Combinatorial native MS and LC-MS/MS approach reveals high intrinsic phosphorylation of human Tau but minimal levels of other key modifications.

Authors:

Friedel Drepper[1,2,¶], Jacek Biernat[3,¶], Senthillvelrajan Kaniyappan[3,4], Helmut E. Meyer[5,6], Eva Maria Mandelkow[3,7], Bettina Warscheid[1,2*], Eckhard Mandelkow[3,7*]

[1] Biochemistry and Functional Proteomics, Institute of Biology II, University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany
[2] Signalling Research Centres BIOSS and CIBSS, University of Freiburg
[3] DZNE (German Center for Neurodegenerative Diseases), Venusberg-Campus 1, Building 99, 53127 Bonn, Germany
[4] Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn, Venusberg-Campus 1, Building 82, 53127 Bonn, Germany.
[5] Medical Proteome -Center, Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany
[6] Leibniz-Institute für Analytical Sciences (ISAS), Biomedical Research Dortmund, Germany
[7] CAESAR Research Center, Ludwig-Erhard-Allee 2, 53127 Bonn, Germany

* Corresponding authors:
Eckhard Mandelkow, E-mail: mand@mpasmb.desy.de
Bettina Warscheid, E-Mail: bettina.warscheid@biologie.uni-freiburg.de

¶ Friedel Drepper, Jacek Biernatare joint first authors

Running title: Native MS reveals high intrinsic phosphorylation of Tau protein

Key words:
Tau protein, native mass spectrometry, LC-MS, phosphorylation, Alzheimer disease

Abstract

Abnormal changes in the neuronal microtubule-associated protein Tau, such as hyperphosphorylation and aggregation, are considered hallmarks of cognitive deficits in Alzheimer disease. Hyperphosphorylation is thought to take place before aggregation, and therefore it is often assumed that phosphorylation predisposes Tau towards aggregation. However, the nature and extent of phosphorylation has remained ill-defined. Tau protein contains up to 85 potential phosphorylation sites (80 Ser/Thr, and 5 Tyr P-sites), many of which can be phosphorylated by various kinases because the unfolded structure of Tau makes them accessible. However, limitations in methods (e.g. in mass spectrometry of phosphorylated peptides, or antibodies against phospho-epitopes) have led to conflicting results regarding the overall degree of phosphorylation of Tau in cells. Here we present results from a new approach, that is based on native mass spectrometry analysis of intact Tau expressed in a eukaryotic cell system (Sf9) which reveals Tau in different phosphorylation states. The extent of phosphorylation is remarkably heterogeneous with up to ~20 phosphates (P_i) per molecule and distributed over 51 sites (including all P-sites published so far and additional 18 P-sites). The medium phosphorylated fraction P_m showed overall occupancies centered at 8 P_i (± 5 P_i) with a bell-shaped distribution, the highly phosphorylated fraction P_h had 14 P_i (± 6 P_i). The distribution of sites was remarkably asymmetric (with 71% of all P-sites located in the C-terminal half of Tau). All phosphorylation sites were on Ser or Thr residues, but none on Tyr. Other known posttranslational modifications of Tau were near or below our detection limit (e.g. acetylation, ubiquitination). None of the Tau fractions self-assemble readily, arguing that Tau aggregation is not promoted by phosphorylation per se but requires additional factors.

Introduction

Tau (*MAPT,* Uniprot P10636) is a developmentally regulated protein promoting microtubule-based functions like axonal transport in neurons. The properties of Tau can be altered by various post-translational modifications (PTMs), most conspicuously by phosphorylation (Wang and Mandelkow, 2016). The aggregation of Tau into amyloid-like insoluble filaments is one of the hallmarks of neurodegenerative diseases called tauopathies, including Alzheimer disease (AD). Early observations suggested a high level of phosphorylation of Tau preceding that of aggregation (Grundke-Iqbal et al., 1986, Braak et al., 1994) which was therefore interpreted to promote aggregation and is used as an early marker of neuronal degeneration. Tau-targeted therapies, aiming at reducing Tau levels, Tau distribution or Tau modifications have emerged as potential strategies for treating tauopathy in patients (DeVos et al., 2017, Yanamandra et al., 2013, Sigurdsson, 2018). Likewise, quantitative diagnostics including the analysis of soluble Tau from cerebrospinal fluid is gaining importance (Barthelemy et al., 2019). These examples underscore the need for improved quantitative diagnostics.

Initial quantitative assessments of Tau-bound phosphate yielded ~8 P_i per Tau molecule in AD Tau (Ksiezak-Reding et al., 1992, Kopke et al., 1993). This was compared with fetal Tau at ~7 P_i per Tau (Kenessey and Yen, 1993) and adult human cytosolic Tau at ~2 P_i (Ksiezak-Reding et al., 1992, Kopke et al., 1993), interpreted to mean that the extent of Tau phosphorylation is abnormally high in AD. In parallel, phosphorylation dependent antibodies and phospho-peptide mapping approaches were developed to identify phospho-epitopes on Tau (e.g.(Augustinack et al., 2002); for a current list see Alzforum.org). However, these antibody-based methods were not well suited for the determination of the occupancy of the P-sites and the overall state of phosphorylation.

The advent of Edman degradation and mass spectrometry (MS) resulted in increased sensitivity and specificity of detection of PTMs on Tau protein. In particular the combination of high performance liquid chromatography with tandem MS (HPLC-MS/MS) resulted in

enhanced coverage of the Tau sequence with newly identified P-sites (review,(Hanger et al., 2009, Hanger, 2020)).

However, to address the issue of hyperphosphorylation versus aggregation, an experimental system was needed which enabled high phosphorylation of defined Tau isoforms. This was achieved with the expression of full-length human Tau (2N4R) in transfected Sf9 cells which yields high protein levels (up to 230 µM in cells) and high diversity of P-sites, as judged by antibody reactivity and MALDI-TOF MS analysis (Tepper et al., 2014). Depending on cell treatment (without or with phosphatase inhibitor okadaic acid), the overall occupancy was estimated around 12 P_i or 20 P_i. Hence the fractions were termed P12 or P20, renamed in this paper to P_m and P_h to indicate medium versus high phosphorylation state, and in contrast to unphosphorylated P_o Tau expressed in *E.coli* **(Fig. 1A-C)**. None of the phosphorylated Tau proteins showed an enhanced tendency for aggregation (Tepper et al., 2014).

Nonetheless, the same fractions were studied by a novel MS-based assay, FLEXITau (Mair et al., 2016), designed to determine the locations and occupancies of all P-sites quantitatively by comparison with isotopically labeled peptide forms. This procedure yielded overall occupancies of only 7 and 8 P_i per Tau molecule in P_m-Tau and P_h-Tau, respectively, with a broad spread (±5 around the mean), and up to 23 observed P-sites. Other MS approaches, designed to identify P-sites in Tau from AD brains and CSF, revealed ~30 P-sites but not the overall occupancies (Sato et al., 2018).

In the present work, we employed native MS to determine the extent of Tau phosphorylation in cells and to clarify the variations between different methods. MS of intact proteins, such as top-down MS (Donnelly et al., 2019) and native MS (Boeri Erba and Petosa, 2015, Hernandez and Robinson, 2007), allow quantification of different proteoforms without requiring prior proteolytic cleavage into peptide fragments or comparison with reference standards (Schaffer et al., 2019). The procedure revealed that full-length Tau in the P_m fraction contained 8± 5 P_i, and in the P_h fraction 14± 6 P_i. Subsequent analysis of phosphorylated peptides revealed up to 51 P-sites in Tau, with variable site occupancies. Finally, the method sensitively detected a low level of different cellular proteins associated with Tau.

Results

Analysis of intact full-length phospho-Tau by native MS

Analysis of tryptic peptides by HPLC-MS/MS reveals sites of protein phosphorylation, but does not directly monitor the extent of phosphorylation per molecule (i.e. overall occupancy) because the detection efficiency varies between phosphorylated and non-phosphorylated peptides. This problem is circumvented by native MS of full-length Tau (Boeri Erba and Petosa, 2015). We compared the phosphorylation status of two differently phosphorylated Tau fractions expressed in eukaryotic Sf9 cells with unphosphorylated Tau from *E.coli* bacteria **(Fig. 1B)**. **Fig. 2** shows native MS spectra of unphosphorylated control Tau (termed P_o, expressed in *E. coli*, bottom), Tau from the fraction P_m-Tau (medium level phosphorylation, expressed in Sf9 cells, middle) and highly phosphorylated fraction P_h-Tau (in presence of phosphatase inhibitor okadaic acid, top). For P_o-Tau the main signals were recorded between *m/z* 2500 and 3500. By charge state deconvolution a series of signals corresponding to charges +16 to +13 can be related to a molecular weight of 45724 Da (**Fig. 2**, open red circles; see also **Fig. 3**). Native MS spectra of P_m-Tau and P_h-Tau show several matching signals in the range below *m/z* 2750 and above *m/z* 3500 (**Fig. 2**, middle and top spectrum). In between, broad peaks are visible for both P_m-Tau and P_h-Tau which can be assigned to a series of charge states +16 to +14 (red squares for P_m-Tau and filled red circles for P_h-Tau). Both series are shifted toward higher *m/z* compared to the corresponding charge states of the control Tau P_o. Further peak series in P_o-Tau were attributed to a trimeric complex (grey circles, bottom trace) of the *E. coli* periplasmic chaperone protein Skp

(molecular mass of full-length Skp 17688 Da, reduced to 15691 Da after cleavage of 20 residue signal peptide (see Table 3)), and to an unknown protein of 59642 ± 5 Da (grey squares, top and middle traces) copurified with P_m-Tau and P_h-Tau from Sf9 cells.

Native MS spectra in the range of *m/z* 2800 to 3500 (**Fig. 3**, left) were analyzed by charge state deconvolution to generate zero-charge mass spectra revealing further differences between these Tau protein forms (**Fig. 3**, right). For P_o-Tau from *E.coli* (bottom), a sharp peak indicates a molecular weight of 45724 Da which is 126 Da lower than the MW of unmodified full-length Tau (2N4R, calculated at 45850 Da). Taking into account the mass accuracy of ± 5 Da this can be explained by a cleavage of the N-terminal methionine (131 Da; Fig. 3). For P_h-Tau from Sf9 cells (**Fig. 3**, top), a broader peak in the zero-charge mass spectrum displays a center mass of 46883 Da and up to six additional peaks on both sides equally spaced by approximately 80 Da. The center mass would correspond to an N-terminally acetylated htau40 carrying 14 ± 6 P_i groups.

The spectrum of P_m-Tau was less well resolved and possibly disturbed by a feature at *m/z* 2750, which is present in the P_h-Tau spectrum at much lower intensity. A charge state deconvolution results in a center mass of 46408 Da $\pm$ 400 Da and , indicating 8 ± 5 P_i (Fig. 3, middle row). Of note, although the native MS spectrum of P_m-Tau was less well defined, all three charge states of +16 to +14 are clearly shifted towards lower *m/z* values compared to the P_h-Tau spectrum, but remained at higher *m/z* values compared to the corresponding charge states in the control spectrum of P_0-Tau (see **Fig. 2** and **Fig. 3**).

Identification of peptide sequences and posttranslational modifications by LC-MS/MS

(a) Phosphorylation

For protein identification and detection of PTMs the protein samples were subjected to reduction of cysteine residues, alkylation of free thiol groups and digestion with trypsin followed by LC-MS/MS analysis. The identified tryptic peptides covered 95% of the Tau sequence and allowed us to detect sites of posttranslational modifications of the Tau proteins. We therefore analyzed our LC-MS/MS data for the following amino acid modifications: acetylation of the protein N-terminus, phosphorylation of serine, threonine or tyrosine, acetylation of lysine, and ubiquitination of lysine. For purified P_m-Tau and P_h-Tau, only N-terminally acetylated forms of peptides containing the protein N-terminus (without Met1) were detected. In contrast, for control P_0-Tau these two peptides were identified in their non-modified, non-acetylated N-terminal peptide form. We therefore deduced from the protein sequence (residues 2-441) average masses of 45718.2 Da for the non-N-terminally acetylated *E.coli* Tau (P_0) and 45760.2 Da for the N-terminally acetylated form of htau40 (P_m and P_h) in the absence of further modifications.

The total number of P-sites on serine and threonine residues on Tau identified at a localization probability higher than 0.75 was 51 sites (32 Ser, 19 Thr). No phospho-tyrosines were detected (for P-sites see Table 1; for procedures to derive site probabilities see Methods). No P-sites were detected for *E. coli* P_0-Tau, as expected. The P-site analysis of Tau from Sf9 cells confirmed the majority of P-sites found in previous studies (Mair et al., 2016, Tepper et al., 2014), including all prominent Alzheimer specific antibody phospho-epitopes, notably those that require pairs of phosphorylated residues (e.g. antibodies AT-8, PHF-1, AT100). Furthermore, the analysis revealed ten additional P-sites (T373, T377, S400, T403, S413, T414, S416, S422, S433, S435, see Table 1) in the C-terminal domain of Tau, which have been reported for Tau from human AD brain (reviews (Hanger et al., 2009, Hanger, 2020)) but were absent in our previous studies. Interestingly six P-sites (T245, S285, S293, S316, S320, S352 , see Table 2) were identified in the repeat domain which are not annotated in the collection of P-sites from AD brains (Hanger et al., 2009, Hanger, 2020).

The highly phosphorylated fraction P_h contains 14 Pi (± 6 Pi) per Tau molecule indicating a maximal value of 20 Pi (higher values are possible, but not resolved in the spectrum). The heterogenous P-site distribution over the entire length of the Tau chain is generally in line with previous findings (Mair et al., 2016)(and see Discussion).

(b) *Other posttranslational modifications*

In contrast to the many identified P-sites, it is remarkable that other well-documented PTMs of Tau reported for human and mouse brain are close or below the detection threshold in Tau from Sf9 cells. This includes lysine acetylation and lysine ubiquitination, both of which have been linked to neurodegeneration in AD (Min et al., 2010, Cohen et al., 2011, Arakhamia et al., 2020) as well as O-GlcNAc modification of serine and threonine, which was reported to prevent phosphorylation and aggregation of Tau (Hart et al., 2011, Yuzwa et al., 2012). Of note, in contrast to the high number of detected P-sites and their high abundancies, other PTMs, if significant at all, were observed only sporadically and in low intensity. Further validation and characterization of such low abundant additional modifications would need specialized approaches and is beyond the scope of the present study.

(c) *Tau from neuronal origin*

In addition to Tau expressed in cell models we studied the phosphorylation status of Tau isolated from brains of transgenic mice, using LC-MS/MS and native MS methods. The LC-MS/MS analysis was performed on neuronal Tau from transgenic mice expressing the mutation Tau^{A152T} (a risk factor for PSP, (Coppola et al., 2012)) purified by immunoprecipitation using the Tau^{A152T} specific antibody 1C5. The analysis showed that human Tau^{A152T} expressed in mice is highly phosphorylated at five sites (T181, T231, S199, S202, S396; P-site probability > 0.75). These sites are consistent with the subset of P-sites in Sf9 cells with the highest occupancy (>50%) (Mair et al., 2016) and with prominent P-sites of Tau in AD brain. However, native-MS did not reveal a specific spectrum of Tau from neuronal origin, mainly because of the inherent heterogeneity of neuronal Tau proteins, low concentration and contamination by other proteins. This problem will require further development.

Proteins co-purifying with Tau identified by LC-MS/MS

The high sensitivity of LC-MS/MS revealed the presence tryptic peptides originating from traces of other proteins accompanying Tau through the purification procedure. The recombinant Tau proteins (P_0 from *E.coli*, P_m and P_h-Tau from Sf9 cells) were purified by making use of the heat stability of Tau protein whereby nearly all other proteins are denatured and precipitated whereas Tau remains in solution because of its hydrophilic nature. The detected small amounts of cellular Tau-binding proteins also passed through the boiling step and were still present in gel filtration column fractions containing the separated Tau proteins. One possibility to explain the presence of these proteins is their high affinity for Tau protein. Thus, the experiments yielded a list of ~~high-affinity~~ Tau-interacting proteins which might become interesting for analyzing possible Tau interactions in other cell types (see **Table 3** for P_0-Tau from *E. coli* and **Table 4** for P_m-Tau and P_h-Tau from Sf9 cells). A substantial part of identified proteins are known for their roles in interacting with RNA or DNA, and with chaperones (see Discussion).

Discussion

Significance of Tau phosphorylation in AD research:
Tau is a neuronal protein whose best-known role is to stabilize microtubules in axons, thus supporting their role as tracks for axonal transport and as struts of axonal shape (Drubin et al., 1985). Tau has a hydrophilic character, contains many charged residues, is highly

soluble and has a mostly unfolded structure; as such it can interact with many cellular components in addition to microtubules (Wang and Mandelkow, 2016). The interactions can be modulated by various PTMs, notably by phosphorylation at Ser, Thr, or Tyr residues (up to 85 potential sites). It has been reported that in physiological conditions human fetal Tau is highly phosphorylated (~7-10 Pi per Tau molecule), compared with human adult Tau (~2 Pi), and high again for human AD patients (7-10 Pi) (Kenessey and Yen, 1993, Kopke et al., 1993, Ksiezak-Reding et al., 1992). However, these values need to be re-interpreted in the light of later discoveries (see below). In hibernating animals, Tau becomes highly phosphorylated in a reversible manner, corresponding to a cyclic regression and reappearance of dendritic trees, which indicates that Tau phosphorylation may play a role for neuronal plasticity (Arendt et al., 2003). Many cell and animal models of Tau have been developed to study its functions. In particular, Sf9 cells served as early examples of Tau's role in the control of process outgrowth via microtubules and actin filaments (Baas et al., 1991, Knops et al., 1991) and the importance of reversible phosphorylation (Biernat et al., 2002). Beyond cell biological issues, the major interest in Tau arises from its property as a hallmark of brain diseases, notably Alzheimer and other tauopathies (DeTure and Dickson, 2019). The finding that Tau becomes aggregated into amyloid filaments in AD and appears hyperphosphorylated triggered an extended search for kinases and phosphatases responsible for Tau's pathological state [reviews, (Martin et al., 2013, Hoffman et al., 2017)], as well as searches for P-sites on Tau (Hanger, 2020), and treatments based on these results (Schneider and Mandelkow, 2008). The aim was to identify P-sites as early indicators of pathology in brain tissue, CSF, or in blood (Barthelemy et al., 2020, Bateman et al., 2020), and as indicators of vulnerable brain circuits affected during Braak stages (Arnsten et al., 2019). In spite of this progress, the relationship between Tau's phosphorylation and aggregation remains enigmatic in view of its exceptional solubility.

Methods development:
In the past, the main tools for identifying phoshorylation sites of proteins were based on tryptic cleavage, combined either with thin-layer chromatography of ^{32}P-labeled phosphopeptides, peptide sequencing (Boyle et al., 1991), or mass spectrometry of peptides (e.g. MALDI-TOF MS) (Hasegawa et al., 1992). For the case of Tau, these methods yielded information on P-sites involved in AD (e.g. (Gustke et al., 1992, Illenberger et al., 1998, Tepper et al., 2014)). A limitation common to these methods was that the extent of phosphorylation of peptides or the whole protein remained uncertain. For peptides, this was partially overcome by using isotope-labeled standards (e.g. (Mair et al., 2016, Barthelemy et al., 2016)). However, the correlation between P-sites and global extent of phosphorylation remained a problem. The early studies of global phospho-occupancy, obtained by ashing and color reaction of samples, revealed average global occupancies but carried no information on occupancy of individual P-sites. Yet knowledge of both the degree of phosphorylation and the main P-sites would be important for assessing the cellular functions (e.g. control of microtubule assembly, phase separation in the cytoplasm) and the pathological state of Tau (e.g. somatodendritic missorting, pathological aggregation). This can be achieved by combining native MS with LC-MS/MS, as described here. We will restrict our discussion to some issues that have received attention in the Tau field during recent years.

Overall extent of phosphorylation:
In previous work (Tepper et al., 2014), we had described three major fractions of phospho-Tau, termed P0, P12, and P20, whose degree of phosphorylation was estimated from the peaks observed by MALDI-TOF MS. Fraction P0 corresponded to the unphosphorylated protein expressed in *E.coli*. Fraction P12 was derived from Sf9 cells after expression of full-length Tau (2N4R) whose mass was consistent with additional ~12 Pi. Fraction P20 was also from Sf9 cells, with phosphatase inhibitor okadaic acid added during preparation, consistent with additional ~20 Pi.

The same types of fractions were studied by Steen and colleagues (Mair et al., 2016) using the FLEXITau method where quantification was done with reference to isotope-labeled full-length Tau. The fractions were denoted as P-Tau (equivalent to the P12 preparation of (Tepper et al., 2014)), but with a lower overall occupancy of 7 P_i, and PP-Tau (equivalent to P20 preparation) with an overall occupancy of 8 P_i. In addition, the method revealed the individual occupancies of 17 sites, including those of the major Tau antibodies (Fig. 3, 4 in (Mair et al., 2016)).

The apparent discrepancies between the overall occupancies observed in the earlier MALDI-TOF MS and FLEXITau experiments are now superseded by the results from native MS. The fractions are now denoted as P_0, P_m, and P_h for zero, medium, and high state of phosphorylation (in order to distinguish the different numerical values). The common control value is P_0-Tau from *E.coli* (Fig. 3C), which remains without modification, apart from N-terminal processing by removal of Met 1. The medium state, reflecting normal kinase/PPase activity in cells (Fig. 3B), is now resolved into a series of peaks centered around a maximum at +8 P_i per molecule, with ~4 sub-maxima resolved on either side (range +4 to +13 P_i). The high state (Fig. 3A), reflecting normal kinase but inhibited PPase activity, appears as a series of peaks centered around a maximum at +14 P_i, with ~6 sub-maxima resolved on either side (range +6 to +20 P_i).

The main conclusions are: The previous MALDI-TOF MS values agree roughly with the upper limits of the peaks resolved by native MS, the FLEXITau values reflect the lower limits. Of note, a closer inspection of the MALDI-TOF MS results (Tepper et al., 2014) shows that the experimental errors of the reported masses were up to +/-195 Da, corresponding to +/-2.5 P_i for individual fractions and thus +/-5 P_i for mass differences. On the other hand, the FLEXITau method achieved accurate occupancies of individual P-sites. However, sequence coverage was limited (75%) so that the degree of Tau phosphorylation was underestimated (Table 2, Table S1; note that the newly covered sequences in this study contain 14 additional P-sites).

It is remarkable that Tau molecules in the many P-states are separated by 1 P_i unit. For a multi-domain protein (or multi-protein complex) with specific functional P-sites one would expect well-defined stoichiometries of phosphorylation (by specific kinases) whereby assembly or activity are controlled; examples are ribosomes or proteasomes (Guo et al., 2017, Mikulik et al., 2011). By contrast, in case of Tau all domains have a natively unfolded character, with an unusually high fraction of phosphorylatable residues (85/441=19%) accessible to multiple kinases and phosphatases. The degree of phosphorylation therefore appears to depend on a phosphorylation "tone" rather than on specific functional modifications. The balance between kinase and phosphatase activities depends on the experimental conditions of Tau purification which can change rapidly. Thus, during postmortem delay, ATP is depleted (making kinases inactive), and $PP2A_{cat}$ is activated, which makes adult human Tau appear to be in a low state of phosphorylation. Conversely, rapid isolation of brain tissue (e.g. with fetal Tau or brain biopsy) reveals a high state of phosphorylation. In this scheme, the high phosphorylation of AD-Tau (in spite of a long postmortem delay) is explained by Tau aggregation which protects them against PPases. Similarly, lowering the temperature (as in anesthesia) decreases PPase activity, with a lesser effect on kinases, causing the impression of hyperphosphorylation (Planel et al., 2007, Matsuo et al., 1994, Wang et al., 2015). Thus, although there are specific functional or diagnostic P-sites in Tau they probably have only a small effect on the average degree of phosphorylation. By implication, the changes in phosphorylation tone would affect many cellular proteins, not just Tau, so that other proteins might also serve as diagnostic markers for AD neurodegeneration, even when they do not appear in aggregated form.

If multiple phosphorylation of a natively unfolded protein like Tau is largely a statistical property, it follows that each state of phosphorylation represents an ensemble of many Tau species, phosphorylated at different sites, depending on activity and accessibility of

kinases/PPases. Certain sites may appear as reliable markers of disease, but this is not necessarily coupled to a specific function (such as aggregation or MT binding). Consistent with this, we find that none of the phosphorylation states led to a pronounced increase in Tau aggregation (Tepper et al., 2014) even though Tau aggregation and phosphorylation seem to be co-regulated during AD progression. However, the extent of phosphorylation will change some bulk properties such as charge, which could turn from net positive to net negative. This would change the interactions with other proteins (e.g. the cytoskeleton via microtubules and actin filaments), RNAs, or membrane surfaces, as well as the folding of Tau from a "paperclip" to other conformations (Jeganathan et al., 2008). In particular, phosphorylation promotes the propensity of Tau to undergo liquid-liquid phase transitions (Wegmann et al., 2018).

Asymmetric distribution and hotspots of phosphorylation:
The sequence of Tau contains several points where phosphorylatable residues are clustered and are phosphorylated, either with high occupancy at a single residue, or with intermediate occupancy at two or three nearby residues. Some of these hotspots of phosphorylation are prominent as epitopes of antibodies raised against AD-Tau. The antibodies may bind to a single site within a cluster, independently of nearby P-sites (e.g. pT181 by antibody AT270, or pS231 by antibody AT180 (Amniai et al., 2011), or require a pair of nearby phosphorylated residues (e.g. S202+T205 (AT8, (Goedert et al., 1995)), S212+S214 (AT100, (Zheng-Fischhofer et al., 1998)), S396+S404 (PHF1, (Otvos et al., 1994)). These clusters are represented by the epitopes of antibodies AT270, AT8, AT100, AT180, 12E8, and PHF1, and together they could represent a substantial fraction of Tau phosphorylation (~30% or more; Fig. 5 and (Mair et al., 2016), Fig. 4b). NMR studies revealed some local structural characteristics within this phosphorylated cluster region in Tau (Lippens et al., 2016).

It is remarkable that all the clusters of phosphorylatable residues mentioned above are located in the C-terminal half of Tau which is responsible for MT interactions and PHF aggregation (Fig. 4). This reflects the asymmetric sensitivity of the domains towards phosphorylation. We can divide up the sequence into an N-terminal 1/3 (approx. residues 1-150, acidic character) and the C-terminal 2/3 (residues 151-441, basic character, including the proline-rich and repeat domains). The total number of 51 detected phosphorylated residues (Fig. 4) represents 51/85=60% of all phosphorylatable residues (S, T, Y). However, the N-terminal 1/3 had only 10/27=37% potential sites phosphorylated, whereas in the C-terminal 2/3 the fraction of phosphorylated vs. potential sites was almost twice as large, 41/58=71%. Since all Tau domains are largely unstructured and mobile in solution (Mukrasch et al., 2009), it is unlikely that the difference in sensitivity to kinases arises from a protection of sites via folding (especially since the N-terminal part is known for its role as a "fuzzy coat" (Jakes et al., 1991). Other conceivable explanations are (1) some kinases of Tau may prefer target motifs of basic character, and/or (2) kinases of Tau may first bind to acidic domains before phosphorylating other parts of the Tau chain. Direct evidence for this scenario comes from NMR studies of the kinase MARK which binds to the N-terminal region of Tau and then phosphorylates the KXGS motifs in the repeat domain (Schwalbe et al., 2013).

Tau phosphorylation versus aggregation in AD:
The hotspots of Tau phosphorylation in Sf9 cells coincide well with the epitopes of the antibodies mentioned above, raised initially against PHFs purified from Alzheimer brain. The coincidence of aggregated and phosphorylated states in AD tissue led to the hypothesis that (hyper-) phosphorylation predisposes Tau for aggregation, and set off an extended search for kinases of Tau (Martin et al., 2013, Kosik and Shimura, 2005). However, this effort remained inconclusive, as no kinase with pronounced pro-aggregant activity was found (in contrast to FTDP17 Tau mutations which are clearly pro-aggregant), and some phosphorylation sites were even protective against aggregation (Schneider et al., 1999). The high solubility of Tau from Sf9 cells illustrates the lack of a direct causal link between phosphorylation and aggregation. In support of this, although AD-diagnostic antibodies react with Tau monomers from Sf9 cells (see (Tepper et al., 2014), this is not per se an indicator of

cell degeneration except in diseases such as AD and transgenic mice with Tau pathology. Indirect causal links may include phosphorylation-dependent conformational changes combined with proteolytic cleavage which exposes aggregation-prone domains of Tau (e.g. repeat domain). Examples are cleavage by calpain or caspases (Wang et al., 2007, Flores et al., 2018) or non-enzymatic cleavage in aging cells which causes the heterogeneous "smear" in SDS gels of AD Tau (Watanabe et al., 2004). In support of this, incipient neurodegeneration correlates with the appearance of soluble tau oligomers rather than Tau-containing tangles (Lasagna-Reeves et al., 2012, Kaniyappan et al., 2017).

Tau phosphorylation in cell models:

As mentioned earlier, the functions of Tau in cytoskeletal remodeling, studied initially in PC12 cells and neurons (Drubin et al., 1985), were well replicated in other eukaryotic cell types of human or animal origin including Sf9 cells (Knops et al., 1991). However, one may ask whether this also involves similar P-sites. In previous work we addressed this issue by studying CHO cells transfected with human Tau (2N4R), and a mouse neuroblastoma cell line (LAN-5) expressing endogenous Tau (Illenberger et al., 1998). The cells were metabolically labeled with ^{32}P, followed by phosphopeptide mapping. There was a striking overlap of P-sites with the current results, especially involving the clusters of P-sites listed above. Phosphorylation became enhanced during mitosis, corresponding to the activation of the cell cycle kinase cdc2, a proline-directed kinase. Note that this phosphorylation occurred during a physiological process, without any sign of pathology or Tau aggregation, indicating that the pattern of phosphorylation serves a normal biological role. The same conclusion can be drawn from the analysis of hibernating animals where retraction and regrowth of neuronal networks are accompanied by similar changes in Tau phosphorylation (von der Ohe et al., 2007, Arendt et al., 2003).

Other types of PTMs in Sf9 cells:

So far we have focused on phosphorylation because it is the most conspicuous feature of Tau modification and is at the focus of biomedical research. On the other hand, research on Tau has unearthed a large variety of modifications, including chemical modifications of residues, cleavage by various proteases, or specific interactions with other regulatory components (Morris et al., 2015). For the case of Sf9 cells, the striking observation is the very low level of other modifications that have received attention in the field:

Phosphorylation of Tyr is absent or below detectability, in contrast to the abundance of Ser/Thr phosphorylation. The 5 Tyr-residues in Tau (residues 18, 29, 197, 310, 394) can be phosphorylated by common Tyr kinases in the brain, e.g. Fyn, Src, Lck, Abl, and the proline-rich regions contain several binding sites for associated SH3 domains of the type PXXP (Derkinderen et al., 2005, Hanger et al., 2009, Ittner et al., 2010). This makes Tau interesting as a carrier of cellular signaling. However, at least in Sf9 cells, the Tyr-dependent signaling affects at best a minute fraction of Tau.

Acetylation of Lys: This modification is important for gene regulation, but occurs in the cytoplasm as well, notably on microtubules and associated proteins (e.g. K174 in Tau) and has been implicated in neuronal damage of aging cells (Sohn et al., 2016). The low level of acetylated Tau can be explained by the high activity of HDACs in Sf9 cells which makes them highly radiation-resistant (Sharma et al., 2016).

Ubiquitination of Lys is a frequent Tau modification in aging or AD brains and reflects the cell's attempt to remove misfolded or aggregated Tau (Kosik and Shimura, 2005, Yu et al., 2019). Ubiquitin binding sites have recently been revealed by high-resolution cryo-EM imaging of Tau filaments (Arakhamia et al., 2020). However, this modification is not detectable in Sf9 cells. A simplified interpretation of these results is that Sf9 cells are "healthy" in a cellular sense: Ubiquitination of Tau is very low, consistent with the absence of aggregates. Likewise, acetylation is low as well. The phosphorylation tone is high, comparable with that of fetal neurons in absence of aggregates. Further evidence for the

healthy state is that the Sf9 cells develop extensions analogous to those of differentiating neurons, stabilized by bundles of parallel microtubules and microtubule associated Tau (Biernat and Mandelkow, 1999, Knops et al., 1991). This implies that the increased Tau phosphorylation is not necessarily detrimental and occurs under both physiological and pathological conditions.

Tau-associated proteins:
In both cell types studied here (*E.coli* and Sf9), LC-MS/MS and native MS revealed traces of numerous proteins which remained associated with Tau through the purification (see Tables 3 and 4, Figures 2 and 3). This agrees well with the multiple published interactions of Tau (Mandelkow and Mandelkow, 2012). Many of the accessory proteins belong to the class of ribonucleoprotein complexes (Gunawardana et al., 2015, Maziuk et al., 2017). Examples are proteins involved in mRNA metabolic processes, e.g. translation initiation, and translation elongation (Wheeler et al., 2019). Tau expression can differentially shift both the transcriptome and the nascent proteome, and the synthesis of ribosomal proteins is reversibly dependent on Tau levels (Meier et al., 2016, Koren et al., 2019). The presence of DNA binding proteins is reminiscent of nuclear functions of Tau in DNA protection under stress conditions (Multhaup et al., 2015). Finally, there are numerous examples of interactions between Tau and chaperones (Karagoz et al., 2014), a class that is prominent among the identified associated proteins. For example, the periplasmic chaperone protein Skp from *E.coli* interacts with membrane proteins, thus maintaining the solubility of early folding intermediates during passage through the periplasm (Chen and Henning, 1996). Skp is similar to Prefoldin/GimC, a cytosolic chaperone present in eukarya and archaea (Cowan and Lewis, 2001). Prefoldin, a protein chaperone used in protein folding complexes, works as a transfer protein in conjunction with a molecule of chaperonin to correctly fold other nascent proteins, such as the cytoskeletal proteins actin and tubulin.

In summary, by combining the results from native MS analysis of intact Tau and peptide identification using LC-MS/MS we have been able to specify the state of modified Tau in eukaryotic Sf9 cells in comparison to the nearly unmodified state of Tau expressed in *E.coli*. Among the identified PTMs on Tau from Sf9 cells, extensive phosphorylation at Ser/Thr residues is by far the dominant modification (up to 20 P_i per molecule), whereas other well-known modifications are present only in minute proportions (acetylation, Tyr phosphorylation, ubiquitination). The S/T phosphorylation has a remarkably asymmetric distribution, whereby a low proportion of residues is modified in the acidic N-terminal domain of Tau, in contrast to a high proportion in the basic middle to C-terminal domains, including all of the sites detected by antibodies against Alzheimer phospho-Tau. This indicates that a substantial degree of phosphorylation (achieved by multiple kinases) is presumably a property of normal soluble Tau in eukaryotic cells, reflecting a phosphorylation tonus dominated by net activity of kinases over phosphatases. By implication, this would be expected to affect many other proteins as well, notably natively unfolded proteins accessible to diverse kinases (Oka et al., 2011, Wisniewski et al., 2010). The results may help to identify further markers of disease states in neurons or elsewhere.

Materials and Methods

Protein preparation and purification
Tau protein (clone htau40, largest isoform in human CNS, 441 residues, Uniprot Id P10636 isoform F) expressed in *E. coli* or Sf9 cells, prepared and purified as described in (Tepper et al., 2014). Expression in either system yielded Tau concentrations up to ~230 µM without causing aggregation. In the case of Sf9 cells the expressed Tau protein was purified from cell extracts by making use of the heat stability of Tau. Sf9 Cells were incubated for 3 days at 27 °C and collected directly for preparation of phosphorylated hTau40 protein in lysis buffer (50 mM Tris-HCl, pH 7.4, 500 mM NaCl, 10% glycerol, 1% Nonidet P-40, 5mM DTT, 10mM EGTA, 20mM NaF, 1 mM orthovanadate, 5 µM microcystin, 10 µg/ml each of

protease inhibitors leupeptin, aprotinin, and pepstatin) in a ratio of 1 g of Sf9 pellet to 10 ml of lysis buffer. This procedure yielded "P_m-Tau" (alias P12-Tau in Tepper, 2014; or P-Tau in Mair, 2016). To increase the phosphorylation even further, Sf9 cells were treated before harvesting for 1 h with 0.2 µM okadaic acid (OA; a phosphatase inhibitor, Enzo-LifeScience). Next, after centrifugation the cells were resuspended in lysis buffer and boiled in a water bath at 100°C for 10 min. By this treatment nearly all proteins were denatured and precipitated, except for Tau which stays soluble. The cell debris was removed by centrifugation.
The supernatant containing soluble Tau protein was concentrated in Millipore Amicon Ultra-4-centrifugal filter units (molecular mass cutoff of 3 kDa). This procedure yielded "P_h-Tau" (alias P20-Tau in (Tepper et al., 2014), or PP-Tau in (Mair et al., 2016)). The material was then applied to a size exclusion column Superdex G200 (GE Healthcare) and eluted with PBS buffer (pH 7.4; 1 mM DTT). For further experiments, the fractions containing Tau protein were pooled and concentrated again in Amicon Ultra-4-centrifugal filter units. Finally, the concentrated protein was rebuffered into 200 mM ammonium acetate, pH 7,6.

Analysis of human Tau proteins by native MS

Purified proteins in 200 mM ammonium acetate, pH 7.6 were analyzed by nanoflow electrospray ionization mass spectrometry using a Synapt HDMS (Waters and MS Vision) equipped with a 32,000 *m/z* range quadrupole. 3 µl of sample were introduced with an in-house manufactured gold-coated capillary needle (borosilicate thin wall with filament, OD 1.0 mm, ID 0.78 mm; Harvard Apparatus). The capillary voltage was set to 1.3 kV and the cone voltage to 180 V. The trap cell was filled with argon gas at a flow rate of 1 ml/min and collisional activation was performed by applying acceleration voltages to the trap cell. The transfer cell was kept at a pressure of about 20% of that in the trap cell and a low acceleration voltage of 5 V. Spectra were calibrated externally using caesium iodide. Charge state spectra were deconvoluted by the program UniDec version 2.7.3 (Marty et al., 2015).

Identification of peptide sequences and posttranslational modifications by LC-MS/MS

In-solution tryptic digestion of proteins was performed as described previously (Morgenstern et al., 2017). In brief, Tau protein samples were subjected to reduction of cysteine residues with 5 mM Tris(2-carboxyethyl)phosphine, alkylation of free thiol groups with 50 mM iodoacetamide/50 mM NH_4HCO_3 and digested with trypsin overnight at 37°C. Peptide mixtures were analyzed by nano-HPLC-ESI-MS/MS using a Q Exactive instrument directly coupled to an UltiMate 3000 RSLCnano HPLC system (both Thermo Fisher Scientific, Dreieich, Germany) equipped with a Nanospray Flex ion source with DirectJunction (Thermo Fisher Scientific). Peptides were washed and preconcentrated on PepMap™ C18 precolumns (5 mm x 300 µm inner diameter; Thermo Scientific) and separated using AcclaimPepMap™ RSLC columns (50 cm x 75 µm inner diameter; pore size, 100 Å; particle size, 2 µm) at a flow rate of 250 nl/min and 40 - 43°C. For peptide elution, binary solvent systems were used consisting of 0.1% (v/v) formic acid (FA) (solvent A) and 0.1% (v/v) FA/86% (v/v) acetonitrile (solvent B) with a gradient of 4- 42% solvent B in 50 min, 42 - 95% in 5 min and 5 min at 95%. The Q Exactive was operated with the following parameters: MS survey scans ranging from *m/z* 375 - 1,700 at a resolution (R) of 70,000 (at *m/z* 200), an automatic gain control (AGC) of 3×10^6 ions, and a maximum injection time (IT) of 60 ms. A TOP12 method was used for higher energy collisional dissociation (HCD) of precursor peptides ($z \geq 2$) in the orbitrap applying a normalized collision energy of 28%, an AGC of 1×10^5 ions and max. IT of 120 ms. The dynamic exclusion time for previously fragmented precursors was set to 45 s.
Data were analysed using Maxquant software version 1.5.5.1 (Tyanova et al., 2016) searching against the sequence of human Tau-F (Uniprot Id P10636-8) and UniProt organism specific sequence databases for *E. coli* and *Spodoptera frugiperda* (version 2018_08, 4,344 and 26,434 entries, respectively) allowing a maximum of four missed cleavages of trypsin, a mass tolerance of 4.5 ppm for precursor and 20 ppm for fragment ions. Methionine oxidation, acetylation of lysine and protein N-terminus as well as phosphorylation at serine, threonine and tyrosine residues were specified as variable modifications and carbamidomethylation of cysteine as fixed modification. Additional

searches were performed with an extended range of variable modifications including ubiquitination (diglycine at lysine) and O-GlcNAc modification of serine and threonine. However, these modifications were observed only sporadically and in low intensity and borderline significance and were therefore excluded from the final search. The lists of both peptide and protein identifications were filtered applying a threshold for the false discovery rate of <0.01.

Conflict of interest
The authors declare that they have no conflicts of interest with the contents of this article.

References:

AMNIAI, L., LIPPENS, G. & LANDRIEU, I. 2011. Characterization of the AT180 epitope of phosphorylated Tau protein by a combined nuclear magnetic resonance and fluorescence spectroscopy approach. *Biochem Biophys Res Commun,* 412**,** 743-6.

ARAKHAMIA, T., LEE, C. E., CARLOMAGNO, Y., DUONG, D. M., KUNDINGER, S. R., WANG, K., WILLIAMS, D., DETURE, M., DICKSON, D. W., COOK, C. N., SEYFRIED, N. T., PETRUCELLI, L. & FITZPATRICK, A. W. P. 2020. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell,* 180**,** 633-644 e12.

ARENDT, T., STIELER, J., STRIJKSTRA, A. M., HUT, R. A., RUDIGER, J., VAN DER ZEE, E. A., HARKANY, T., HOLZER, M. & HARTIG, W. 2003. Reversible paired helical filament-like phosphorylation of tau is an adaptive process associated with neuronal plasticity in hibernating animals. *J Neurosci,* 23, 6972-81.

ARNSTEN, A. F. T., DATTA, D., LESLIE, S., YANG, S. T., WANG, M. & NAIRN, A. C. 2019. Alzheimer's-like pathology in aging rhesus macaques: Unique opportunity to study the etiology and treatment of Alzheimer's disease. *Proc Natl Acad Sci U S A.*

AUGUSTINACK, J. C., SCHNEIDER, A., MANDELKOW, E. M. & HYMAN, B. T. 2002. Specific tau phosphorylation sites correlate with severity of neuronal cytopathology in Alzheimer's disease. *Acta Neuropathol,* 103**,** 26-35.

BAAS, P. W., PIENKOWSKI, T. P. & KOSIK, K. S. 1991. Processes induced by tau expression in Sf9 cells have an axon-like microtubule organization. *J Cell Biol,* 115, 1333-44.

BARTHELEMY, N. R., FENAILLE, F., HIRTZ, C., SERGEANT, N., SCHRAEN-MASCHKE, S., VIALARET, J., BUEE, L., GABELLE, A., JUNOT, C., LEHMANN, S. & BECHER, F. 2016. Tau Protein Quantification in Human Cerebrospinal Fluid by Targeted Mass Spectrometry at High Sequence Coverage Provides Insights into Its Primary Structure Heterogeneity. *J Proteome Res,* 15, 667-76.

BARTHELEMY, N. R., LI, Y., JOSEPH-MATHURIN, N., GORDON, B. A., HASSENSTAB, J., BENZINGER, T. L. S., BUCKLES, V., FAGAN, A. M., PERRIN, R. J., GOATE, A. M., MORRIS, J. C., KARCH, C. M., XIONG, C., ALLEGRI, R., MENDEZ, P. C., BERMAN, S. B., IKEUCHI, T., MORI, H., SHIMADA, H., SHOJI, M., SUZUKI, K., NOBLE, J., FARLOW, M., CHHATWAL, J., GRAFF-RADFORD, N. R., SALLOWAY, S., SCHOFIELD, P. R., MASTERS, C. L., MARTINS, R. N., O'CONNOR, A., FOX, N. C., LEVIN, J., JUCKER, M., GABELLE, A., LEHMANN, S., SATO, C., BATEMAN, R. J., MCDADE, E. & DOMINANTLY INHERITED ALZHEIMER, N. 2020. A soluble phosphorylated tau signature links tau, amyloid and the evolution of stages of dominantly inherited Alzheimer's disease. *Nat Med,* 26**,** 398-407.

BARTHELEMY, N. R., MALLIPEDDI, N., MOISEYEV, P., SATO, C. & BATEMAN, R. J. 2019. Tau Phosphorylation Rates Measured by Mass Spectrometry Differ in the Intracellular Brain vs. Extracellular Cerebrospinal Fluid Compartments and Are Differentially Affected by Alzheimer's Disease. *Front Aging Neurosci,* 11**,** 121.

BATEMAN, R. J., BARTHELEMY, N. R. & HORIE, K. 2020. Another step forward in blood-based diagnostics for Alzheimer's disease. *Nat Med,* 26**,** 314-316.

BIERNAT, J. & MANDELKOW, E. M. 1999. The development of cell processes induced by tau protein requires phosphorylation of serine 262 and 356 in the repeat domain and is inhibited by phosphorylation in the proline-rich domains. *Mol Biol Cell,* 10**,** 727-40.

BIERNAT, J., WU, Y. Z., TIMM, T., ZHENG-FISCHHOFER, Q., MANDELKOW, E., MEIJER, L. & MANDELKOW, E. M. 2002. Protein kinase MARK/PAR-1 is required for neurite outgrowth and establishment of neuronal polarity. *Mol Biol Cell,* 13**,** 4013-28.

BOERI ERBA, E. & PETOSA, C. 2015. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci,* 24**,** 1176-92.

BOYLE, W. J., VAN DER GEER, P. & HUNTER, T. 1991. Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol,* 201**,** 110-49.

BRAAK, E., BRAAK, H. & MANDELKOW, E. M. 1994. A sequence of cytoskeleton changes related to the formation of neurofibrillary tangles and neuropil threads. *Acta Neuropathol,* 87**,** 554-67.

CHEN, R. & HENNING, U. 1996. A periplasmic protein (Skp) of Escherichia coli selectively binds a class of outer membrane proteins. *Mol Microbiol,* 19**,** 1287-94.

COHEN, T. J., GUO, J. L., HURTADO, D. E., KWONG, L. K., MILLS, I. P., TROJANOWSKI, J. Q. & LEE, V. M. 2011. The acetylation of tau inhibits its function and promotes pathological tau aggregation. *Nat Commun,* 2**,** 252.

COPPOLA, G., CHINNATHAMBI, S., LEE, J. J., DOMBROSKI, B. A., BAKER, M. C., SOTO-ORTOLAZA, A. I., LEE, S. E., KLEIN, E., HUANG, A. Y., SEARS, R., LANE, J. R., KARYDAS, A. M., KENET, R. O., BIERNAT, J., WANG, L. S., COTMAN, C. W., DECARLI, C. S., LEVEY, A. I., RINGMAN, J. M., MENDEZ, M. F., CHUI, H. C., LE BER, I., BRICE, A., LUPTON, M. K., PREZA, E., LOVESTONE, S., POWELL, J., GRAFF-RADFORD, N., PETERSEN, R. C., BOEVE, B. F., LIPPA, C. F., BIGIO, E. H., MACKENZIE, I., FINGER, E., KERTESZ, A., CASELLI, R. J., GEARING, M., JUNCOS, J. L., GHETTI, B., SPINA, S., BORDELON, Y. M., TOURTELLOTTE, W. W., FROSCH, M. P., VONSATTEL, J. P., ZAROW, C., BEACH, T. G., ALBIN, R. L., LIEBERMAN, A. P., LEE, V. M., TROJANOWSKI, J. Q., VAN DEERLIN, V. M., BIRD, T. D., GALASKO, D. R., MASLIAH, E., WHITE, C. L., TRONCOSO, J. C., HANNEQUIN, D., BOXER, A. L., GESCHWIND, M. D., KUMAR, S., MANDELKOW, E. M., WSZOLEK, Z. K., UITTI, R. J., DICKSON, D. W., HAINES, J. L., MAYEUX, R., PERICAK-VANCE, M. A., FARRER, L. A., ALZHEIMER'S DISEASE GENETICS, C., ROSS, O. A., RADEMAKERS, R., SCHELLENBERG, G. D., MILLER, B. L., MANDELKOW, E. & GESCHWIND, D. H.

2012. Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer's diseases. *Hum Mol Genet,* 21**,** 3500-12.

COWAN, N. J. & LEWIS, S. A. 2001. Type II chaperonins, prefoldin, and the tubulin-specific chaperones. *Adv Protein Chem,* 59**,** 73-104.

DERKINDEREN, P., SCALES, T. M., HANGER, D. P., LEUNG, K. Y., BYERS, H. L., WARD, M. A., LENZ, C., PRICE, C., BIRD, I. N., PERERA, T., KELLIE, S., WILLIAMSON, R., NOBLE, W., VAN ETTEN, R. A., LEROY, K., BRION, J. P., REYNOLDS, C. H. & ANDERTON, B. H. 2005. Tyrosine 394 is phosphorylated in Alzheimer's paired helical filament tau and in fetal tau with c-Abl as the candidate tyrosine kinase. *J Neurosci,* 25**,** 6584-93.

DETURE, M. A. & DICKSON, D. W. 2019. The neuropathological diagnosis of Alzheimer's disease. *Mol Neurodegener,* 14**,** 32.

DEVOS, S. L., MILLER, R. L., SCHOCH, K. M., HOLMES, B. B., KEBODEAUX, C. S., WEGENER, A. J., CHEN, G., SHEN, T., TRAN, H., NICHOLS, B., ZANARDI, T. A., KORDASIEWICZ, H. B., SWAYZE, E. E., BENNETT, C. F., DIAMOND, M. I. & MILLER, T. M. 2017. Tau reduction prevents neuronal loss and reverses pathological tau deposition and seeding in mice with tauopathy. *Sci Transl Med,* 9**,** 10.

DONNELLY, D. P., RAWLINS, C. M., DEHART, C. J., FORNELLI, L., SCHACHNER, L. F., LIN, Z., LIPPENS, J. L., ALURI, K. C., SARIN, R., CHEN, B., LANTZ, C., JUNG, W., JOHNSON, K. R., KOLLER, A., WOLFF, J. J., CAMPUZANO, I. D. G., AUCLAIR, J. R., IVANOV, A. R., WHITELEGGE, J. P., PASA-TOLIC, L., CHAMOT-ROOKE, J., DANIS, P. O., SMITH, L. M., TSYBIN, Y. O., LOO, J. A., GE, Y., KELLEHER, N. L. & AGAR, J. N. 2019. Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat Methods,* 16**,** 587-594.

DRUBIN, D. G., FEINSTEIN, S. C., SHOOTER, E. M. & KIRSCHNER, M. W. 1985. Nerve growth factor-induced neurite outgrowth in PC12 cells involves the coordinate induction of microtubule assembly and assembly-promoting factors. *J Cell Biol,* 101**,** 1799-807.

FLORES, J., NOEL, A., FOVEAU, B., LYNHAM, J., LECRUX, C. & LEBLANC, A. C. 2018. Caspase-1 inhibition alleviates cognitive impairment and neuropathology in an Alzheimer's disease mouse model. *Nat Commun,* 9**,** 3916.

GOEDERT, M., JAKES, R. & VANMECHELEN, E. 1995. Monoclonal antibody AT8 recognises tau protein phosphorylated at both serine 202 and threonine 205. *Neurosci Lett,* 189**,** 167-9.

GRUNDKE-IQBAL, I., IQBAL, K., TUNG, Y. C., QUINLAN, M., WISNIEWSKI, H. M. & BINDER, L. I. 1986. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci U S A,* 83**,** 4913-7.

GUNAWARDANA, C. G., MEHRABIAN, M., WANG, X., MUELLER, I., LUBAMBO, I. B., JONKMAN, J. E., WANG, H. & SCHMITT-ULMS, G. 2015. The Human Tau Interactome: Binding to the Ribonucleoproteome, and Impaired Binding of the Proline-to-Leucine Mutant at Position 301 (P301L) to Chaperones and the Proteasome. *Mol Cell Proteomics,* 14**,** 3000-14.

GUO, X., HUANG, X. & CHEN, M. J. 2017. Reversible phosphorylation of the 26S proteasome. *Protein Cell,* 8**,** 255-272.

GUSTKE, N., STEINER, B., MANDELKOW, E. M., BIERNAT, J., MEYER, H. E., GOEDERT, M. & MANDELKOW, E. 1992. The Alzheimer-like phosphorylation of tau protein reduces microtubule binding and involves Ser-Pro and Thr-Pro motifs. *FEBS Lett,* 307**,** 199-205.

HANGER, D. P. 2020. Table of Tau phosphorylation sites. *Web page of King's college London*.

HANGER, D. P., ANDERTON, B. H. & NOBLE, W. 2009. Tau phosphorylation: the therapeutic challenge for neurodegenerative disease. *Trends Mol Med,* 15**,** 112-9.

HART, G. W., SLAWSON, C., RAMIREZ-CORREA, G. & LAGERLOF, O. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem,* 80**,** 825-58.

HASEGAWA, M., MORISHIMA-KAWASHIMA, M., TAKIO, K., SUZUKI, M., TITANI, K. & IHARA, Y. 1992. Protein sequence and mass spectrometric analyses of tau in the Alzheimer's disease brain. *J Biol Chem,* 267**,** 17047-54.

HERNANDEZ, H. & ROBINSON, C. V. 2007. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc,* 2**,** 715-26.

HOFFMAN, A., TALESKI, G. & SONTAG, E. 2017. The protein serine/threonine phosphatases PP2A, PP1 and calcineurin: A triple threat in the regulation of the neuronal cytoskeleton. *Mol Cell Neurosci,* 84**,** 119-131.

ILLENBERGER, S., ZHENG-FISCHHOFER, Q., PREUSS, U., STAMER, K., BAUMANN, K., TRINCZEK, B., BIERNAT, J., GODEMANN, R., MANDELKOW, E. M. & MANDELKOW, E. 1998. The endogenous and cell cycle-dependent phosphorylation of tau protein in living cells: implications for Alzheimer's disease. *Mol Biol Cell,* 9**,** 1495-512.

ITTNER, L. M., KE, Y. D., DELERUE, F., BI, M., GLADBACH, A., VAN EERSEL, J., WOLFING, H., CHIENG, B. C., CHRISTIE, M. J., NAPIER, I. A., ECKERT, A., STAUFENBIEL, M., HARDEMAN, E. & GOTZ, J. 2010. Dendritic function of tau mediates amyloid-beta toxicity in Alzheimer's disease mouse models. *Cell,* 142**,** 387-97.

JAKES, R., NOVAK, M., DAVISON, M. & WISCHIK, C. M. 1991. Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's disease. *EMBO J,* 10**,** 2725-9.

JEGANATHAN, S., HASCHER, A., CHINNATHAMBI, S., BIERNAT, J., MANDELKOW, E. M. & MANDELKOW, E. 2008. Proline-directed pseudo-phosphorylation at AT8 and PHF1 epitopes induces a compaction of the paperclip folding of Tau and generates a pathological (MC-1) conformation. *J Biol Chem,* 283**,** 32066-76.

KANIYAPPAN, S., CHANDUPATLA, R. R., MANDELKOW, E. M. & MANDELKOW, E. 2017. Extracellular low-n oligomers of tau cause selective synaptotoxicity without affecting cell viability. *Alzheimers Dement,* 13**,** 1270-1291.

KARAGOZ, G. E., DUARTE, A. M., AKOURY, E., IPPEL, H., BIERNAT, J., MORAN LUENGO, T., RADLI, M., DIDENKO, T., NORDHUES, B. A., VEPRINTSEV, D. B., DICKEY, C. A., MANDELKOW, E., ZWECKSTETTER, M., BOELENS, R., MADL, T. & RUDIGER, S. G. 2014. Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. *Cell,* 156**,** 963-74.

KENESSEY, A. & YEN, S. H. 1993. The extent of phosphorylation of fetal tau is comparable to that of PHF-tau from Alzheimer paired helical filaments. *Brain Res,* 629**,** 40-6.

KNOPS, J., KOSIK, K. S., LEE, G., PARDEE, J. D., COHEN-GOULD, L. & MCCONLOGUE, L. 1991. Overexpression of tau in a nonneuronal cell induces long cellular processes. *J Cell Biol,* 114**,** 725-33.

KOPKE, E., TUNG, Y. C., SHAIKH, S., ALONSO, A. C., IQBAL, K. & GRUNDKE-IQBAL, I. 1993. Microtubule-associated protein tau. Abnormal phosphorylation of a non-paired helical filament pool in Alzheimer disease. *J Biol Chem,* 268**,** 24374-84.

KOREN, S. A., HAMM, M. J., MEIER, S. E., WEISS, B. E., NATION, G. K., CHISHTI, E. A., ARANGO, J. P., CHEN, J., ZHU, H., BLALOCK, E. M. & ABISAMBRA, J. F. 2019. Tau drives translational selectivity by interacting with ribosomal proteins. *Acta Neuropathol,* 137**,** 571-583.

KOSIK, K. S. & SHIMURA, H. 2005. Phosphorylated tau and the neurodegenerative foldopathies. *Biochim Biophys Acta,* 1739**,** 298-310.

KSIEZAK-REDING, H., LIU, W. K. & YEN, S. H. 1992. Phosphate analysis and dephosphorylation of modified tau associated with paired helical filaments. *Brain Res,* 597**,** 209-19.

LASAGNA-REEVES, C. A., CASTILLO-CARRANZA, D. L., SENGUPTA, U., GUERRERO-MUNOZ, M. J., KIRITOSHI, T., NEUGEBAUER, V., JACKSON, G. R. & KAYED, R. 2012. Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau. *Sci Rep,* 2**,** 700.

LIPPENS, G., LANDRIEU, I., SMET, C., HUVENT, I., GANDHI, N. S., GIGANT, B., DESPRES, C., QI, H. & LOPEZ, J. 2016. NMR Meets Tau: Insights into Its Function and Pathology. *Biomolecules,* 6.

MAIR, W., MUNTEL, J., TEPPER, K., TANG, S., BIERNAT, J., SEELEY, W. W., KOSIK, K. S., MANDELKOW, E., STEEN, H. & STEEN, J. A. 2016. FLEXITau: Quantifying Post-translational Modifications of Tau Protein in Vitro and in Human Disease. *Anal Chem,* 88**,** 3704-14.

MANDELKOW, E. M. & MANDELKOW, E. 2012. Biochemistry and cell biology of tau protein in neurofibrillary degeneration. *Cold Spring Harb Perspect Med,* 2**,** a006247.

MARTIN, L., LATYPOVA, X., WILSON, C. M., MAGNAUDEIX, A., PERRIN, M. L., YARDIN, C. & TERRO, F. 2013. Tau protein kinases: involvement in Alzheimer's disease. *Ageing Res Rev,* 12**,** 289-309.

MARTY, M. T., BALDWIN, A. J., MARKLUND, E. G., HOCHBERG, G. K., BENESCH, J. L. & ROBINSON, C. V. 2015. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal Chem,* 87**,** 4370-6.

MATSUO, E. S., SHIN, R. W., BILLINGSLEY, M. L., VAN DEVOORDE, A., O'CONNOR, M., TROJANOWSKI, J. Q. & LEE, V. M. 1994. Biopsy-derived adult human brain tau is phosphorylated at many of the same sites as Alzheimer's disease paired helical filament tau. *Neuron,* 13**,** 989-1002.

MAZIUK, B., BALLANCE, H. I. & WOLOZIN, B. 2017. Dysregulation of RNA Binding Protein Aggregation in Neurodegenerative Disorders. *Front Mol Neurosci,* 10**,** 89.

MEIER, S., BELL, M., LYONS, D. N., RODRIGUEZ-RIVERA, J., INGRAM, A., FONTAINE, S. N., MECHAS, E., CHEN, J., WOLOZIN, B., LEVINE, H., 3RD, ZHU, H. & ABISAMBRA, J. F. 2016. Pathological Tau Promotes Neuronal Damage by Impairing Ribosomal Function and Decreasing Protein Synthesis. *J Neurosci,* 36**,** 1001-7.

MIKULIK, K., BOBEK, J., ZIKOVA, A., SMETAKOVA, M. & BEZOUSKOVA, S. 2011. Phosphorylation of ribosomal proteins influences subunit association and translation of poly (U) in Streptomyces coelicolor. *Mol Biosyst,* 7**,** 817-23.

MIN, S. W., CHO, S. H., ZHOU, Y., SCHROEDER, S., HAROUTUNIAN, V., SEELEY, W. W., HUANG, E. J., SHEN, Y., MASLIAH, E., MUKHERJEE, C., MEYERS, D., COLE, P. A., OTT, M. & GAN, L. 2010. Acetylation of tau inhibits its degradation and contributes to tauopathy. *Neuron,* 67**,** 953-66.

MORGENSTERN, M., STILLER, S. B., LUBBERT, P., PEIKERT, C. D., DANNENMAIER, S., DREPPER, F., WEILL, U., HOSS, P., FEUERSTEIN, R., GEBERT, M., BOHNERT, M., VAN DER LAAN, M., SCHULDINER, M., SCHUTZE, C., OELJEKLAUS, S., PFANNER, N., WIEDEMANN, N. & WARSCHEID, B. 2017. Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell Rep,* 19**,** 2836-2852.

MORRIS, M., KNUDSEN, G. M., MAEDA, S., TRINIDAD, J. C., IOANOVICIU, A., BURLINGAME, A. L. & MUCKE, L. 2015. Tau post-translational modifications in wild-type and human amyloid precursor protein transgenic mice. *Nat Neurosci,* 18**,** 1183-9.

MUKRASCH, M. D., BIBOW, S., KORUKOTTU, J., JEGANATHAN, S., BIERNAT, J., GRIESINGER, C., MANDELKOW, E. & ZWECKSTETTER, M. 2009. Structural polymorphism of 441-residue tau at single residue resolution. *PLoS Biol,* 7**,** e34.

MULTHAUP, G., HUBER, O., BUEE, L. & GALAS, M. C. 2015. Amyloid Precursor Protein (APP) Metabolites APP Intracellular Fragment (AICD), Abeta42, and Tau in Nuclear Roles. *J Biol Chem,* 290**,** 23515-22.

OKA, T., TAGAWA, K., ITO, H. & OKAZAWA, H. 2011. Dynamic changes of the phosphoproteome in postmortem mouse brains. *PLoS One,* 6**,** e21405.

OTVOS, L., JR., FEINER, L., LANG, E., SZENDREI, G. I., GOEDERT, M. & LEE, V. M. 1994. Monoclonal antibody PHF-1 recognizes tau protein phosphorylated at serine residues 396 and 404. *J Neurosci Res,* 39**,** 669-73.

PEREZ-RIVEROL, Y., CSORDAS, A., BAI, J., BERNAL-LLINARES, M., HEWAPATHIRANA, S., KUNDU, D. J., INUGANTI, A., GRISS, J., MAYER, G., EISENACHER, M., PEREZ, E., USZKOREIT, J., PFEUFFER, J., SACHSENBERG, T., YILMAZ, S., TIWARY, S., COX, J., AUDAIN, E., WALZER, M., JARNUCZAK, A. F., TERNENT, T., BRAZMA, A. & VIZCAINO, J. A. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res,* 47**,** D442-D450.

PLANEL, E., RICHTER, K. E., NOLAN, C. E., FINLEY, J. E., LIU, L., WEN, Y., KRISHNAMURTHY, P., HERMAN, M., WANG, L., SCHACHTER, J. B., NELSON, R. B., LAU, L. F. & DUFF, K. E. 2007. Anesthesia leads to tau hyperphosphorylation through inhibition of phosphatase activity by hypothermia. *J Neurosci,* 27, 3090-7.

SATO, C., BARTHELEMY, N. R., MAWUENYEGA, K. G., PATTERSON, B. W., GORDON, B. A., JOCKEL-BALSAROTTI, J., SULLIVAN, M., CRISP, M. J., KASTEN, T., KIRMESS, K. M., KANAAN, N. M., YARASHESKI, K. E., BAKER-NIGH, A., BENZINGER, T. L. S., MILLER, T. M., KARCH, C. M. & BATEMAN, R. J. 2018. Tau Kinetics in Neurons and the Human Central Nervous System. *Neuron,* 97**,** 1284-1298 e7.

SCHAFFER, L. V., MILLIKIN, R. J., MILLER, R. M., ANDERSON, L. C., FELLERS, R. T., GE, Y., KELLEHER, N. L., LEDUC, R. D., LIU, X., PAYNE, S. H., SUN, L., THOMAS, P. M., TUCHOLSKI, T., WANG, Z., WU, S., WU, Z., YU, D., SHORTREED, M. R. & SMITH, L. M. 2019. Identification and Quantification of Proteoforms by Mass Spectrometry. *Proteomics,* 19, e1800361.

SCHNEIDER, A., BIERNAT, J., VON BERGEN, M., MANDELKOW, E. & MANDELKOW, E. M. 1999. Phosphorylation that detaches tau protein from microtubules (Ser262, Ser214) also protects it against aggregation into Alzheimer paired helical filaments. *Biochemistry,* 38, 3549-58.

SCHNEIDER, A. & MANDELKOW, E. 2008. Tau-based treatment strategies in neurodegenerative diseases. *Neurotherapeutics,* 5**,** 443-57.

SCHWALBE, M., BIERNAT, J., BIBOW, S., OZENNE, V., JENSEN, M. R., KADAVATH, H., BLACKLEDGE, M., MANDELKOW, E. & ZWECKSTETTER, M. 2013. Phosphorylation of human Tau protein by microtubule affinity-regulating kinase 2. *Biochemistry,* 52**,** 9068-79.

SHARMA, K., KUMAR, A. & CHANDNA, S. 2016. Constitutive hyperactivity of histone deacetylases enhances radioresistance in Lepidopteran Sf9 insect cells. *Biochim Biophys Acta,* 1860**,** 1237-46.

SIGURDSSON, E. M. 2018. Tau Immunotherapies for Alzheimer's Disease and Related Tauopathies: Progress and Potential Pitfalls. *J Alzheimers Dis,* 66**,** 855-856.

SOHN, P. D., TRACY, T. E., SON, H. I., ZHOU, Y., LEITE, R. E., MILLER, B. L., SEELEY, W. W., GRINBERG, L. T. & GAN, L. 2016. Acetylated tau destabilizes the cytoskeleton in the axon initial segment and is mislocalized to the somatodendritic compartment. *Mol Neurodegener,* 11**,** 47.

TEPPER, K., BIERNAT, J., KUMAR, S., WEGMANN, S., TIMM, T., HUBSCHMANN, S., REDECKE, L., MANDELKOW, E. M., MULLER, D. J. & MANDELKOW, E. 2014. Oligomer Formation of Tau Hyperphosphorylated in Cells. *J Biol Chem,* 289, 34389-34407.

TYANOVA, S., TEMU, T. & COX, J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc,* 11, 2301-2319.

VON DER OHE, C. G., GARNER, C. C., DARIAN-SMITH, C. & HELLER, H. C. 2007. Synaptic protein dynamics in hibernation. *J Neurosci,* 27, 84-92.

WANG, Y. & MANDELKOW, E. 2016. Tau in physiology and pathology. *Nat Rev Neurosci,* 17, 5-21.

WANG, Y., ZHANG, Y., HU, W., XIE, S., GONG, C. X., IQBAL, K. & LIU, F. 2015. Rapid alteration of protein phosphorylation during postmortem: implication in the study of protein phosphorylation. *Sci Rep,* 5**,** 15709.

WANG, Y. P., BIERNAT, J., PICKHARDT, M., MANDELKOW, E. & MANDELKOW, E. M. 2007. Stepwise proteolysis liberates tau fragments that nucleate the Alzheimer-like aggregation of full-length tau in a neuronal cell model. *Proc Natl Acad Sci U S A,* 104**,** 10252-7.

WATANABE, A., HONG, W. K., DOHMAE, N., TAKIO, K., MORISHIMA-KAWASHIMA, M. & IHARA, Y. 2004. Molecular aging of tau: disulfide-independent aggregation and non-enzymatic degradation in vitro and in vivo. *J Neurochem,* 90**,** 1302-11.

WEGMANN, S., EFTEKHARZADEH, B., TEPPER, K., ZOLTOWSKA, K. M., BENNETT, R. E., DUJARDIN, S., LASKOWSKI, P. R., MACKENZIE, D., KAMATH, T., COMMINS, C., VANDERBURG, C., ROE, A. D., FAN, Z., MOLLIEX, A. M., HERNANDEZ-VEGA, A., MULLER, D., HYMAN, A. A., MANDELKOW, E., TAYLOR, J. P. & HYMAN, B. T. 2018. Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J,* 37.

WHEELER, J. M., MCMILLAN, P., STROVAS, T. J., LIACHKO, N. F., AMLIE-WOLF, A., KOW, R. L., KLEIN, R. L., SZOT, P., ROBINSON, L., GUTHRIE, C., SAXTON, A., KANAAN, N. M., RASKIND, M., PESKIND, E., TROJANOWSKI, J. Q., LEE, V. M. Y., WANG, L. S., KEENE, C. D., BIRD, T., SCHELLENBERG, G. D. & KRAEMER, B. 2019. Activity of the poly(A) binding protein MSUT2 determines susceptibility to pathological tau in the mammalian brain. *Sci Transl Med,* 11.

WISNIEWSKI, J. R., NAGARAJ, N., ZOUGMAN, A., GNAD, F. & MANN, M. 2010. Brain phosphoproteome obtained by a FASP-based method reveals plasma membrane protein topology. *J Proteome Res,* 9**,** 3280-9.

YANAMANDRA, K., KFOURY, N., JIANG, H., MAHAN, T. E., MA, S., MALONEY, S. E., WOZNIAK, D. F., DIAMOND, M. I. & HOLTZMAN, D. M. 2013. Anti-tau antibodies that block tau aggregate seeding in vitro markedly decrease pathology and improve cognition in vivo. *Neuron,* 80**,** 402-14.

YU, A., FOX, S. G., CAVALLINI, A., KERRIDGE, C., O'NEILL, M. J., WOLAK, J., BOSE, S. & MORIMOTO, R. I. 2019. Tau protein aggregates inhibit the protein-folding and vesicular trafficking arms of the cellular proteostasis network. *J Biol Chem,* 294**,** 7917-7930.

YUZWA, S. A., SHAN, X., MACAULEY, M. S., CLARK, T., SKOROBOGATKO, Y., VOSSELLER, K. & VOCADLO, D. J. 2012. Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nat Chem Biol,* 8**,** 393-9.

ZHENG-FISCHHOFER, Q., BIERNAT, J., MANDELKOW, E. M., ILLENBERGER, S., GODEMANN, R. & MANDELKOW, E. 1998. Sequential phosphorylation of Tau by glycogen synthase kinase-3beta and protein

kinase A at Thr212 and Ser214 generates the Alzheimer-specific epitope of antibody AT100 and requires a paired-helical-filament-like conformation. *Eur J Biochem,* 252**,** 542-52.

Abbreviations and definitions:

htau40, Tau-2N4R	largest Tau isoform in human CNS, 441 residues
P_i	Phosphate group (on Ser, Thr, or Tyr)
MS	mass spectrometry
LC-MS/MS	Liquid chromatography-mass spectrometry
MALDI TOF-MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
Sf9 cells	*Spodoptera frugiperda* cells
CSF	Cerebrospinal fluid
HDAC	Histone deacetylase
PTM	posttranslational modification
site occupancy:	number of modifications (e.g. phosphorylation) at a given residue position, relative to number of molecules (max. 100%)
overall occupancy:	total number of modifications of a certain type (e.g. phosphorylation) per tau molecule (up to ~20)

Table 1: Phosphorylation sites on Tau protein (fractions P$_m$-Tau and P$_h$-Tau) identified by LC-MS/MS. Note that phosphorylation is observed only on Ser or Thr, not on Tyr.

Peptide No.	Residues	P-Site	Antibody against P-epitope	Peptide sequences
1	6-24	T17		QEFEVMEDHAG**T**YGLGDRK
2	25-44	T39		DQGGYTMHQDQEGD**T**DAGLK
3	24-67	S46		KDQGGYTMHQDQEGDTDAGLKE**S**PLQTPTEDGSEEPGSETSDAK
4	24-67	T50		KDQGGYTMHQDQEGDTDAGLKESPLQ**T**PTEDGSEEPGSETSDAK
5	45-67	T52		ESPLQTP**T**EDGSEEPGSETSDAK
6	45-67	S61		ESPLQTPTEDGSEEPG**S**ETSDAK
7	68-87	T71		STP**T**AEDVTAPLVDEGAPGK
8	68-87	T76		STPTAEDV**T**APLVDEGAPGK
9	88-126	T111		QAAAQPHTEIPEGTTAEEAGIGD**T**PSLEDEAAGHVTQAR
10	127-143	S131		MVSK**S**KDGTGSDDKKAK
11	164-174	T169		GQANA**T**RIPAK
12	171-194	T175		IPAK**T**PPAPKTPPSSGEPPKSGDR
13	175-194	T181	AT270	TPPAPK**T**PPSSGEPPKSGDR
14	175-194	S184		TPPAPKTPP**S**SGEPPKSGDR
15	175-194	S185		TPPAPKTPPS**S**GEPPKSGDR
16	195-209	S199		SGYS**S**PGSPGTPGSR
17	195-209	S202	AT8	SGYSSPG**S**PGTPGSR
18	195-221	T205	AT8	SGYSSPGSPG**T**PGSRSRTPSLPTPPTR
19	195-221	S208		SGYSSPGSPGTPG**S**RSRTPSLPTPPTR
20	210-224	S210		**S**RTPSLPTPPTREPK
21	210-221	T212	AT100	SR**T**PSLPTPPTR
22	212-221	S214	AT100	TP**S**LPTPPTR
23	212-224	T217		TPSLP**T**PPTREPK
24	226-240	T231	AT180	VAVVR**T**PPKSPSSAK
25	226-242	S235		VAVVRTPPK**S**PSSAKSR
26	225-240	S237		KVAVVRTPPKSP**S**SAK
27	225-242	S238		KVAVVRTPPKSPS**S**AKSR
28	243-254	T245		LQ**T**APVPMPDLK
29	255-267	S258		NVK**S**KIGSTENLK
30	260-267	S262	12E8	IG**S**TENLK
31	281-290	S285		KLDL**S**NVQSK
32	281-290	S289		KLDLSNVQ**S**K
33	291-317	S293		CG**S**KDNIKHVPGGGSVQIVYKPVDLSK
34	299-317	S305		HVPGGG**S**VQIVYKPVDLSK
35	299-321	S316		HVPGGGSVQIVYKPVDL**S**KVTSK
36	318-340	S320		VT**S**KCGSLGNIHHKPGGGQVEVK
37	322-343	S324		CG**S**LGNIHHKPGGGQVEVKSEK
38	350-370	S352		VQ**S**KIGSLDNITHVPGGGNKK
39	354-370	S356	12E8	IG**S**LDNITHVPGGGNKK
40	370-379	T373		KIE**T**HKLTFR
41	370-383	T377		KIETHKL**T**FRENAK
42	384-406	S396	PHF1	AKTDHGAEIVYK**S**PVVSGDTSPR
43	386-406	S400		TDHGAEIVYKSPVV**S**GDTSPR
44	384-406	T403		AKTDHGAEIVYKSPVVSGD**T**SPR
45	396-406	S404	PHF1	SPVVSGDT**S**PR
46	407-438	S413		HLSNVS**S**TGSIDMVDSPQLATLADEVSASLAK
47	407-438	T414		HLSNVSS**T**GSIDMVDSPQLATLADEVSASLAK
48	407-438	S416		HLSNVSSTG**S**IDMVDSPQLATLADEVSASLAK
49	407-438	S422		HLSNVSSTGSIDMVD**S**PQLATLADEVSASLAK
50	407-438	S433		HLSNVSSTGSIDMVDSPQLATLADEVS**A**SLAK
51	407-438	S435		HLSNVSSTGSIDMVDSPQLATLADEVSA**S**LAK

Tab. 2: P-sites in human Tau (2N4R) expressed in Sf9 cells by three independent studies: Tepper et al. (2014) [MALDI-TOF MS], Mair et al. (2016) [FLEXITAU] & current work [native MS and LC-MS/MS]. Col 1= potential P-sites in full-length Tau (residues S, T, Y), col 2-4 = observed P-sites in Tau from Sf9 cells, col. 5= P-sites in Tau from human AD brain (Hanger 2020). Red = additional sites observed in current work, blue = sites observed in previous but not in present study.

Potential Site S,T, Y in htau 2N4R	P-Sites obs. Tepper 2014	P-Sites obs. Mair 2016	P-Sites obs. This work	P-Sites reported AD brain (Hanger, 2020)
T 17			T 17	
Y 18				Y 18
T 30				T 30
T 39			T 39	T 39
S 46			S 46	S 46
T 50			T 50	
T 52			T 52	
S 61			S 61	
S 68	S 68	S 68		S 68
T 71				
T 69	T 69	T 69		T 69
T 71	T 71		T 71	T 71
T 76	T 76	T 76	T 76	
T 111			T 111	
S 113				S 113
T 123				T 123
S 131	S 131		S 131	
T 135	T 135			
T 149	T 149			T 149
T 153	T 153	T 153		T 153
T 169			T 169	
T 175		T 175	T 175	T 175
T 181		T 181	T 181	T 181
S 184		S 184	S 184	S 184
S 185			S 185	S 185
S 191				S 191
Y 197				Y 197
S 198				S 198
S 199		S 199	S 199	S 199
S 202	S 202	S 202	S 202	S 202
T 205	T 205	T 205	T 205	T 205
S 208			S 208	S 208
S 210		S 210	S210	S 210
T 212	T 212	T 212	T 212	T 212
S 214	S 214	S 214	S 214	S 214
T 217		T 217	T 217	T 217
T 220				T 220
T 231	T 231	T 231	T 231	T 231
S 235	S 235	S 235	S 235	S 235
S 237			S 237	S 237
S 238			S 238	S 238
S 241				S 241
T 245			T 245	
S 258			S 258	S 258
S 262	S 262	S 262	S 262	S 262
S 285			S 285	
S 289			S 289	S 289
S 293		S 293	S 293	
S 305	S 305		S 305	S 305
Y 310	Y 310			
S 316	S 316		S 316	
S 320			S 320	
S 324		S 324	S 324	S 324
S 352			S 352	
S 356	S 356	S 356	S 356	S 356
T 373			T 373	
T 377			T 377	
Y 394				Y 394
S 396	S 396	S 396	S 396	S 396
S 400			S 400	S 400
T 403			T 403	T 403
S 404	S 404	S 404	S 404	S 404
S 409				S 409
S 412				S 412
S 413			S 413	S 413
T 414			T 414	T 414
S 416			S 416	S 416
S 422			S 422	S 422
T 427				T 427
S 433			S 433	S 433
S 435			S 435	S 435

Table 3: Ten most abundant Tau-associated proteins identified by LC-MS/MS analysis in samples of purified P$_o$-Tau expressed in *E.coli*. Protein samples were digested with trypsin followed by LC-MS/MS. Data were analyzed using MaxQuant software and Uniprot organism specific databases for *H. sapiens*, *E.coli* and *Spodoptera frugiperda* (SPOFR). iBAQ = intensity-based absolute quantification correcting for differences in the number of predicted peptides per protein. Only the 10 most abundant proteins per sample are listed. Average of two biological replicates. Note the frequency of RNA- or DNA-binding proteins copurified with Tau.

UniProt Id	Source	Gene name	Protein name	Pept. count	Seq. cov. [%]	Mol. wt [kDa]	iBAQ intensity/E+10 ctrl-Tau	rank ctrl-Tau
P10636-8	HUMAN	MAPT	**Tau**	102	95.9	45.849	3.10	1
P0A7N9	ECOLI	rpmG	**50S ribosomal protein L33**	8	80	6.3715	0.62	2
P0AEU7	ECOLI	skp	**Chaperone protein Skp**	16	76.4	17.688	0.35	3
P0A7M9	ECOLI	rpmE	**50S ribosomal protein L31**	8	90	7.871	0.20	4
P0ADN2	ECOLI	yifE	**UPF0438 protein YifE**	21	87.5	13.133	0.19	5
P0ACF0	ECOLI	hupA	**DNA-binding protein HU-alpha**	10	71.1	9.5349	0.14	6
P60624	ECOLI	rplX	**50S ribosomal protein L24**	17	79.8	11.316	0.10	7
P0ADK8	ECOLI	yibL	**Uncharacterized protein**	10	53.3	13.696	0.04	8
P0ACF4	ECOLI	hupB	**DNA-binding protein HU-β**	8	68.9	9.2255	0.04	9
P68191	ECOLI	sra	**ribosome-associated protein**	4	57.8	5.0958	0.04	10

Table 4A: 20 most abundant proteins co-purifying with Tau identified by LC-MS/MS in fractions of P_m-Tau and P_h-Tau from Sf9 cells. For protein functions see Tab. 4B.

UniProt Id	Source	Gene name	Peptide count	Seq. cov. [%]	Mol. mass [kDa]	iBAQ P_m-Tau	iBAQ P_h-Tau	rank P_m Tau	rank P_h Tau
P10636-8	HUMAN	MAPT	102	95.9	45.849	1.8E+10	2.6E+10	1	1
A0A2H1VYE4	SPOFR	SFRICE_014287	4	20.0	12.545	4.1E+09	1.1E+09	2	5
Q8WQJ0	SPOFR	SFRICE_016047	20	100	11.525	1.8E+09	1.2E+09	4	2
A0A1C7D1B9	SPOFR	CALM2	19	99.3	16.810	1.8E+09	1.1E+09	3	3
P62925	SPOFR	eIF-5A	22	86.9	17.525	1.2E+09	1.1E+09	5	4
A0A2H1W3F3	SPOFR	SFRICE_023497	14	48.8	22.506	1.1E+09	7.7E+08	6	6
A0A2H1W368	SPOFR	SFRICE_003581	27	80.4	19.592	9.0E+08	6.9E+08	7	7
A0A2H1W7P4	SPOFR	SFRICE010198.2	14	83.5	11.217	7.5E+08	4.3E+08	9	13
I7F433	SPOFR	CFL	22	92.6	16.995	5.4E+08	5.3E+08	12	9
Q962R7	SPOFR	RpS12	14	84.2	14.892	3.6E+08	6.5E+08	18	8
A0A2H1WD76	SPOFR	SFRICE_001903	35	64.5	43.714	5.6E+08	4.5E+08	11	12
A0A2H1VL41	SPOFR	SFRICE_006583	9	59.2	13.341	5.3E+08	4.7E+08	13	10
A0A2H1VF43	SPOFR	SFRICE_021632	14	56.8	12.194	7.6E+08	2.4E+08	8	26
A0A2H1W542	SPOFR	SFRICE_011912	8	46.9	6.968	5.0E+08	4.2E+08	14	14
A0A2H1VA11	SPOFR	SFRICE_005530	9	73.9	16.921	5.9E+08	3.2E+08	10	17
A0A2H1WWX	SPOFR	SFRICE_031205	8	78.7	10.157	4.7E+08	4.0E+08	15	16
A0A2H1W5E9	SPOFR	SFRICE_008465	10	82.7	12.446	3.7E+08	4.1E+08	16	15
A0A2H1WQB1	SPOFR	SFRICE_004596	13	90.0	10.196	2.4E+08	4.7E+08	27	11
A0A2H1V4G9	SPOFR	SFRICE_023010	12	38.3	21.015	3.7E+08	2.9E+08	17	19
A0A2H1VTT2	SPOFR	SFRICE_010125	6	62.1	14.986	2.8E+08	3.0E+08	23	18
A0A2H1WKM22	SPOFR	SFRICE_038968	2	64.9	6.6093	3.1E+08	2.5E+08	20	24
A0A2H1VSI0	SPOFR	SFRICE_022351	15	23.6	46.453	3.6E+08	1.9E+08	19	28
Q962Q2	SPOFR	RpS28	8	61.5	7.3184	1.0E+08	2.8E+08	75	20

Table 4B: Known functions of proteins co-purified with Tau.

Protein ID	Source	Gene name	Function
A0A2H1VYE4	SPOFR	SFRICE_014287	uncharacterized protein
Q8WQJ0	SPOFR	SFRICE_016047	60S acidic ribosomal protein P2; structural constituent of ribosome; translational elongation
A0A1C7D1B9	SPOFR	CALM2	calcium ion binding
P62925	SPOFR	eIF-5A	ribosome binding; translation elongation factor activity
A0A2H1W3F3	SPOFR	SFRICE_023497	uncharacterized protein
A0A2H1W368	SPOFR	SFRICE_003581	transcription factor BTF3 (protein inferred by homology)
A0A2H1W7P4	SPOFR	SFRICE 010198.2	ATP binding prot (chaperon)
I7F433	SPOFR	CFL	cofilin; actin binding protein
Q962R7	SPOFR	RpS12	40S ribosomal protein S12
A0A2H1WD76	SPOFR	SFRICE_001903	RNA binding; protein predicted; unreviewed
A0A2H1VL41	SPOFR	SFRICE_006583	DNA binding protein; component of nucleus (predicted)
A0A2H1VF43	SPOFR	SFRICE_021632	uncharacterized protein
A0A2H1W542	SPOFR	SFRICE_011912	translation machinery associated protein 7
A0A2H1VA11	SPOFR	SFRICE_005530	proton-transporting ATP synthase activity; ATP synthesis proton coupled transport
A0A2H1WWX8	SPOFR	SFRICE_031205	chaperon mediated protein transport
A0A2H1W5E9	SPOFR	SFRICE_008465	protein N-linked glycosylation, protein in ER in oligosaccharyl transferase complex; integral component of membrane
A0A2H1WQB1	SPOFR	SFRICE_004596	DNA binding; protein predicted; unreviewed
A0A2H1V4G9	SPOFR	SFRICE_023010	
A0A2H1VTT2	SPOFR	SFRICE_010125	heme binding; metal ion binding
A0A2H1WKM2	SPOFR	SFRICE_038968	translation initiation factor activity
A0A2H1VSI0	SPOFR	SFRICE_022351	poly-ADP-D ribose binding protein; ubiquitin protein ligase; zinc ion binding
Q962Q2	SPOFR	RpS28	40S ribosomal protein S28

Supplement Table S1

Sequence coverage of Tau peptides in Mair 2016 (M) and current work (D) with all P-sites found.

Red letters= P-sites observed in current work. Blue=sites observed in Mair 2016 study. Note that the newly covered sequences in this study (symbol ***) contain 14 observed phosphorylation sites (red letters S, T). Sequences covered previously (Mair2016) but not in current study (symbol +++) contain only 1 observed phosphorylation site (blue, T153).

Thus, the excess of new sequence coverage is sufficient to explain the higher occupancy in the current study.

grey: missing residues in sequence coverage.

*** Gaps in (M)

+++ Gaps in (D);

--- Gaps in both.

```
M   1 MAEPRQEFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLK
D     MAEPRQEFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLK
      -----
M  45 ESPLQTPTEDGSEEPGSETSDAKSTPTAE
D     ESPLQTPTEDGSEEPGSETSDAKSTPTAE

M  74 DVTAPLVDEGAPGKQAAAQPHTEIPEGTT
D     DVTAPLVDEGAPGKQAAAQPHTEIPEGTT

M 103 AEEAGIGDTPSLEDEAAGHVTQARMVSKSKDGTGSDDKKAKGADGKTK
D     AEEAGIGDTPSLEDEAAGHVTQARMVSKSKDGTGSDDKKAKGADGKTK
                             ****************-------
M 151 IATPRGAAPPGQKGQANATRIPAKTPPAPKTPPSSGEPPKSGDRSGY
D     IATPRGAAPPGQKGQANATRIPAKTPPAPKTPPSSGEPPKSGDRSGY
      +++++++++++++                            ****
M 198 SSPGSPGTPGSRSRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRL
D     SSPGSPGTPGSRSRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRL
            *****
M 244 QTAPVPMPDLKNVKSKIGSTENLKHQPGGGK
D     QTAPVPMPDLKNVKSKIGSTENLKHQPGGGK
                 *****    -------
M 275 VQIINKKLDLSNVQSKCGSKDNIKHVPGGGS
D     VQIINKKLDLSNVQSKCGSKDNIKHVPGGGS
      ++++++          ********
M 306 VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ
D     VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ
                 ****
M 337 VEVKSEKLDFKDRVQSKIGSLDNITHVPGGGN
D     VEVKSEKLDFKDRVQSKIGSLDNITHVPGGGN
              +++---****
M 369 KKIETHKLTFRENAKAKTDHGAEIVYKSPVVS
D     KKIETHKLTFRENAKAKTDHGAEIVYKSPVVS
        *****
M 401 GDTSPRHLSNVSSTGSIDMVDSPQLATLADEVSASLAKQGL
D     GDTSPRHLSNVSSTGSIDMVDSPQLATLADEVSASLAKQGL
            *****************************---
```

Figure legends

Fig. 1: Overview: Phosphorylation of hTau-2N4R expressed in Sf9 cells and *E. coli*.
(A) Diagram of domains of hTau40 (2N4R or Tau-F, Uniprot Id P10636-8), the largest human isoform of CNS Tau consisting of 441 residues with the three alternatively spliced inserts N1, N2 and R2. The N-terminal half represents the projection domain, and the C-terminal half contains the four pseudo-repeats (R1-R4) which - in combination with their flanking domains P2 and R' - represent the microtubule assembly domain.
(B) Schematic representation of production of unphosphorylated Tau (P_0) in *E. coli* (prokaryote) and hyperphosphorylated Tau (P_m and P_h) Tau in Sf9 cells (eukaryote). Okadaic acid (OA) treatment increases the phosphorylation of Tau which yields P_h -Tau.
(C) Amino acid sequence of Tau (2N4R, 441 residues) showing phosphorylation sites identified in this study (*red and purple letters*; *purple=Pro directed)*) and further potential phosphorylation sites (not detected so far, *blue letters*) (total of 85 sites, 45 Ser, 35 Thr, 5 Tyr). Note that only a minority of phosphorylation sites of 37% was detected in the N-terminal part of the sequence (up to residue ~150, compared to 71% in the C-terminal part of the protein. Note also that none of the five Tyr residues was phosphorylated.
(D) SDS-PAGE analysis stained with Coomassie blue showing P_0 Tau purified from *E. coli*, P_m and P_h Tau purified from Sf9 cells. Note the upward shift in Mr value with increasing phosphorylation, from 55 kDa (P_0-Tau) to 68 kDa for P_h-Tau (compared with the theoretical MW values of ~45850 Da and ~46863 Da). This shift is characteristic for AD-Tau.

Fig. 2: Analysis of hTau-2N4R expressed in Sf9 cells by native MS. 20 µM of purified proteins in 200 mM ammonium acetate, pH 7.6, were analyzed by nanoflow electrospray ionization quadrupole time-of-flight mass spectrometry. Representative mass spectra for the hyperphosphorylated states P_h-Tau (top) and P_m-Tau (middle) as well as unphosporylated P_0-Tau (bottom) are shown.
Top, main signals were assigned to charge state series of P_h-Tau (MW of 46883 Da; filled red circles) and a copurified component with a molecular weight of 59642 Da (grey squares).
Middle, spectrum of P_m-Tau displaying peaks assigned to a series of equivalent charge states, but shifted towards lower *m/z* compared to those for P_h-Tau (red squares). Further signals match closely with those in the spectrum of P_h-Tau.
Bottom, spectrum of control Tau P_0 consisting of a series of charge states indicating a molecular weight of 45,724 Da (open red circles). A further series was assigned to a MW of 47,085 Da. It was attributed to a trimeric species (grey circles), as it decomposes upon increasing collisional activation into highly charged (centered at +9, below *m/z* 2,000) monomeric and charge-stripped dimeric species (charges +6 to +8). Its monomeric mass of 15,691 Da matches the theoretical mass of *the E. coli* chaperone protein *skp* (see **Table 3**), taking into account removal of its N-terminal 20 amino acids signal peptide.

Fig. 3: Phosphorylation of hTau-2N4R expressed in Sf9 cells analyzed by native MS.
Left, signals in the range between 2800 and 3500 *m/z* attributed to Tau proteins.
Right, corresponding charge state-deconvoluted spectra. The mass spectrum of P_h-Tau (top) displays a center mass of 46,883 Da (filed red circles) and additionally six peaks on both sides equally spaced by ~80 Da. P_m-Tau (middle) displays peaks of similar width shifted towards lower *m/z* compared to those for P_h-Tau, which can be assigned to the equivalent charge states and indicate a molecular weight of approximately 46,408 Da (red squares). The spectrum of control Tau P_0 consists of a series of charge states resulting in a comparatively sharp peak at a mass of 45,724 Da (45,850 -131Da +/- 5 of missing Met-1; open red circles).
A further series was assigned to a MW of 47,085 Da and was attributed to a trimeric species (grey circles), as it decomposes upon increasing collisional activation into highly charged (centered at +9, below *m/z* 2,000) monomeric and charge-stripped dimeric species (charges +6 to +8). Its monomeric mass of 15,691 Da matches the theoretical mass of *the E. coli* chaperone protein *skp* (see **Table 3**), taking into account removal of its N-terminal 20 amino acids signal peptide.

Fig. 4: Bar diagram of full-length human Tau protein and distribution of phosphorylation sites. Left, no phosphorylation observed; right, phosphorylation sites observed in Sf9 cells.

Fig. 5: Bar diagram of full-length human Tau protein with complementary presentation of total phosphorylation sites acquired by Tau in Sf9 cells (left) and phosphorylated sites identified in AD Tau (right).

Fig. 1 A-D

Phosphorylation of hTau-2N4R expressed in Sf9 cells and E.coli
Overview of Tau domains, sequence, preparations

A. **2N4R**

repeat domain

| NT | N1 | N2 | | P1 | P2 | R1 | R2 | R3 | R4 | R` | CT |

1   45 74   151   198   244   368   401   441

projection domain   MT assembly domain

D. M P_0 P_h P_m

kD
100
70
55

B.   Expression   Treatment   Purification

Sf9 cells

+OA ———————————→ P_h

−OA ———————————→ P_m

E.coli cells   ———————————→ P_0

C.   **AA Sequence of 2N4R hTau**   **Subdomains**

```
  1 MAEPRQEFEVMEDHAGTYGLGDRKDQGGYTMHQDQEGDTDAGLK        NT
 45 ESPLQTPTEDGSEEPGSETSDAKSTPTAE                       N1
 74 DVTAPLVDEGAPGKQAAAQPHTEIPEGTT                        N2
103 AEEAGIGDTPSLEDEAAGHVTQARMVSKSKDGTGSDDKKAKGADGKTK
151 IATPRGAAPPGQKGQANATRIPAKTPPAPKTPPSSGEPPKSGDRSGY      P1
198 SSPGSPGTPGSRSRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRL       P2
244 QTAPVPMPDLKNVKSKIGSTENLKHQPGGGK                      R1
275 VQIINKKLDLSNVQSKCGSKDNIKHVPGGGS                      R2
306 VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ                      R3
337 VEVKSEKLDFKDRVQSKIGSLDNITHVPGGGN                     R4
369 KKIETHKLTFRENAKAKTDHGAEIVYKSPVVS                     R'
401 GDTSPRHLSNVSSTGSIDMVDSPQLATLADEVSASLAKQGL            CT
```
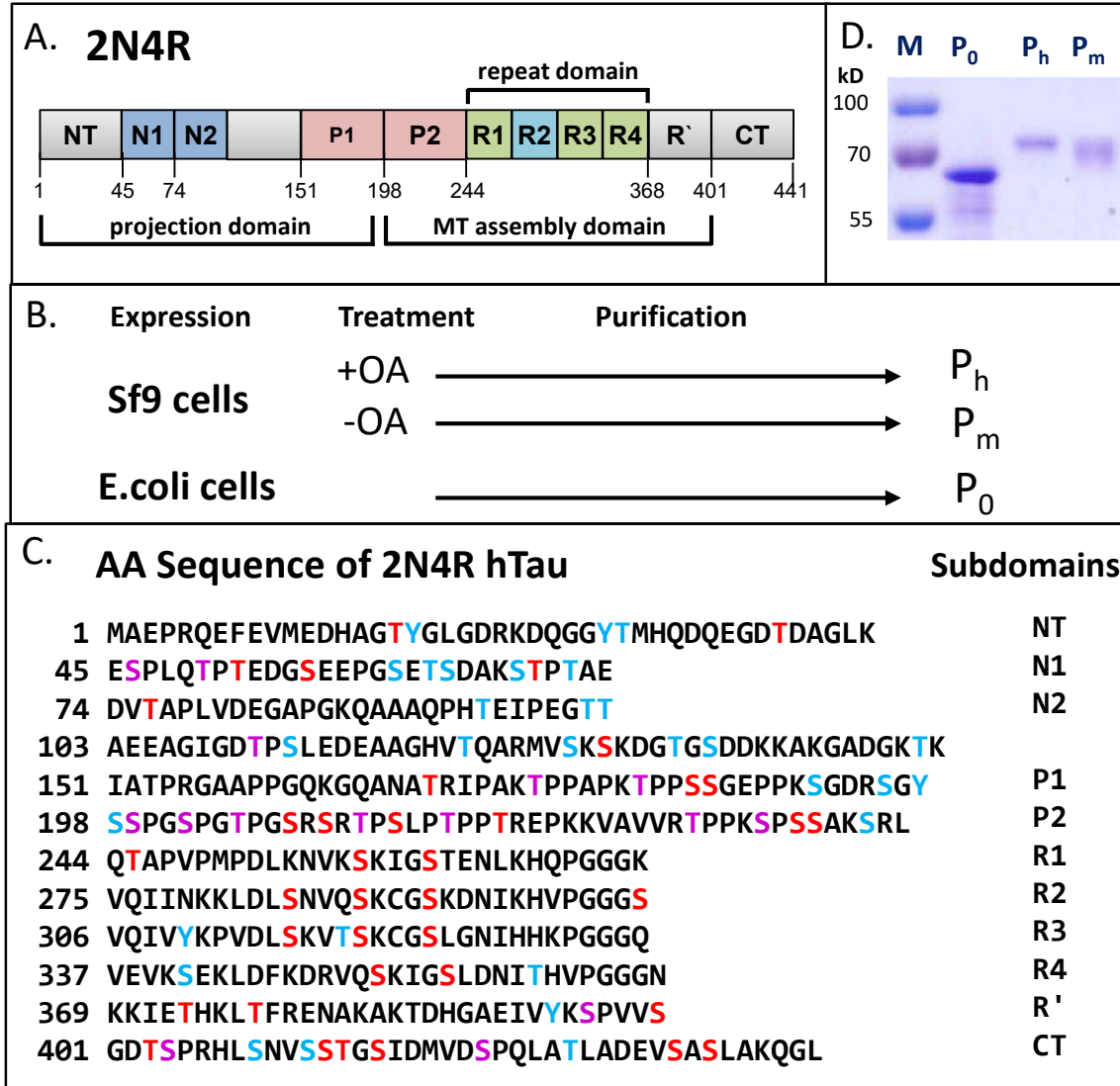
Fig. 2

Analysis of hTau-2N4R expressed in Sf9 cells by native MS: charge state assignment
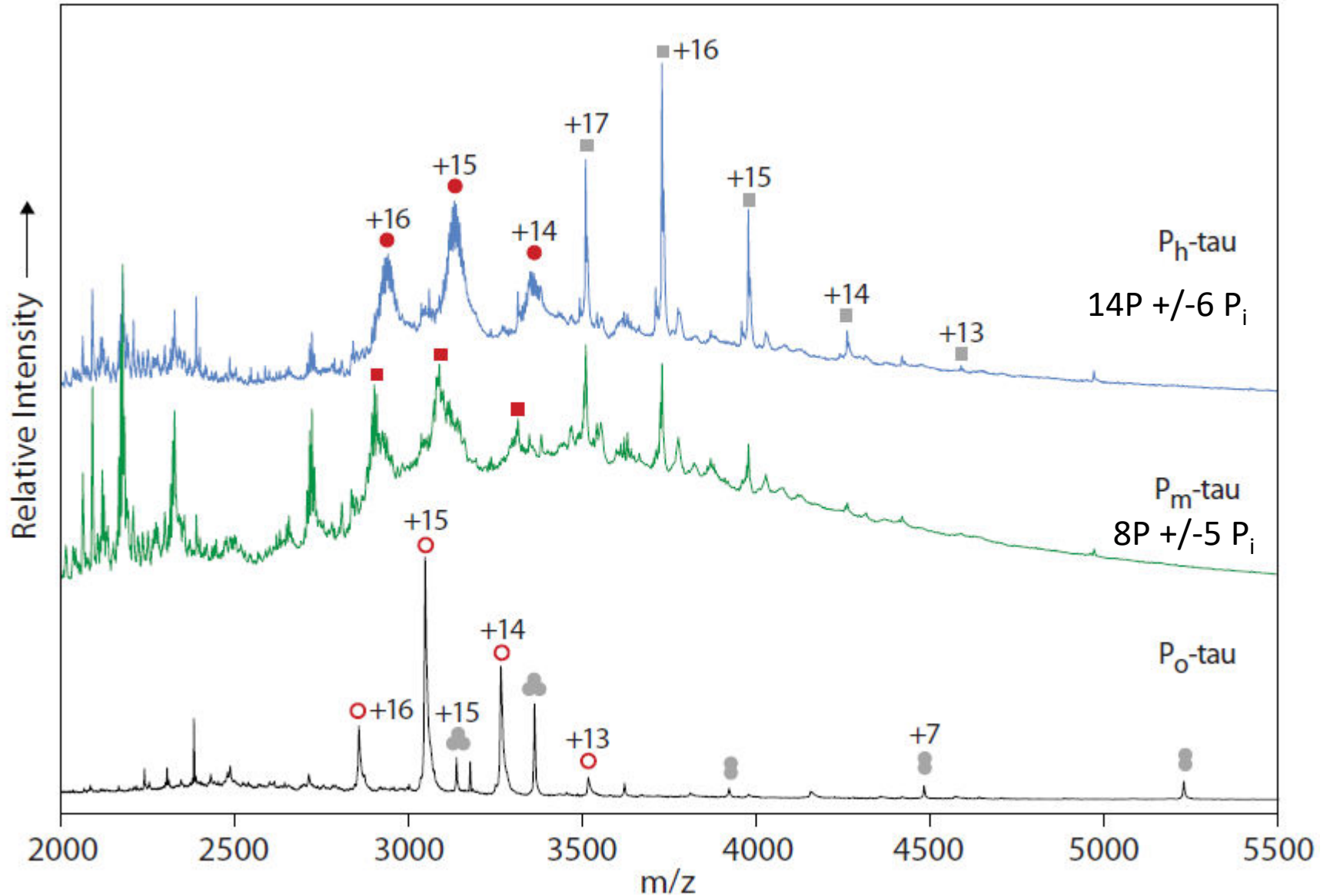
Fig. 3 Analysis of hTau-2N4R expressed in Sf9 cells by native MS: charge state-deconvoluted spectra
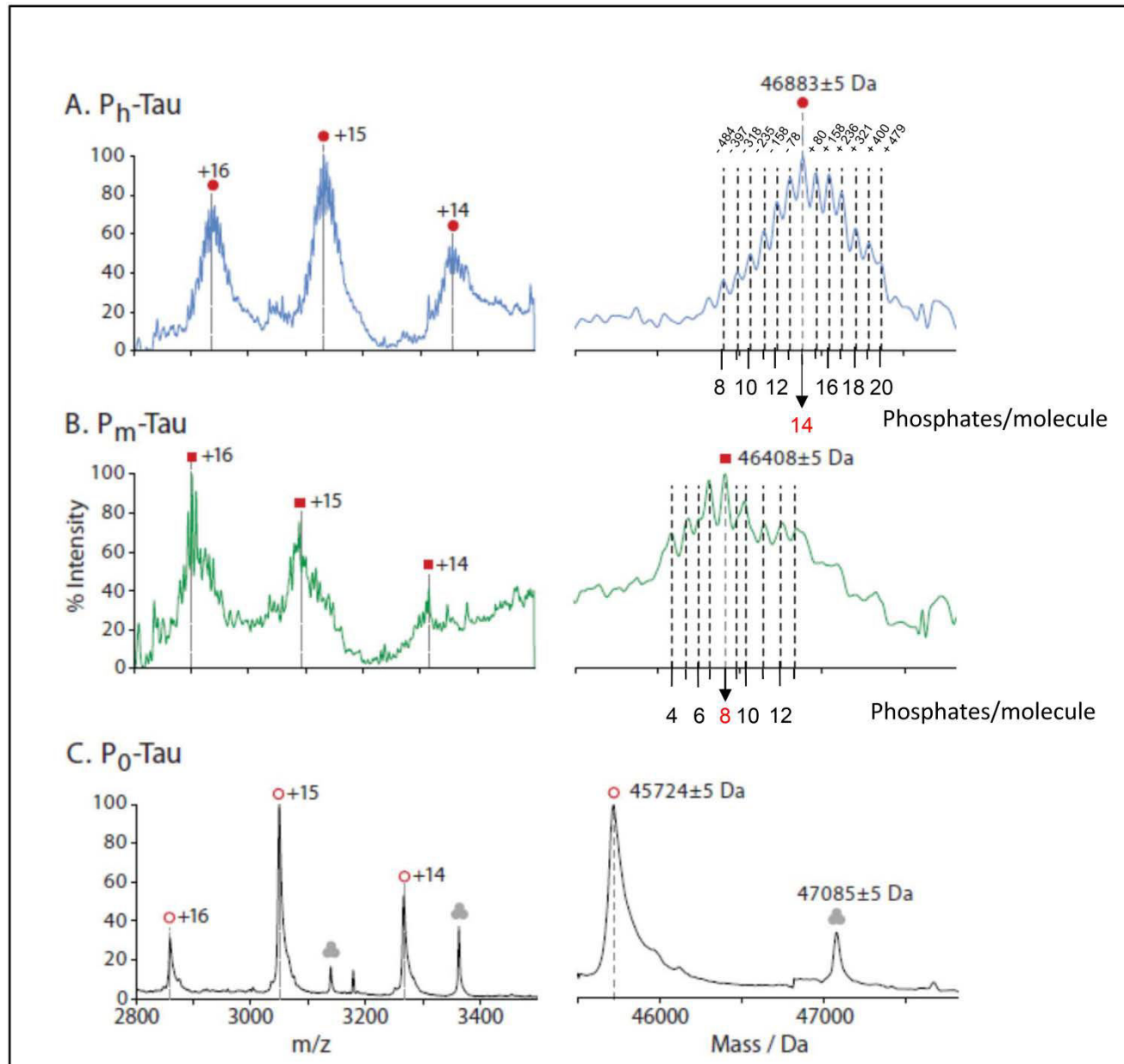
Fig. 4

Potential vs observed phosphorylation sites of Tau-2N4R from Sf9 cells

no phosphorylation observed, mainly in N-terminal half (blue)

phosphorylation observed, mainly in C-terminal half (red)
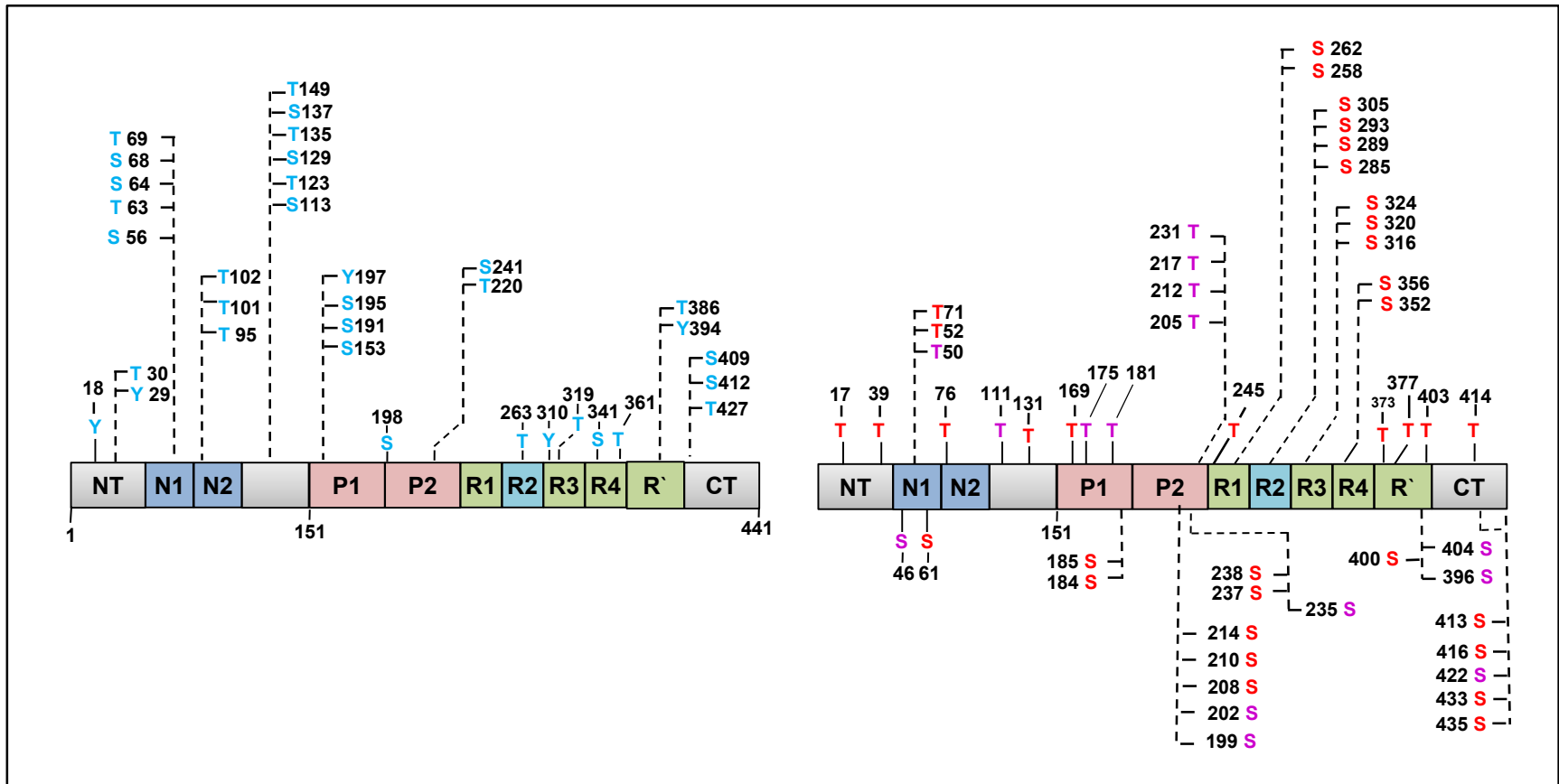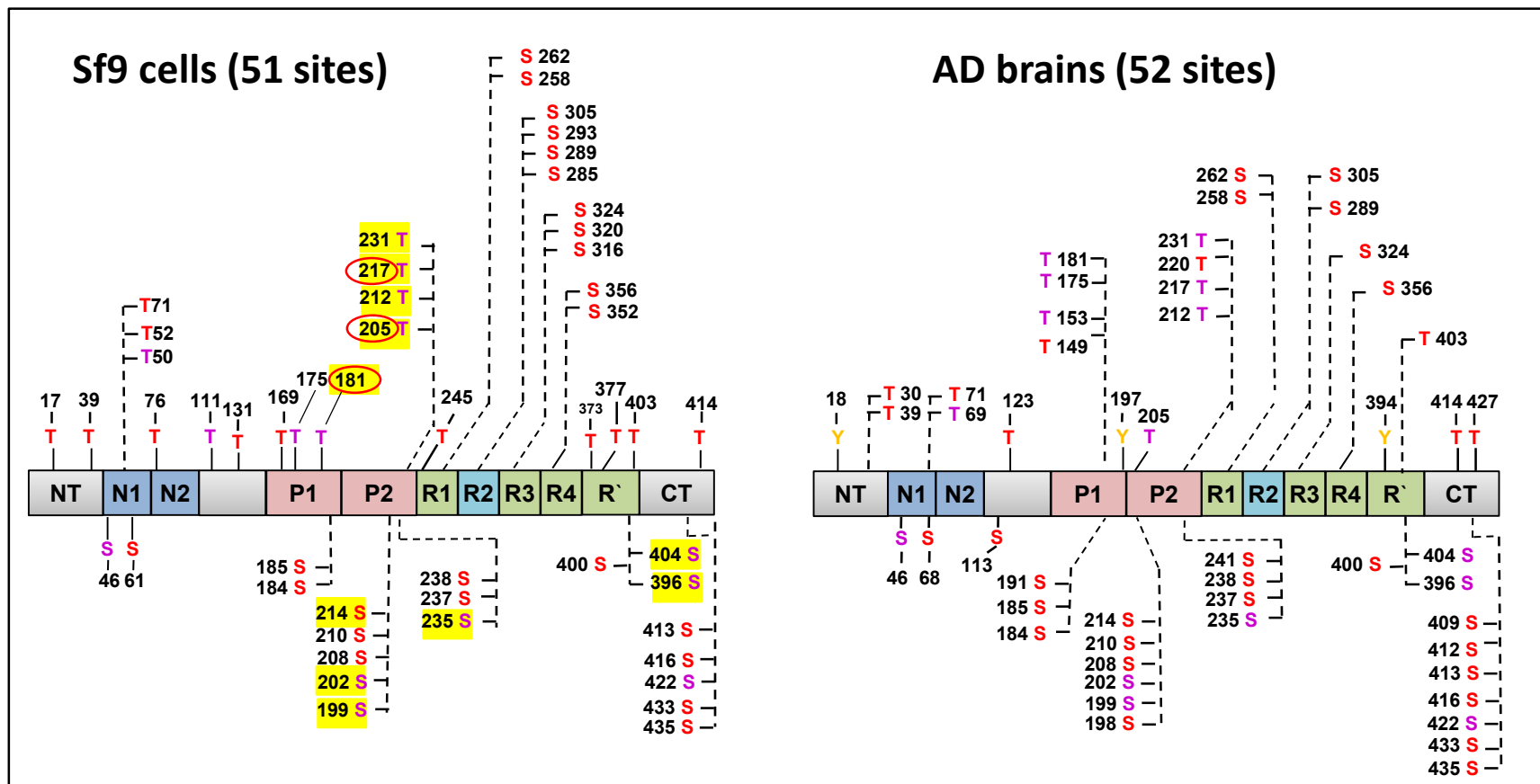sites in SP or TP motifs (purple)

Fig. 5

hTau phosphorylation sites observed in Sf9 cells vs. AD brain

>50 sites, mainly in C-terminal half, overlap 68%

red: S, T sites, purple SP or TP motifs
yellow highlight: high occupancy (Mair 2016)
circles: important for AD diagnostics (Barthelemy 2020)