recently, the phase-hydrogen deuterium exchange (HDX) was coupled to MS in order to determine gas-phase structures of proteins. HDX involves labeling of solvent exposed acidic amino acids with deuterons from the surrounding gas. The gas-phase protein structural analysis has shown that while proteins exist in specific folds in solution, they may exist as a heterogeneous population in the gas phase. To simplify the HDX analysis, the coupling to Differential Mobility Spectrometry (DMS) was established. DMS is an ion mobility strategy that allows for pre-selection of ions based on their high- and low-field mobilities. In an ideal case, the DMS coupled to HDX would allow for determining the liquid-phase structure of a particular protein and evaluating global structure of protein in solution. DMS-HDX was used to determine charge distribution (800-2400 m/z) and compensation voltages of nTau and pTau (Zhu et al. 2016). While it was found that nTau existed in an unfolded state, pTau existed in a mix of states. DMS-HDX showed a constant Relative Deuterium Uptake (0.9-1.0 range) for nTau but greater variation for pTau (0.6-1.0 range). Results indicate that following PTM, proteins may exhibit large differences in the relative contributions of exchange-equivalent and nonequivalent states. The relative population of these states may be unique to certain PTMs or certain phosphorylation patterns.
By combining the HDX-MS hybrid approach with the Framework Rigidity Optimized Dynamics Algorithm New (FRODAN) program, the theoretical tau ensembles were generated for nTau and pTau proteins (Zhu et al. 2015). Exceptional data was rendered, as seen in Figure 7, showing conformers of nTau and pTau spanning large conformational landscapes. Figure 7 displays the least to most representative conformers, from left to right, associated with nTau (A) and pTau (B). The high-ranking conformer of nTau (Pearson coefficient >0.27) adopted a compact conformation, while that of pTau (Pearson coefficient >0.16) had an extended conformation. The high-ranking conformer of pTau showed full exposure of hexapeptide II (305-315) and sequestration of hexapeptide I (274-284) (Figure 7B), unlike nTau (Figure 7A).

**Figure 7.** Histograms showing relative distribution of structures in agreement with HDX data. Proposed structures of (A) native tau and (B) pTau (the high-ranking conformer is shown to the right) (Zhu et al. 2015).
Data indicate that hyperphosphorylation induces structural changes (release of global fold, development of new intramolecular interactions within the MTBR, exposure of hexapeptide II, and sequestration of hexapeptide I) which may contribute to tau amyloidogenesis.

Ion mobility mass spectrometry (IM-MS) has wide applications, including conformational analysis of proteins. IM-MS allows for detection of multiple conformations of the same species, or oligomers of different sizes. IM-MS was also used to measure relative populations of tau monomers, dimers, and oligomers (Bernstein et al. 2009; Eschmann et al. 2015; Levine et al. 2015). Using the R2/WT peptide (273-284 sequence (Figure 2)), IM-MS profile was analyzed and arrival time distributions were compared for monomer and oligomer (Ganguly et al. 2015). The peptide existed in two charged states (n/z =1/2 and 1/3) and had a significant population of monomers with comparable conformation. Electrospray ionization (ESI)-q-MS was also used to evaluate aggregation of peptides in presence of heparin. Figure 8A shows very little aggregation within 15 min, but a significant aggregation within two hours. The peaks at m/z values above 1000 (Figure 8B) are indicative of soluble tau oligomers. Notably, the amount of aggregates decreased after 2 days which may indicate formation of mature insoluble aggregates, which MS was unable to detect.

MS is a powerful technique for monitoring early-stages of tau aggregation, as well as characterization of structural features induced by PTMs. The structural information obtained by MS may be used to render gas-phase structures of various tau forms.
Figure 8. ESI-q-MS spectra of R2/wt peptide in the presence of heparin at (A) 15 min and (B) 1.5 h (Eschmann et al. 2015).

4. FRET-BASED SPECTROSCOPIC METHODS

Spectroscopic methods, including fluorescence, circular dichroism, and Fourier-transform infrared spectroscopies have been commonly used to evaluate global structure and aggregation of tau protein. While these methodologies allow for some structural insight, an alternative approach may provide additional information required for identifying structure-function relationship of IDPs. Förster resonance energy transfer (FRET) is a spectroscopic assay which detects energy transfer between two fluorophores: an excited donor (D) and a fluorophore acceptor (A) (Figure 9A). The energy transfer is extremely sensitive and dependent on the donor-acceptor (D-A) distance, which is affected by inter- or intramolecular binding interactions. The major requirement for an appropriate D-A pair is for fluorescence emission spectrum of D to overlap with an excitation (absorption) spectrum of A (Figure 9A). In addition to FRET, fluorescence methodologies based on intrinsic and extrinsic probes are commonly used for evaluation of protein structure and misfolding (Chen et al. 2008).
Intrinsic probes are commonly Tyr or Trp residues, and modulation of their fluorescence emission may be used to evaluate structural protein changes. While tau has several Tyr residues, overall change in Tyr fluorescence is less sensitive to conformational flexibility within a protein, compared to other non-biological fluorophores. In addition, all tau isoforms lack Trp residues, limiting their fluorescence use. Thus, by chemically labeling tau protein with fluorescence groups, new avenues have opened up for protein analysis. For example, by site-directed mutagenesis, D-A fluorophores can be introduced into tau. D-A distance in intramolecular FRET (Figure 9B) can be used to probe protein-protein interactions, such as between tubulin and tau.

**Figure 9.** Schematic of (A) D-A fluorescence spectra, (B) intramolecular FRET used to measure tau and tubulin interactions (PDB 1TUB), and (C) intramolecular FRET for evaluating conformational changes within tau.

In addition, D-A distance in intramolecular FRET is useful for probing conformational changes of tau as a function of PTM or aggregation (Figure 9C). Furthermore, fluorescence emissions
intensities for D and A can be used to calculate FRET efficiency (ET\(_\text{eff}\)) and determine relative distances.

D-A labels within tau were used to probe interactions or tau with itself and with MTs. Single molecule FRET (smFRET) was used to elucidate conformational changes of wild-type (WT) and mutated tau to tubulin heterodimers (Elbaum-Garfinkle et al. 2014). All repeat sequences have a conserved PGGG segment, and the Pro mutations within it are linked to a neurodegenerative gene, FTDP-17. Three biologically-relevant mutations were used within the K16 tau construct (198-372): P301L, P332L, and P301L/P332L (Figure 2). P301L tau mutant, which can induce Alzheimer’s disease pathology, has been characterized with greater aggregation propensity compared to WT, and is present in a conserved PGGG segment of R2. P332L mutation is present in a conserved PGGG sequence of R3 domain. In addition, R2 and R3 repeat domains were identified as key sequences in tau aggregation due to their propensity for β-sheet formation. With fluorophores spanning the MTBR, similar FRET efficiencies were observed between mutated and non-mutated constructs (Figure 10A). In presence of tubulin, all tau constructs induced decrease in FRET efficiencies as a result of an extended structure due to binding to tubulin heterodimers, as depicted in Figure 10B. However, P332L and P301L/P332L mutants had a more extended structure and greater tubulin binding affinity than WT and P301L (Figure 10A). It has been proposed that tau adopts an α-helix, and that Pro at the beginning of the PGGG sequence may terminate α-helical structure. The Leu in the LGGG sequence of a mutant may extend an α-helical structure further increasing the “footprint” of mutant tau. A more extended structure in tau mutants has been associated with lower FRET efficiencies due to greater binding to a tubulin dimer (Li et al. 2015).
Figure 10. (A) $E_{T\text{eff}}$ for nTau (WT) and its mutants in absence or presence of tubulin and (B) the schematic illustration of nTau conformation before and after tubulin binding and the resulting change in $E_{T\text{eff}}$ (Elbaum-Garfinkle et al. 2014; Copyright (2014) National Academy of Sciences).

However, greater binding of mutants to tubulin heterodimers was followed by reduction of MT polymerization. It is possible that mutations in K16 constructs allow them to be more sensitive to conformational differences between polymerized tubulin and tubulin heterodimer. Conformational differences of tau and its mutants with tubulin have given a new insight into the importance of tau structure and conformation for binding to tubulin.

While intermolecular probing of tau has benefits, intramolecular probing via FRET can unveil key contact points in the protein itself. FRET-based evaluation of 2N4R with conformers associated with PHFs (2N3R and 0N3R) has led to a rejection of random coil structure within R2 and R3 (Jeganathan et al. 2006). The random-coil assumption has been previously rejected in 2N4R as well (Von Bergen et al. 2006). Furthermore, low FRET efficiency was observed in mutants with various labeling positions, indicating lack of interactions between N-terminus and repeat domains. By labeling and inserting probes within N- and C-termini, an incredibly short
distance was calculated (20.8 Å) which was significantly shorter than the proposed distance (170 Å). The proposed structure of nTau exhibits the following features: 1) compact R domains, 2) partially helical C-terminal tail, and 3) hairpin folding between the N-terminus, C-terminus, and R domains. The data point towards a possible hairpin-like folding of tau protein - a ‘paperclip’-like formation.

Additional intramolecular smFRET of 2N4R have also found that N- and C-termini lack a random coil structure (Elbaum-Garfinkle et al. 2012). The smFRET assay indicated that tau’s compact conformation may be attributed to electrostatic interactions (between negatively charged termini and positively charged MTBR/PRD). Further studies in solution with heparin confirmed previous hypotheses that heparin promotes aggregation of MTBR and its flanking region by increasing solvent exposure. Specifically, the MTBR and PRD experienced compaction, but the interactions within termini and interactions between termini and MTBR were minimized. Comparison of the MTBR (244-354) fragment with the R2-R3 (291-322) repeat domain revealed that the MTBR is more compact in isolation than it is in 2N4R. Tau MTBR fragment is more prone to aggregation than 2N4R. Thus, unlike previous findings, smFRET suggested that tau conformation, based upon termini distance, was more S-shaped than the previously proposed ‘paperclip’ shape. The two proposed structures of nTau, the ‘paperclip’ and S-shaped, are illustrated in Figure 11.

Further FRET studies were carried out with pseudophosphorylated 2N4R containing the Ser199Glu/Ser202Glu/Thr205Glu (AT8), Thr212Glu/Ser213Glu (AT100), Ser396Glu/Ser404Glu (PHF1) mutations (Figure 2) (Jeganathan et al. 2008). These sites are prevalent in Alzheimer’s disease and may induce conformational changes of tau leading to its dysfunction (Goedert et al. 1995; Zheng-Fischhofer et al. 1998). While phosphorylation at
AT100 epitope reduces MT binding, pseudophosphorylation at AT100 site does not change the distance between the N-terminus and R repeat domains. Pseudophosphorylation within the AT8 epitope induced an immense drop in FRET efficiency, which was indicative of N-terminus ‘swinging away’ from C-terminus, undoing the proposed ‘paperclip’-like formation. Further, PHF1 mutation increased the distance between R repeat domains and C-terminus, but also positioned C-terminus in closer proximity to N-terminus. Individually, AT8 and PHF1 pseudophosphorylations expose the repeat domains of tau. The pseudophosphorylations within the AT100 and PHF1 epitopes, in combination, lead to compaction of tau, specifically N-terminus coming closer to R repeat domains. Notably, the AT8 and PHF1 mutant combination showed marked increase in FRET efficiency, indicating further compaction of tau via N-terminus. A tau mutant with all three pseudophosphorylations indicated a more extreme compaction of protein via N- and C-termini near the R repeat domains. The compact structures observed in certain hyperphosphorylated tau forms correspond to the pathological conformations.

Intramolecular FRET assays allow for probing of 1) conformational landscape of tau, 2) determination of key dynamic domains involved in tau-MT interactions, and 3) onset of tau aggregation. These molecular strategies may also allow for evaluation of more complex conformers of tau resulting from PTMs.

5. COMPUTATIONAL APPROACHES

IDPs lack ordered structure, limiting their crystallographic characterization, and thus their solution structures remain largely unidentified. Computational methodologies have allowed for visualization of IDPs and have also been extended to tau peptides and proteins.

The Ac-VQIVYK-OMe peptide sequence of tau, adopts a β-strand conformation, forms parallel β-sheet with other such peptides, and is key to amyloidogenic tau aggregation. Molecular mechanics and density-functional theory were used to evaluate 10,000 conformers, most of which contained the stable β-sheet motif, which was corroborated by FT-IR studies (Vaden et al. 2009). Similarly, replica exchange molecular dynamics (MD) and forward-flux sampling simulations of VQIINK and VQIVYK hexapeptides (Figure 2) have provided insight into the mechanism, rates, and free energy landscapes during aggregation (Ganguly et al. 2015; Smit et
al. 2016). In the gas-phase, the VQIINK peptide does not form β-sheet fibrils, however, the VQIVYK peptide does.

Prediction models were used to statistically determine highly ordered or disordered amino acid distribution within tau. By using 2N4R as a model system, algorithm predictions indicated largely disordered structure of tau except for sequences including R2-R3 repeat and C-terminus (Jeganathan et al. 2006). The ordered behavior predicted for hydrophobic R2 and R3 regions matches well with the solution-state NMR studies of tau, which have shown that these domains were prone to β-sheet formation (Mukrasch et al. 2005). Accelerated MD studies on R1 through R4 found that R1-R3 have a type I β-turn, which matched well with NMR data (Mukrasch et al. 2007). The β-turn motif is highly resistant to denaturing conditions with urea. Other prediction models were also used to determine the extent of amyloidogenic propensities based on hydrophobicity of tau residues. The 20 sequences were identified to have amyloid-like aggregation propensities, with the highest values associated with GYTMHQDQ (N-terminus) and VQIVYK fragments (Moore et al. 2011). Tau sequences studied included the following repeats: (X)$_n$Z, Z(X)$_n$Z ($n \geq 2$) or (XZ)$_n$ ($n \geq 2$) ($X =$ hydrophobic residue, $Z =$ charged or polar residue). Alternating hydrophobic and charged or polar residues within peptide strands promotes Amyloid-like aggregation (Caplan et al. 2000).

The R2/wt fragment (273-284) (Figure 2), which contains a highly aggregating sequence, has been recently analyzed using MD simulations. MD simulations revealed that R2/wt fragment is highly unstructured with interconverting conformers, which are largely extended or compact (Larini et al. 2013). The compact conformation was based on combination of H-bonding in hairpins and helices, and salt-bridges between Asp and Lys groups. Specifically, the salt-bridge interaction between Lys274 and Asp283, located at opposite ends of peptide, is the most
significant. Further, it was found that intrinsically disordered peptides fail to adopt unique structures due to competition between H-bonding and salt-bridges. Figure 12 depicts the normalized probability distribution of finding extended or compact conformations. The R2/wt (Figure 12A) and R2/ΔK280 (Figure 12B) mutant peptides were compared and it was discovered that both peptides exist in a mixture of conformations, but mostly as a compact conformation (darker spots in the plot). The compact conformation was energetically favorable due to presence of H-bonding and salt-bridges.

**Figure 12.** Normalized probability distribution of conformation of one molecule of (A) R2/wt and (B) R2/ΔK280 mutant, as determined by molecular dynamics simulations. (C) The extended and compact conformations of a dimer of R2/wt (Reproduced from Larini et al. 2013 with permission from the PCCP Owner Societies).
Additionally, the mutant has a slightly greater amount of extended conformations indicating that mutation of Lys280 leads to reduction in salt-bridge formation, which has significant impact on overall conformational population. The extended structure of R2/ΔK280 mutant facilitates aggregation into fibrils through pre-structuring of the peptide. The respective dimers of these peptides were simulated, and the extended and compact conformations are shown in Figure 12C. Even the compact conformations of dimers are less prone to aggregation than extended conformations.

MD simulations were also used to evaluate the role of external guest molecules on peptide conformation. For example, urea may denature the globular structure of R2/wt, while a crowding agent, trimethylamine N-oxide (TMAO), may stabilize globular structure. Figure 13 shows the schematic of R2/wt peptide conformations in extended or compact forms (Levine et al. 2015). The conformations are dependent on environmental conditions. For example, urea may break H-bonding leading to extended conformation. However, the crowding agent, TMAO, may lead to increased H-bonding and a more compact structure.
**Figure 13.** Proposed schematic of the R2/wt tau peptide in equilibrium between extended and compact conformations (Levine et al. 2015; Copyright (2015) National Academy of Sciences).

Simulations of various isoforms of nTau and pTau were carried out using multiscale MD approach (Xu et al. 2016). Simulated nTau filaments (Figure 14) exhibit a compact core and an extended outer flanking region surrounding the core. However, upon hyperphosphorylation of both N- and C-termini, there is an expansion of the core. In addition, a simulated view of pTau (Figure 14) shows that phosphosites are predominantly located in flanking region of filament.

![nTau and pTau filaments](image)

**Figure 14.** Simulated nTau and p Tau filaments showing Ser, Thr, Tyr phosphosites (as spheres) within N-(cyan) and C-termini (orange) (the filament core is represented as yellow) (Reproduced with permission from (Xu, L., Zheng, J., Margittai, M., Nussinov, R., and Ma, B. 2016. How does hyperphosphorylation promote tau aggregation and modulate filament structure and stability? ACS Chem. Neurosci. 7(5): 565-575) Copyright (2012) American Chemical Society).

Hyperphosphorylation causes expansion of the filament core due to electrostatic attraction between N-terminus and R repeat domains, as well as repulsion between N- and C-termini. This expansion exposes the MTBR to solvent as described above. Specifically, the amyloid-like filament growth may be induced by exposure of R2 and R3, both of which have high aggregation propensities. In turn, the accessibility of R repeat domains promotes...
aggregation. MD studies also revealed that conformations of pTau associated with Parkinson’s disease and Lewy bodies dementia have a similar hard filament core, but different conformation of the soft outer layers (flanking regions). This finding is significant, since a differential diagnosis of certain neurodegenerative diseases still remains a challenge. While many diseases may share a common biomarker, the biomarker structure and conformation may be vastly different from disease to disease.

Computational and theoretical approaches have elucidated conformational landscapes of tau forms and generated enticing images of tau, revealing its dynamic nature. However, these methodologies require a careful selection of the system size, force field applied, and time scale of conformational changes.

6. CONCLUSION AND FUTURE OUTLOOK

In this review, we summarized recent methodologies and approaches for probing tau’s conformational landscapes despite its dynamic character. Specifically, experimental and modeling platforms were used to identify the native structure of nTau and its various conformers as a result of PTMs. Current methodologies and approaches provide some insight into how and why tau behaves and works in such an elusive manner.

All methods and approaches are in agreement in that nTau does not adopt an extended structure but is rather folded into a “paperclip” or S-shaped conformation. The overall conformation is stabilized by a number of intramolecular points of contacts, including H-bonding and salt-bridges. In addition, it has been discovered that nTau exists in an ensemble of conformations, which is not unusual for IDPs. It has also been identified that tau adopts a local α-helical structure within R2 and R3 domains when bound to tubulin. When tau binds tubulin
dimers, tau adopts an extended structure (containing an α-helix) and allows for tubulin polymerization. Specifically, Pro amino acid terminates the α-helix and defines the surface area used in tubulin binding. The greater the surface area, the stronger the binding to tubulin and lower the MT polymerization.

Upon phosphorylation, pTau undergoes a dramatic conformational change, wherein the “paperclip” or S-shaped structures open up to expose MTBR domain (R repeats). The degree of conformational flexibility is dependent on the phosphorylation type and location of the phosphate groups. For example, hyperphosphorylation in the first segment of PRD promotes conformational change towards a more open structure, while hyperphosphorylation within the later segment of PRD has minimal effect on conformation. In addition, hyperphosphorylation within C-terminus also promoted extended tau conformation. The opening up or extending conformation of hyperphosphorylated tau exposes R repeat domains and promotes aggregation. Specifically, the exposure of R3 repeat and sequestration of R2 repeat triggers the β-sheet formation leading to misfolding and aggregation. The critical hyperphosphorylation which opens-up tau conformation and exposes R repeat domains, leading to aggregation, is within PRD domain and C-terminus. However, some hyperphosphorylation patterns may induce a more compact conformation, while still leading to pathological tau forms. For example, the electrostatic interactions between N-terminus and R domains promote compaction of the core but still expose the MTBR for aggregation. In addition, critical salt-bridge interactions involving Lys residues have been identified, which alongside H-bonding, dictate the overall structure of tau. For example, the R2 repeat domain adopts a largely compact conformation, but its Lys280 mutant has a slightly greater amount of extended conformations. Open conformations have a greater propensity to aggregate than compact conformation. Data indicate that tau conformation
is largely dependent on electrostatic charge perturbations induced by specific PTM patterns. But, it is still unclear how synergistic PTMs (combined phosphorylation and acetylation or others) regulate tau structure and its function.

Despite methodological advances, there are still challenges surrounding IDPs, such as tau, and the role of PTMs on protein conformation. For example, a method-driven evaluation may lead to compartmentalization of data interpretation. These unidirectional methodologies may often be limited in their intrinsic value. However, when used in conjunction with another method, they may reveal key interactions of misfolded proteins. Specifically, this ‘hybrid’ approach may illuminate structural landscapes of tau and unravel ‘tangled mysteries’ of neurodegeneration. For example, by combining FRET analysis with structural information from NMR, the proposed solution tau model may gain conformational accuracy over individual assay evaluations. In addition, creating a phospho-specific FRET assay may allow for precise evaluation of hyperphosphorylation of tau and its impact on tau conformation.

Also, existing methodologies and approaches may be extended to structural analysis of pTau through careful selection of protein kinases. The tunability in kinase-catalyzed hyperphosphorylation may result in discrete tau conformers, which could be characterized by methods such as NMR or MS. Additionally, transient conformational changes during phosphorylation may be monitored by FRET, allowing for identification of key phospho-sites.

Currently, it is unclear which PTM leads to tau aggregation, therefore further studies on aggregation propensities of post-translationally modified tau are needed. By using spectroscopy or MS, early-onset of tau misfolding can be monitored. The morphologies of insoluble tau aggregates may also be characterized by microscopy. In a complex system, several different
PTM combinations should be considered in conformational studies as they more closely mimic biological setting of tau.

Given intra- and extracellular roles of tau, the interactions between nTau, post-translationally modified tau, and their various conformers with other biological targets must be addressed. For example, recent reports on n- or p-Tau interacting with amyloid-β or DNA indicate the importance of tau biomolecular interactions beyond MTs (Manczak et al. 2013; Padmaraju et al. 2010). It is currently unknown which tau conformer is prone to binding other endogenous molecules. For identification of tau binding partners, intermolecular FRET can be used, but may require an extensive labeling approach.

In the field of IDPs, there is a plethora of open-ended questions which may only be answered or addressed through design of novel methodologies and approaches for structural analysis. The continuously evolving structural interpretations of IDPs necessitate formulation of methods and approaches leading to a paradigm shift in this field. Design of novel platforms for evaluating protein misfolding will advance mechanistic understanding of misfolding diseases and search for viable therapeutics.
Figure Captions

Figure 1. Schematic illustration of 6 tau isoforms.

Figure 2. Schematic illustration of specific peptide sequences and key phosphorylation sites introduced by GSK-3β, MARK and Fyn protein kinases.

Figure 3. Schematic of α/β-tubulin subunits within MT and tau prior to and post PTM. PTMs lead to tau modifications and MT depolymerization.

Figure 4. Methods and approaches applied to probe tau’s structure.

Figure 5. Proposed solution structure of 2N4R determined by solution-state NMR. α-Helical structures H1 (114-123) and H2 (428-437) and β-structures B2 (274-284), B3 (305-315), and B4 (336-345) are displayed in gray and light green, respectively. PPP (175-184), PP (216-223) and P (232-239) are polyproline II structures shown in purple. N and C represent the respective terminus of tau (Mukrasch et al 2009).

Figure 6. Binding model based on NMR studies showing interactions between TauF4 fragment with tubulin heterodimers. In U-turn conformation, TauF4 promotes tubulin dimerization. Further MT growth requires extended TauF4 conformation.

Figure 7. Histograms showing relative distribution of structures in agreement with HDX data. Proposed structures of (A) native tau and (B) pTau (the high-ranking conformer is shown to the right) (Zhu et al. 2015).

Figure 8. ESI-q-MS spectra of R2/wt peptide in the presence of heparin at (A) 15 min and (B) 1.5 h (Eschmann et al. 2015).

Figure 9. Schematic of (A) D-A fluorescence spectra, (B) intramolecular FRET used to measure tau and tubulin interactions (PDB 1TUB), and (C) intramolecular FRET for evaluating conformational changes within tau.

Figure 10. (A) ET_{eff} for nTau (WT) and its mutants in absence or presence of tubulin and (B) the schematic illustration of nTau conformation before and after tubulin binding and the resulting change in ET_{eff} (Elbaum-Garfinkle et al. 2014; Copyright (2014) National Academy of Sciences).


Figure 12. Molecular dynamics simulations. Normalized probability distribution of conformation of one molecule of (A) R2/wt and (B) R2/ΔK280 mutant in solution. (C) The extended and compact conformations of a dimer of R2/wt (Reproduced from Larini et al. 2013 with permission from the PCCP Owner Societies).
**Figure 13.** Proposed schematic of the R2/wt tau peptide in equilibrium between extended and compact conformations (Levine et al. 2015; Copyright (2015) National Academy of Sciences).

**Figure 14.** Simulated nTau and pTau filaments showing Ser, Thr, Tyr phosphosites (as spheres) within N-(cyan) and C-termini (orange) (the filament core is represented as yellow) (Reproduced with permission from (Xu, L., Zheng, J., Margittai, M., Nussinov, R., and Ma, B. 2016. How does hyperphosphorylation promote tau aggregation and modulate filament structure and stability? ACS Chem. Neurosci. 7(5): 565-575) Copyright (2012) American Chemical Society).
7. REFERENCES


Fig 1

139x96mm (300 x 300 DPI)
Fig 2

173x99mm (300 x 300 DPI)
Fig 3

134x62mm (300 x 300 DPI)
Fig 5

103x98mm (300 x 300 DPI)
Fig 6

108x76mm (300 x 300 DPI)
Fig 8

44x63mm (300 x 300 DPI)
Fig 9

165x97mm (300 x 300 DPI)
Fig 10

160x52mm (300 x 300 DPI)
Fig 11

167x54mm (300 x 300 DPI)
Fig 12

108x111mm (300 x 300 DPI)
Fig 13

116x74mm (300 x 300 DPI)
Fig 14

nTau

pTau

69x42mm (300 x 300 DPI)