Time-resolved multirotational dynamics of single solution-phase tau proteins reveals details of conformational variation

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-11-2018-006971.R1
Article Type:	Paper
Date Submitted by the Author:	08-Dec-2018
Complete List of Authors:	Foote, Alexander; University of Wisconsin - Madison, Chemistry Manger, Lydia; University of Wisconsin Madison Graduate School, Chemistry Holden, Michael; University of Colorado Denver School of Medicine, Pharmacology Margittai, Martin; University of Denver, Department of Chemistry & Biochemistry Goldsmith, Randall; University of Wisconsin Madison, Chemistry

PCCP

PAPER

Time-resolved multirotational dynamics of single solution-phase tau proteins reveals details of conformational variation

nReceived 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Alexander K. Foote[a], Lydia H. Manger[a], Michael R. Holden[b], Martin Margittai[b], and Randall H. Goldsmith*[a]

Intrinsically disordered proteins (IDPs) are crucial to many cellular processes and have been linked to neurodegenerative diseases. Single molecules of tau, an IDP associated with Alzheimer's disease, are trapped in solution using a microfluidic device, and a time-resolved fluorescence anisotropy decay is recorded for each molecule. Multiple rotational components are resolved and a novel k-means algorithm is used to sort the molecules into two families of conformations. Differences in rotational dynamics suggest a change in the rigidity and steric hindrance surrounding a sequence (^{306}VQIVYK^{311}) which is central to paired helical filament formation. This single-molecule approach can be applied to other IDPs to resolve heterogeneous populations and underlying differences in conformational dynamics.

Introduction:

Tau protein is an intrinsically disordered protein (IDP) at the center of the etiology of Alzheimer's disease. Tau plays a vital role in the human central nervous system by stabilizing microtubules in the axons of neurons[1] and consists of an N-terminal projection domain, a proline rich region, a region made up of 3 or 4 microtubule-binding repeats (MTBRs), and a C-terminal tail[2]. In its pathogenic form the MTBR region loses its microtubule affinity, adopts a β-sheet structure, and aggregates into filaments associated with Alzheimer's disease and other tauopathies[3–6]. As tau progresses along its aggregation pathway, it passes through a loosely defined 'soluble' state in which the protein may be hyperphosphorylated, mislocalized, conformationally changed, and/or oligomeric before becoming fibrillar[7]. Importantly, it has been suggested that it is these soluble tau forms which are toxic, rather than the more conspicuous insoluble aggregates[7–9]. This toxicity underscores how critical it is to study tau in the solution-phase, in order to understand Alzheimer's disease etiology.

Studies of solution-phase, monomeric IDPs have shown, through experiment and computation, that they form heterogeneous ensembles of structures in aqueous solutions[10–13]. Due to this heterogeneous nature of IDPs, the utilization of bulk measurements has an inherent limitation: the individual conformations are difficult to independently characterize and resolve[14]. For example, NMR studies have also shown clear evidence of multiple conformations of tau existing in solution, but teasing out the molecular details of these conformers is difficult and indirect[13]. In contrast, solution-phase single-molecule measurements are powerful methods for characterizing heterogeneous conformations within, and interactions involving, IDPs[11,14–17]. Multiple solution-phase Förster resonance energy transfer (FRET) experiments at the single-molecule level have revealed unexpected behaviors of IDPs including simultaneously existing populations which evolve over time[18] or multiple static subpopulations[19]. This observation of subpopulations of IDPs could not have been directly observed at the bulk level and demonstrates the relevance of single-molecule approaches to the study of IDPs.

In order to probe the conformations and dynamics of tau protein, we performed single-molecule time-resolved fluorescence anisotropy on solution-phase molecules. Fluorescence anisotropy involves first exciting a molecule with polarized light, which preferentially excites molecules with transition dipoles oriented parallel with the excitation's polarization. Fluorescence emission is then partitioned into parallel and perpendicular polarization components and collected in separate detection channels. By using time-correlated single-photon counting to record the arrival time of each photon (i.e. a time-resolved measurement as opposed to a steady-state one) we can observe how the ratio of parallel to perpendicular components evolves over a nanosecond timescale, revealing the nature of the molecule's rotational diffusion with subnanosecond resolution[20]. The timescales of this rotational diffusion can then report on conformational dynamics of the protein. Time-resolved anisotropy has been employed to image environments in cells[21], distinguish different single molecules in solution[22,23], and watch rotational dynamics of tethered single molecules exposed to solution[24]. A difficulty of using time-resolved anisotropy measurements at the single-molecule level arises from partitioning the limited number of collected photons into their respective time-resolved bins,

---

[a] Department of Chemistry, University of Wisconsin-Madison
1101 University Ave, Madison, WI 53706 (USA)

[b] Department of Chemistry & Biochemistry, University of Denver
2199 S. University Blvd., Denver, CO, 80208 (USA)

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x resulting in a noisy – and thus hard to fit – time-resolved decay. Complex protein segmental motions or varying rigidity of structures will show up as additional rotational components in these anisotropy decays. Without sufficient photons, these subtle effects cannot be accurately quantified due to limited statistics. In previous time-resolved anisotropy measurements of single tethered proteins, limited photon counts (and subsequently low S:N ratios) required modeling the parallel and perpendicular decays as single exponentials, making it difficult to quantify multiple rotational components[24]. Analysis that is capable of taking multiple rotational components into account at the single-molecule level is critical to understanding complex protein motions, but has remained elusive.

A common way of yielding more photons (and thus achieving higher S:N ratios) is to immobilize the molecule and study it for a prolonged period of time. However, computational and empirical studies show that immobilization can perturb the protein's conformational dynamics,[25–27] and this issue will likely be even more severe when studying immobilized IDPs. In the absence of immobilization, solution-phase experimental methods which involve examining freely diffusing proteins as they transiently pass through a confocal volume[22] result in limited observation times (<10 ms), which are inadequate if we wish to observe multiple rotation parameters since too few photons are collected.

To circumvent this constraint, we utilize an anti-Brownian electrokinetic (ABEL) trap that allows trapping and monitoring of single fluorescent molecules in solution in a non-perturbative manner over the course of multiple seconds[28]. Briefly, an excitation beam is rapidly scanned in a predetermined pattern through the trapping region of a microfluidic device; each detected photon is correlated to the beam's position, allowing application of real-time feedback voltages that induce position-compensating electrokinetic flows, effectively canceling the Brownian motion of the molecule. The ABEL trap has been used to observe mobility constants and non-linear hydrodynamic interactions of DNA oligomers[29,30], hybridization and melting rates between single-strand and double-strand DNA[31], ATP stoichiometry in a multisubunit enzyme[32], spectroscopic properties and functional roles of pigments in monomeric and multimeric photosynthetic antenna protein[26,33–35], electron spin resonance in fluorescence nanodiamonds[36], and even spectral properties of objects as small as single organic dyes[37,38], all while in the solution-phase. Most recently, time-resolved anisotropy has been used to quantify simple mixtures of ABEL-trapped DNA oligomers[23].

A recent study using steady-state fluorescence anisotropy of tau protein in an ABEL trap showed that the full-length isoform of tau (htau40) has at least two stable families of conformations with differences in protein compaction which can be resolved by steady-state anisotropy [17]. However, steady-state anisotropy measurements are unable to determine the different rotational parameters that define that compaction. Because compaction is an important IDP characteristic that likely regulates inter-protein interaction (and thus pathological aggregation), determining the structural details of the two states and what molecular degrees of freedom differentiate them is of high importance. Being able to measure the complete time-resolved anisotropy of each molecule would provide a substantially more information-rich investigation.

In this study, we use an ABEL trap to take prolonged time-resolved fluorescence anisotropy measurements of freely diffusing monomeric htau40 at the single-molecule level. We then apply a rigorous analytical procedure for examining the anisotropy decays that takes advantage of the increased number of photons and allows classification of multiple rotational parameters, a quantification that has not been accomplished at the single-molecule level. Upon obtaining these rotational parameters, we developed a novel clustering algorithm that appropriately weights these multiple parameters when classifying molecules into protein subpopulations. These experimental and numerical advances result in orders-of-magnitude more photons in our time-resolved decays, quantitative extraction of multiple rotational components from these now statistically robust decays, and distinction between conformational subpopulations within a heterogeneous mixture of freely diffusing single protein molecules based on their extracted rotational components. Thus, substantially more information is available to understand the complex behavior of an important IDP.

Data collection and Time-resolved Decays:

Details of the experimental procedures can be found in the SI. Samples of htau40 were prepared and labeled with Atto647N in the third MTBR as previously described[17], and single tau proteins were trapped in a buffered solution containing 25% glycerol. The fluorescence was resolved into two polarization components which were collected on separate detectors (Figure 1a). Analyzed events were trapped for an average of 2.7 seconds (SI Figure 13). Each detected photon has two arrival times associated with it: the macro-time, assigned as the time difference between the beginning of the acquisition run and the arrival of the photon (with a timescale of ms-s), and the micro-time, the time difference between the most recent laser pulse and the arrival of the photon (with a timescale of ps-ns). Data acquisition occurred in 10-minute intervals, and photon macro-times were recorded and binned into 10 ms bins for analysis (Figure 1a). After binning, a change-point-finding algorithm[39] was run on the binned data to determine where intensity changes occurred (Figure 1a). Changepoints that resulted in an event longer than 1 second, and that had an intensity above a threshold were attributed to the htau40 protein (see SI). For each of these events, the steady-state fluorescence anisotropy was calculated and histogrammed (Figure 1b). It is apparent from this histogram that there is more than one underlying anisotropy population in htau40, as shown previously[17].

In order to analyze the time-resolved components, the micro-times associated with each detected photon that occurred during a trapping event were binned into 16 ps time bins to create the parallel and perpendicular channels' time-resolved decays (Figure 1c). It is important to note that the steady-state anisotropy (Figure 1b) is time-averaged over an entire trapping event, and does not have the resolving power to characterize the multicomponent rotational dynamics that define the two populations. Time-resolved measurements on the nanosecond scale, however, can both distinguish and characterize these rotational dynamics. Two trapped htau40 molecules that resulted in a steady-state anisotropy on either ends of the distribution in Figure 1b were analyzed and their parallel and perpendicular decays were used to calculate time-resolved anisotropy decays for each molecule (Figure 1d). It is clear that the molecules displayed conspicuously different conformational dynamics, and further analysis showed that other molecules from the corresponding areas of the histogram also behaved similarly. The time-resolved anisotropy decays contain substantially more information on the conformational dynamics of htau40 than the steady-state anisotropy, information that originated in the time-resolved parallel and perpendicular decays (Figure 1c). Below, these parallel and perpendicular decays are analyzed to gain insight into the multiple rotational parameters, and thus complex structural dynamics of the protein.

The observed decays for the parallel and perpendicular channels $I_{\parallel}(t)$ and $I_{\perp}(t)$ are a sum of the convolution of the individual channel's instrument response functions $IRF_{\parallel}(t)$ and $IRF_{\perp}(t)$ with their corresponding true decays $i_{\parallel}(t)$ and $i_{\perp}(t)$, and each channel's corresponding background $BG_{\parallel}(t)$ and $BG_{\perp}(t)$, such that

$$\begin{aligned} I_{\parallel}(t) &= i_{\parallel}(t) \otimes IRF_{\parallel}(t) + BG_{\parallel}(t) \\ I_{\perp}(t) &= i_{\perp}(t) \otimes IRF_{\perp}(t) + BG_{\perp}(t) \end{aligned} \tag{1}$$

where the true decays are defined as

$$\begin{aligned} i_{\parallel}(t) &= Ae^{-\frac{t-t_{0,\parallel}}{\tau}}\left[1 + 2\left(\alpha_{local}e^{-\frac{t-t_{0,\parallel}}{\rho_{local}}} + \alpha_{int}e^{-\frac{t-t_{0,\parallel}}{\rho_{int}}}\right)\right] \\ i_{\perp}(t) &= Ae^{-\frac{t-t_{0,\perp}}{\tau}}\left[1 - \left(\alpha_{local}e^{-\frac{t-t_{0,\perp}}{\rho_{local}}} + \alpha_{int}e^{-\frac{t-t_{0,\perp}}{\rho_{int}}}\right)\right] \end{aligned} \tag{2}$$

such that A is an amplitude determined by the number of total photons observed during a trapping event, α_{local} and α_{int} are rotational amplitudes of the local and intermediate motions which sum to the fundamental anisotropy, τ is the fluorescence lifetime, ρ_{local} and ρ_{int} are two rotational correlation times corresponding to different protein motions and $t_{0,\parallel}$ and $t_{0,\perp}$ are instrumentation factors accounting for time delays in the two channels (see SI). As described below, the rotational amplitudes (α_{local}, α_{int}) can be connected to cone angles of the respective protein motions. If the IRFs on the two detectors have identical shapes, and each detector has the same time delay, it is apparent that one can sum both equations in 2 to attain a decay which is independent of the rotational dynamics of the system, which, when convoluted with the instrument's IRF, yields the following:

$$I_{\parallel}(t) + 2GI_{\perp}(t) = 3Ae^{-\frac{t-t_0}{\tau}} \otimes IRF(t) \tag{3}$$

Where G is a compensating factor for the different detection efficiencies of the two detection pathways, and is simply a ratio of the two efficiencies. The above equation can be used to fit the lifetime and instrument parameters independent of the rotational parameters, which limits errors that arise from finding local minima in a large, multi-dimensional parameter space (See SI)[24].

Results and Discussion:

The two channels' time-resolved decays were added together to remove their rotational dependence (Eq. 3, Figure 1c - purple) and fit via Maximum Likelihood Estimation (MLE) (See SI, section 1, 2)[40] to obtain a lifetime (τ) and instrument time delay (t_0). We then fit the independent channels' rotational decays using Equation 2 while keeping the now-known values for the lifetime (τ) and instrument parameters ($t_{0,\parallel}$ and $t_{0,\perp}$) constrained. Each molecule's time-resolved decays were fit multiple times with different initial conditions, and the likelihoods of each fit were compared (See SI), with the highest likelihood chosen as the best fit. To test the efficacy of this method, simulated data with comparable parameters and photon counts to our measured data was run through the same fitting process, and the extracted parameters were shown to closely match the input values (SI Figure 12).

After fitting the decays, we extracted fit parameters of τ, α_{local}, ρ_{local}, and ρ_{int} for each trapped molecule. Each of these parameters contains information about a particular element of the molecules' conformational dynamics. The fluorescent lifetime (τ) informs about the microenvironment of the dye, and measuring it allows one to determine whether different lifetimes were contributing to the distinction between the two anisotropy subpopulations observed in Figure 1b. Ideally, the rotational correlation times (ρ_{local} and ρ_{int}), inform on the size of the segment contributing to that rotational motion (with larger segments giving rise to slower rotations). The rotational amplitude (α_{local}) describes the space that each rotational movement can explore (i.e. that segment's rotational freedom), and informs on the restrictions imposed on that segment's rotation due to the rest of the protein. These rotational amplitudes can be converted to a cone angle of rotation by the following[41–43]

$$\frac{\alpha_{int}}{(\alpha_{local}+\alpha_{int})} = S^2 = \left[\frac{1}{2}cos(\theta)\big(1+cos(\theta)\big)\right]^2 \tag{4}$$

where θ is the angle that defines a spatial cone which the segment of protein is able to sample with correlation time ρ_{local}. This cone angle provides an intuitive picture of the flexibility of a region, with a larger angle representing a larger cone of spatial orientations that a dye or proximal protein regions are able to sample.

The distribution in Figure 1b reproduces the two tau conformational populations observed previously using steady-state anisotropy[17]. However, moving beyond this simple method to sorting the trapped molecules into two underlying populations based on their time-resolved (instead of steady-state) parameters reveals correlations between the different population's rotational behaviors, providing structural information on how the two conformational families of tau differ with unprecedented detail. We developed a novel k-means algorithm that utilizes a Gaussian weighting scheme to sort individual molecules into two populations (See SI, Section 4). We tested this sorting algorithm on simulated data where it consistently and correctly sorted overlapping populations which have comparable parameters to our data (Figure 2), granting high confidence in the classification algorithm. In this simulation, events were generated into either a red or a blue population. Each of these populations had four values associated with it (τ, α_{local}, ρ_{local}, and ρ_{int}) and each of these parameters had a Gaussian probability distribution, Figure 2a. Once generated, the dataset was sorted using our new k-means algorithm (Figure 2b) with >94% accuracy. For comparison, the dataset was also sorted using a standard, Euclidean distance based k-means algorithm (Figure 2c) which was found to sort events with only 54% accuracy. Critically, the standard Euclidean method resulted in the analysis ignoring the critical distinction in the parameter $\alpha_{local,}$ and instead created a clearly unphysically divided distinction in ρ_{int}, where a single peak was crudely divided in half. Having confirmed the algorithm's ability to sort events while taking into account their multiple dimensions, we sorted our experimental events into two populations using our measurements of their time-resolved components (τ, α_{local}, ρ_{local}, and ρ_{int}).

After sorting the events into two populations, each of the populations' parameters derived from equation 2 were histogrammed (Figure 3b, d, f, h) and plotted against the corresponding trapping event's steady-state anisotropy (Figure 3i). Importantly, the steady-state anisotropy was calculated independently from our fitting method by using the total number of photons in each detection channel. Although the sorting algorithm only took the time-resolved parameters (τ, α_{local}, ρ_{local}, and ρ_{int}) as inputs, when the steady-state anisotropy values of the two separate populations were plotted as histograms, the populations clearly separate into high (red) and low (blue) steady-state anisotropy populations (Figure 3i). Importantly, this partitioning occurred even though the sorting algorithm did not take the steady-state anisotropy into consideration while sorting these populations, but rather arrived at them organically. To reiterate, we have sorted the solution-phase tau protein into two distinct populations (red and blue) which are described by two rotational motions measured by the time-resolved parameters (τ, α_{local}, ρ_{local}, and ρ_{int}). Thus tau is rotating on two distinct timescales and the details of these rotations are different for the red and blue populations. Fitting the time-resolved decays using MLE gave the time-resolved parameters that describe each population, while sorting those parameters with the k-means algorithm differentiated the red and blue populations from each other.

Figure 3 has several scatter plots (a, c, e, g) which plot the time-resolved parameters (τ, θ_{local}, ρ_{local}, ρ_{int}) against the steady-state anisotropy, revealing correlations between the components and the steady state anisotropy. For example, in Figure 3c one can see that the blue population of θ_{local} values tends upwards and to the left, and the red population tends downwards and to the right. This tendency indicates that higher θ_{local} values (more free rotation) correlate with lower anisotropy values in the blue population, and lower θ_{local} values (more constrained rotation) correlate with higher anisotropy values in the red population. In contrast, the lifetime scatter plot (Figure 3a) shows little vertical correlation, indicating that the lifetime does not correlate with a high or low anisotropy value. From these correlations, we can determine what physical conformational differences between the two subpopulations of htau40 are contributing to the steady-state anisotropy differences.

We will first describe the role of each rotational component (ρ_{local}, θ_{local}, ρ_{int}) before applying them to analysis of htau40. The rotational correlation time, ρ_{local}, and its corresponding rotational cone angle, θ_{local}, are descriptors of a "local" motion, which includes the dye label itself as well as segmental motions of the protein loop in the immediate vicinity of the dye. These two motions were distinguishable with the higher number of photons in bulk time-resolved studies[17] but are not distinguishable here. This detail means that any numerical value associated with the local rotational motion should be interpreted as deriving from a combination of the dye's motion and the motion of the proximal surrounding protein. A smaller ρ_{local} value indicates faster rotation, while a larger ρ_{local} value indicates that the dye and its immediate protein vicinity are moving more slowly. That larger ρ_{local} value would also suggest a larger segment of the protein is contributing to these local motions, which in turn suggests an increase in structure surrounding the label. Simultaneously, a larger θ_{local} implies a larger degree of rotational freedom, indicating that the structural unit surrounding the dye is allowed to sample a wider range of orientations. Meanwhile, the intermediate rotational correlation time (ρ_{int}) informs on a rotational motion involving a larger segment of the protein, or even a global motion of the asymmetric protein. It is important to note that a past bulk study[17] of htau40 revealed three rotational dynamics while we are only examining two: local and intermediate. The reason for this discrepancy is that even though our measurement enables the collection of substantially more photons than in a freely diffusing approach, there are still fewer photons than in a bulk experiment, and too few photons to effectively discern three separate rotational components. We will now describe in greater detail how all of the individual time-resolved components (τ, θ_{local}, ρ_{local}, ρ_{int}) provide details regarding the conformational differences between htau40's two populations.

The lifetime affects the magnitude of the steady-state anisotropy, as a shorter lifetime results in a photon being emitted before it has a chance to depolarize due to rotational motion of the protein or dye. However, because of the high information-content of our time-resolved measurement, one can clearly see that the two subpopulations of htau40 have no significant difference in lifetime (Figure 3a, b). Thus, the conformational changes that happen within the protein which distinguish these two subpopulations do not result in a significant change in the excited-state dynamics of the dye, such as quenching. Most importantly, because the lifetime of the dye is the same between the two populations, the underlying cause of the two populations must derive from differences in their rotational parameters, indicating that there is a significant change in the conformational dynamics between the two populations.

This difference in rotational parameters is captured by our time-resolved measurements, where certain parameters present a conspicuous correlation with steady-state anisotropy (Figure 3c) whereas other parameters appear non-correlated (Figure 3g) or even anti-correlated (Figure 3e). Differences in the values of these parameters across the two populations suggest a conformational difference between the populations that affects the protein's rotational dynamics. The θ_{local} value for the blue population is larger than the θ_{local} for the red population (Figure 3d), indicating a greater amount of local rotational freedom and thus a lower steady-state anisotropy (Figure 3i) which is consistent with previous work[17]. However, the ρ_{local} shows a counter-intuitive correlation when plotted against the steady state anisotropy (Figure 3e), as its values for the blue population are higher than the red population, indicating a slower rate of local rotation. A larger ρ_{local} value in the blue population would normally lead one expect it to have a higher steady-state anisotropy value. Nonetheless, these effects are compensated for by the blue population's θ_{local} value being larger than the red population's θ_{local}, resulting in the blue population's net lower steady-state anisotropy. Finally, though the average ρ_{int} value is slightly larger for the blue population than for the red population (Figure 3h), the difference is small compared to the width of the distribution of values. Simulations suggest that this breadth is due to conformational heterogeneity within the individual protein populations as opposed to a consequence of limited statistical resolution (see SI, end of Section 5), a result consistent with recent NMR measurements[13]. Further analysis of the θ_{local}, ρ_{local}, and ρ_{int} components shows that the θ_{local} is, indeed, the dominant term, and contributes the most to lowering the anisotropy (SI Figure 14), a conclusion that is consistent with numerical simulations derived from ensemble-averaged time-resolved anisotropy[17]. Thus, even as different rotational components may counter-intuitively pull the steady-state anisotropy in different directions, our multiparameter analysis can assess each component's independent contribution.

Having extracted the characteristic rotational parameters of the two populations, it is now possible to describe in detail the differences in structural fluctuations of the protein between the two conformer families. Variations in ρ_{local} and θ_{local} can be attributed to features of the protein's structure within the immediate vicinity of the label (i.e. the third MTBR). The blue population has a larger ρ_{local} value, meaning the dye and proximal protein loop take longer to rotate. We attribute the blue population's larger local rotational correlation time to indicate an increased amount of local structure and consequent rigidity in the MTBRs proximal to the label, thus slowing down the rotational motion due to the segment's increased size (Figure 4, bold blue line). The blue population also has a larger θ_{local} value, suggesting that this segment can sample a larger number of orientations through its local motion (Figure 4, light blue), even as those local motions are slower.

Looking at the red population's rotational values, one can gain additional insight into the populations' conformational differences relative to the blue population. The red population has a smaller ρ_{local}, indicating a smaller principle structural unit contributes to this rotation (Figure 4, bold red line). At the same time, the red population also has a smaller θ_{local} value, which is indicative of the MTBRs of the protein being less able to sample various spatial orientations (Figure 4, light red). Of specific interest here are the θ_{local} and ρ_{local} parameters and the information they contain regarding the nature of the local MTBR region.

It has been shown that assembly of tau filaments is heavily dependent on a hexapeptide $_{306}$VQIVYK$_{311}$ near the beginning of the third MTBR[44,45], and cryo-EM measurements by Fitzpatrick *et al.* have reported that the third and fourth binding repeats adopt a stable, C-shaped architecture which forms the core of the protofilament upon aggregation[46]. This region contains our labeled residue (310), and this structural model is in agreement with our observations for the blue population. The formation of a large, C-shaped structure (Figure 4, bold blue line) would constitute a structural linking and an increase in rigidity of the surrounding MTBR region, resulting in a bulkier segment with slower ρ_{local}. Thus our observation of a slower ρ_{local} is consistent with the formation of a local C-shaped structure, although the exact shape of the solution-phase protein's local structure cannot be directly identified through our rotational measurements. Simultaneously, a larger θ_{local} indicates that this local structure, while more rigid, is able to sample a larger number of spatial orientations (Figure 4, light blue) due to fewer interactions with the surrounding protein. This suggests an exposure of the MTBR, which would facilitate inter-molecular contact.

Our observations of differing rotational dynamics of the two populations is also in agreement with single-molecule FRET measurements by Elbaum-Garfinkle and co-workers[15]. It was found that the aggregation-prone conformation of monomeric tau protein was characterized by a global extension of the N-terminus away from the C-terminus, and a compaction of the MTBR region. Both structural changes were induced upon addition of heparin, and the compaction of the MTBR was also seen upon truncation of the protein (where both heparin addition and truncation are shown to induce aggregation). The ρ_{local} and θ_{local} values of our blue population fit well with this model while still incorporating Fitzpatrick's cryo-EM results[46]: the formation of a rigid, more extended local structural motif (Figure 4, bold blue line) would result in a slowing down of the local rotational movement (i.e. a slower ρ_{local}). Simultaneously, the increased magnitude of θ_{local} in the blue population relates to the increased range of spatial orientations the local structural motif can sample. In Elbaum-Garfinkle's model, the tertiary structure of the aggregation-prone conformer also changes, with N and C-termini extending away from the MTBR. This extension would result in less steric crowding of the MTBR, and thus an increase to the newly-formed local structure's ability to sample various spatial conformations as interactions with the termini are reduced, resulting in our observed increase in θ_{local} (Figure 4, light blue).

Our observed increase in θ_{local} also adds to the picture developed by Fitzpatrick's cryo-EM measurements[46] by adding information on dynamics, which cryo-EM measurements cannot directly describe. The C-shaped structure that Fitzpatrick described is a central feature of tau's aggregation process, and may be underlying the same increase in structure that our θ_{local} is reporting on. Increasing the number of spatial orientations that this compacted structure samples and reducing the steric hindrance of the N and C-termini surrounding the structure (both of which are implied by an increase in θ_{local}) would increase its exposure to other tau proteins, and would be expected to promote aggregation. It is important to note that there are alternative modes of compaction possible besides the specific formation of a C-shaped structure, including interactions between the different MTBR regions. Such alternative modes are hard to distinguish (because of similar rotational parameters), but may lead to distinct fibril populations.

Two families of populations within solution-phase tau are observed, similar to previous observations using steady-state anisotropy[17], where one population (blue) shows rotational conformational dynamics that suggest conformational similarity to an aggregation-prone structure. The increase in the blue population's θ_{local} value indicates a reduction of rotational constraints on the protein region surrounding the third MTBR, a region containing the hexapeptide $_{306}$VQIVYK$_{311}$ and which forms the core of the aggregated protofilament[44–46]. A loss of these constraints is indicative of exposure of the MTBR and the ability to explore a larger set of spatial orientations, two features that could lead to an increased rate of interaction with the MTBRs of other proteins and consequent aggregation. This notion builds on top of previous models in an important way, as it reveals previously unknown local dynamics in the repeat region which may be important in the initial nucleation steps of aggregation.

While these observations match previous models and make intuitive sense for aggregation, further studies are needed to conclusively characterize which subpopulation is responsible for pathological aggregation. For example, while our experiment is comparable to past solution-phase measurements, it should be noted that the experiment of Elbaum-Garfinkle and co-workers[15] focused on two structures with and without heparin, whereas our work shows two structures co-existing in solution. While addition of heparin certainly shifts more of the protein monomers into aggregation-prone conformations and may be needed to stabilize the fibril nucleus, the existence of aggregation-prone structures does not necessitate the presence of heparin. These conformations may still exist in solution even without the addition of heparin, a notion that is consistent with past NMR studies showing multiple conformations existing in solution without heparin[13], even if aggregation without heparin is still slow due to other kinetic requirements. It should also be noted that the cryo-EM measurements are of aggregated tau, not monomers in solution. We expect solution-phase single-molecule experiments to continue to help unravel the mechanistic connection between aggregation-prone monomers, toxic oligomers, and ultimately mature aggregates.

Conclusions

In this work we have measured time-resolved anisotropy decays of solution-phase monomeric tau proteins using an ABEL trap to cancel Brownian motion without covalent attachment to a surface. With a larger number of detected photons from each molecule we have quantitatively extracted two rotational components from the time-resolved decays of each trapped protein and used a k-means algorithm to sort the rotational components into two families of populations of the protein. Examination of these measured parameters suggests that one population has an increased local structure surrounding the third MTBR, a region that forms the core of tau filaments[44,46], as well as an increase in the orientational freedom of that region. The rotational parameters measured in this work suggest an exposure of the local MTBR in one of the populations, offering physical insight into why this structure could be prone to aggregation. This work highlights the power of single-molecule experiments to identify and characterize subpopulations in detail based on their local conformational dynamics. Our analysis method is general and can be extended to other IDPs and many biologically relevant molecules.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

RHG thanks the NIH (R21NS108684), Alzheimer's Association (NIRG-342100), and Great Milwaukee Foundation (Shaw Scientist) and MM thanks the NIH (R21AG051833) for support. We thank K. Heylman for quartz microfabrication. We also thank Dr. Mary Van Vleet and Dr. Desiree Bates for their expertise and help in utilizing High Performance Computing to expedite our analysis.

Notes and references

1. Mandelkow, E. M., Mandelkow, E. *Trends Cell Biol.*, 1998, **8**, 425–427.
2. Gustke, N., Trinczek, B., Biernat, J., Mandelkow, E. M., Mandelkow, E. *Biochemistry*, 1994, **33**, 9511–9522.
3. Crowther, R. A., Goedert, M. *J. Struct. Biol.*, 2000, **130**, 271–279.
4. Ittner, L. M., Götz, J. *Nat. Rev. Neurosci.*, 2011, **12**, 67–72.
5. Wang, Y., Mandelkow, E. *Nat. Rev. Neurosci.*, 2016, **17**, 5–21.
6. Goedert, M., Eisenberg, D. S., Crowther, R. A. *Annu. Rev. Neurosci.*, 2017, **40**, 189–210.
7. Katherine J Kopeikina, Bradley T Hyman, and T. L. S.-J. *Transl. Neurosci.*, 2012, **3**, 223–233.
8. Spires-jones, T. L., Kopeikina, K. J., Koffie, R. M., Hyman, B. T. *J. Mol. Neurosci.*, 2011, **45**, 438–444.
9. Gendron, T. F. *Mol. Neurodegener.*, 2009, **4**, 1–19.
10. Crick, S. L., Jayaraman, M., Frieden, C., Wetzel, R., Pappu, R. V. *Proc. Natl. Acad. Sci.*, 2006, **103**, 16764–16769.
11. Mukhopadhyay, S., Krishnan, R., Lemke, E. A., Lindquist, S., Deniz, A. A. *Proc. Natl. Acad. Sci.*, 2007, **104**, 2649–2654.
12. Wang, X., Vitalis, A., Wyczalkowski, M. A., Pappu, R. V. *Proteins Struct. Funct. Genet.*, 2006, **63**, 297–311.
13. Mukrasch, M. D., Bibow, S., Korukottu, J., Jeganathan, S., Biernat, J., Griesinger, C., Mandelkow, E., Zweckstetter, M. *PLoS Biol.*, 2009, **7**, 0399–0414.
14. Brucale, M., Schuler, B., Samorì, B. *Chem. Rev.*, 2014, **114**, 3281–3317.
15. Elbaum-Garfinkle, S., Rhoades, E. *J. Am. Chem. Soc.*, 2012, **134**, 16607–16613.
16. Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., Sottini, A., Soranno, A., Buholzer, K. J., Nettels, D., Kragelund, B. B., Best, R. B., Schuler, B. *Nature*, 2018, **555**, 61–66.
17. Manger, L. H., Foote, A. K., Wood, S. L., Holden, M. R., Heylman, K. D., Margittai, M., Goldsmith, R. H. *Angew. Chemie - Int. Ed.*, 2017, **56**, 15584–15588.
18. Banerjee, P. R., Mitrea, D. M., Kriwacki, R. W., Deniz, A. A. *Angew. Chemie - Int. Ed.*, 2016, **55**, 1675–1679.
19. Muller-Spath, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L., Nettels, D., Schuler, B. *Proc. Natl. Acad. Sci.*, 2010, **107**, 14609–14614.
20. Lakowicz, J. R. Principles of Fluorescent Spectroscopy, Plenum Press, 1983.
21. Suhling, K., Siegel, J., Lanigan, P. M. P., Lévêque-Fort, S., Webb, S. E. D., Phillips, D., Davis, D. M., French, P. M. W. *Opt. Lett.*, 2004, **29**, 584.
22. Schaffer, J., Volkmer, a, Eggeling, C., Subramaniam, V., Striker, G., Seidel, C. a M. *J. Phys. Chem. A*, 1999, **103**, 331–336.
23. Yang, H., Moerner, W. E. 2018.
24. Hu, D., Lu, H. P. *J. Phys. Chem. B*, 2003, **107**, 618–626.

25. Friedel, M., Baumketner, A., Shea, J.-E. *Proc. Natl. Acad. Sci.*, 2006, **103**, 8396–8401.
26. Goldsmith, R. H., Moerner, W. E. *Nat. Chem.*, 2010, **2**, 179–186.
27. Talaga, D. S., Lau, W. L., Roder, H., Tang, J., Jia, Y., DeGrado, W. F., Hochstrasser, R. M. *Proc. Natl. Acad. Sci.*, 2000, **97**, 13021–13026.
28. Wang, Q., Goldsmith, R. H., Jiang, Y., Bockenhauer, S. D., Moerner, W. E. *Acc. Chem. Res.*, 2012, **45**, 1955–1964.
29. Wang, Q., Moerner, W. E. *ACS Nano*, 2011, **5**, 5792–5799.
30. Cohen, A. E., Moerner, W. E. *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12622–12627.
31. Wang, Q., Moerner, W. E. *Nat. Methods*, 2014, **11**, 555–8.
32. Jiang, Y., Douglas, N. R., Conley, N. R., Miller, E. J., Frydman, J., Moerner, W. E. *Proc. Natl. Acad. Sci.*, 2011, **108**, 16962–16967.
33. Schlau-Cohen, G. S., Wang, Q., Southall, J., Cogdell, R. J., Moerner, W. E. *Proc. Natl. Acad. Sci.*, 2013, **110**, 10899–10903.
34. Wang, Q., Moerner, W. E. *Proc. Natl. Acad. Sci.*, 2015, **112**, 13880–13885.
35. Squires, A. H., Moerner, W. E. *Proc. Natl. Acad. Sci.*, 2017, **114**, 9779–9784.
36. Kayci, M., Chang, H. C., Radenovic, A. *Nano Lett.*, 2014, **14**, 5335–5341.
37. Wang, Q., Moerner, W. E. *J. Phys. Chem. B*, 2013, **117**, 4641–4648.
38. Fields, A. P., Cohen, A. E. *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 8937–8942.
39. Watkins, L. P., Yang, H. *J. Phys. Chem. B*, 2005, **109**, 617–628.
40. Maus, M., Cotlet, M., Hofkens, J., Gensch, T., Schryver, F. C. De *Anal. Chem.*, 2001, **73**, 2078–2086.
41. Lipari, G., Szabo, A. *Biophys. J.*, 1980, **30**, 489–506.
42. Kinosita, K., Ikegami, A., Kawato, S. *Biophys. J.*, 1982, **37**, 461–464.
43. Ellis, J. P., Culviner, P. H., Cavagnero, S. *Protein Sci.*, 2009, **18**, 2003–2015.
44. von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E.-M., Mandelkow, E. *Proc. Natl. Acad. Sci.*, 2000, **97**, 5129–5134.
45. Xie, C., Soeda, Y., Shinzaki, Y., In, Y., Tomoo, K., Ihara, Y., Miyasaka, T. *J. Neurochem.*, 2015, **135**, 19–26.
46. Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M., Scheres, S. H. W. *Nature*, 2017, **547**, 185–190.

Figure 1:

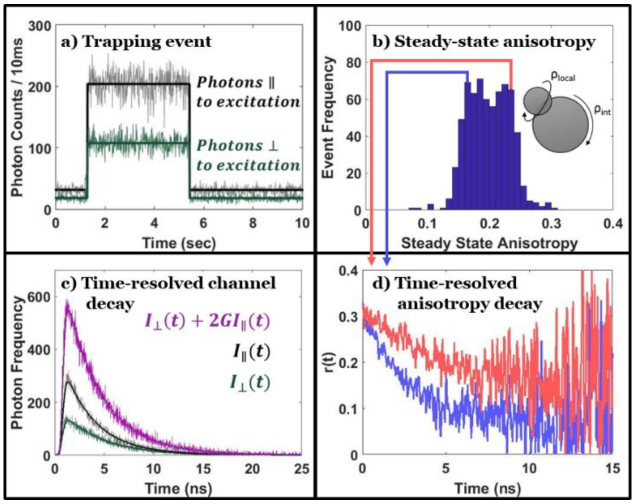

Figure 1. a) Macro-time arrivals of photons binned in 10 ms bins according to their parallel (grey) and perpendicular (green) emission relative to excitation polarization. Trapping events were found using a change-point algorithm (bold line). b) Histogram of steady-state anisotropy values for N = 671 htau40 molecules which can be caused by multiple rotations (inset). c) Micro-time arrivals of photons during a trapping event. Eq 3 is used to create and fit the lifetime decay (purple). Independent parallel (grey) and perpendicular (green) decays are fit (bold line) with Eq 1 and 2. d) Two time-resolved anisotropy decays of single molecules corresponding to events with a low (r = ~0.17, blue) and high (r = ~0.24, red) steady-state anisotropy value.

Figure 2:

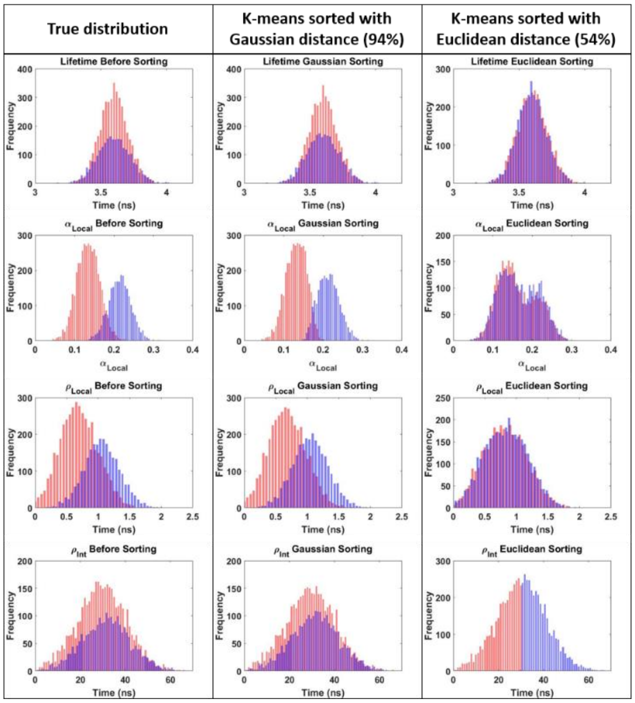

Figure 2. a) Two simulated sets of data (red and blue) described by 4 parameters (τ, ρ_{local}, ρ_{int}, and α_{local}) each, with a total N=6710. The values of the parameters followed two different Gaussian probability functions. The left-most column shows the true distributions. b) The sorted populations output from our Gaussian-based k-means algorithm. Events were sorted into the correct population with 94% accuracy. c) The output when sorting the simulated data using a k-means algorithm that utilizes a Euclidean distance. Events were sorted with only 54% accuracy and unphysical partitioning.

Figure 3:

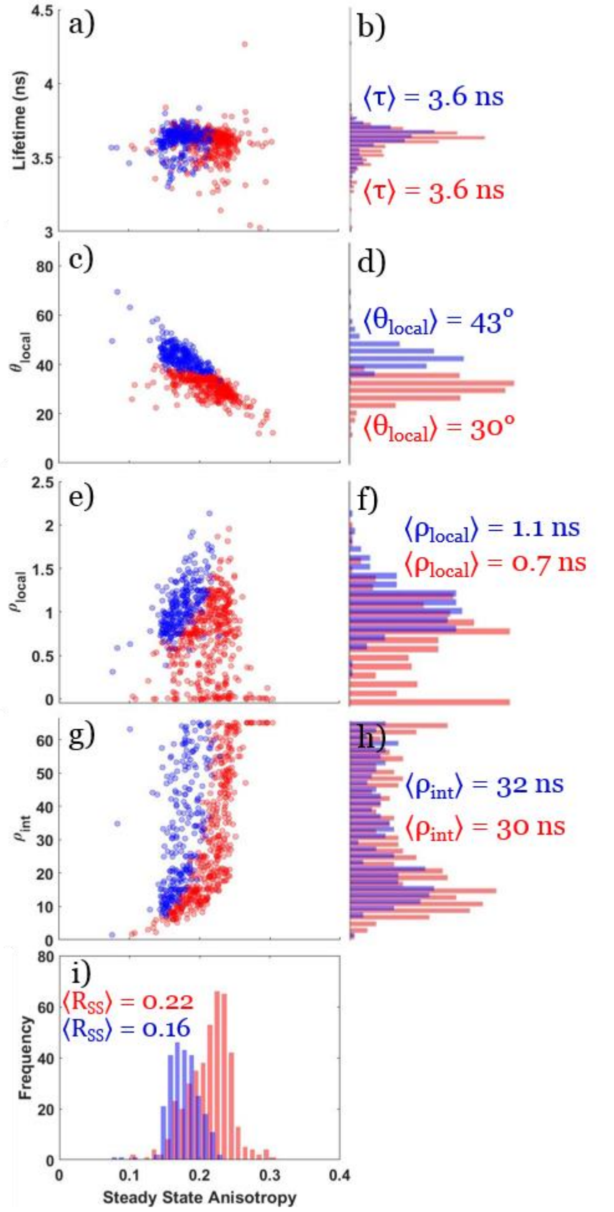

Figure 3. Scatter plots and histograms of the lifetime τ (a-b), and the rotational parameters θ_{local} (c-d), ρ_{local} (e-f), and ρ_{int} (g-h). Scatter plots are scattered against steady state anisotropy (i). All events (N = 671) are sorted into two populations using a k-means algorithm which compares rotational parameters and lifetime measurements.

[TGG1] Figure 4:

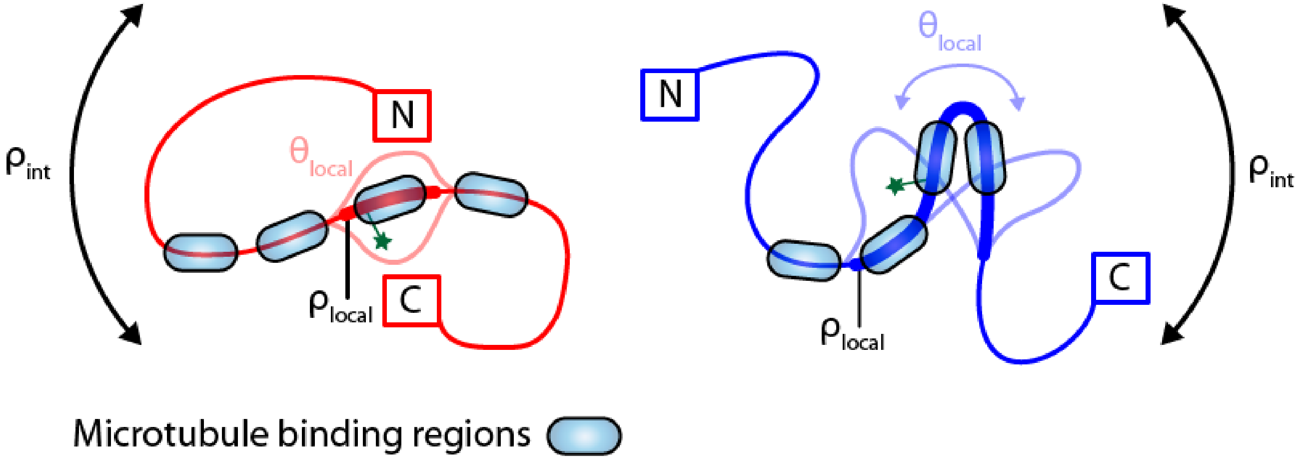

Figure 4. Model of the two families of populations of htau40 based upon previously reported structures[15,46] and measured rotational parameters. The ρ_{local} shows an increased contribution of local structure formation around the dye going from red to blue (bold lines). The θ_{local} shows an increase in the available orientations going from red to blue (transparent lines). The ρ_{int} is caused by a slower rotation of the protein.