

Unraveling the structure of monomeric
and fibrilized 441-residue tau with
NMR spectroscopy

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"doctor rerum naturalium"
der Georg-August-Universität Göttingen

vorgelegt von
Stefan Bibow
aus Berlin

Göttingen 2011

Mitglied des Betreuungsausschusses 1:

Mitglied des Betreuungsausschusses 2:

Mitglied des Betreuungsausschusses 3:

Tag der mündlichen Prüfung:

The truth is an elephant described by three blind men.
The first man touches the tail and says it's a rope.
The second man feels the rough leg and says it's a tree.
The third man feels the trunk and says it's a snake.

(Variation of the the poem "The Blind Men and the Elephant"
by John Godfrey Saxe.)

Affidavit

I hereby declare that this thesis has been written independently and with no other sources and aids than quoted.

Stefan Bibow

Acknowledgements

I would like to thank Prof. Dr. Markus Zweckstetter for the opportunity to investigate this interesting protein. Furthermore, I thank Prof. Christian Griesinger for providing an excellent research environment and for many useful tips.

Dr. Marco Mukrasch, my predecessor on the tau project, was a huge help in introducing me to the project, to scientific procedures and to colleagues. Worth mentioning, he actually showed me how to set-up heteronuclear multidimensional experiments.

I thank Dr. Lukasz Skora for sharing his impressive knowledge about NMR-hardware with me to maintain the spectrometer in the best possible way. He instructed me with passion and patience so that I could continue to successfully compete for the highest Signal-to-Noise ratio in Göttingen. He additionally introduced me to HRMAS and was a respectable contender in Steeldarts.

I would like to express my gratitude to Dr. Nils Lakomek, who helped me with the preparation of RDC gels. I thank Dr. Nasrollah Rezaei-Ghaleh for helping me with the relaxation experiments and many stimulating discussions. I had many fruitful discussions with Shengqi Xiang, Dr. Edward D'Auvergne, Dr. Donghan Lee, Dr. Ullrich Dürr and Dr. Min-Kyu Cho.

I am much obliged to Dr. Antoine Loquet, Dr. Adam Masur, Dr. Mitchell Martinez, David Ban and Eibe Dücker. Discussions with the 'family' during our evening sessions resulted in many interesting ideas.

I am deeply grateful for all the tau-samples, that were produced in the lab of Prof. E. Mandelkow. The high quality allowed the recording of excellent spectra.

I thank Dr. Martin Blackledge for his contribution to the E-mutant paper and for the nice stay in Grenoble.

Finally, I would like to thank all my colleagues in the department NMR-

based structural biology for a nice and friendly atmosphere making it an enjoyable place to work and to discuss.

Abstract

The microtubule-associated, intrinsically disordered protein tau is a key-protein in maintaining the axonal transport in neurons. The association to and the dissociation from microtubules is regulated via a complex pattern of phosphorylation that is still largely unknown. Moreover, a clear relationship between Alzheimer's disease (AD) and tau was established. Tau-mediated neurodegenerative diseases are characterized by insoluble tau-aggregates, termed *paired helical filaments* that eventually assemble into *neurofibrillary tangles*. However, it is unclear, which extracellular or intracellular factors might trigger the transformation from a physiological important protein into a pathological agent.

In this thesis, several aspects of tau were investigated using NMR spectroscopy. In the first part of the results section, paramagnetic relaxation experiments were conducted that allowed the characterization of the global folding of monomeric tau in solution with great detail. The 2288 distance restraints derived from these experiments has allowed for an ensemble structure calculation, that resulted in a residue-specific model.

The second part of the results section investigates the performance of disorder predictors. We found a good correlation of the residue-contacts predicted with IUPred when compared to the contact map of tau. Protein-binding sites prediction with ANCHOR is in very good agreement with the hot-spots of microtubule-binding. For the secondary structure prediction, a combined approach of predictors allowed for the correct secondary structure assignment to the strongest residual secondary structure propensities.

The occurrence of paired helical filaments is accompanied with a hyperphosphorylated state of tau. To avoid incomplete and unspecific phosphorylation, we decided to mimic phosphorylation by replacing serines and threonines with glutamic acid in epitopes of AD-diagnostic antibodies (AT8, AT100 and PHF1). The global and local structural characteristics were

investigated and compared to the wildtype of tau. We demonstrate, that pseudophosphorylation results in an opening of the paperclip conformation. The higher aggregation rate found for this pseudophosphorylated tau-mutant underpins the importance of the transient network of intramolecular interactions. We found, that a weakening of the transient interactions precedes local structural changes.

In the fourth part, the NMR-investigations were conducted on the fuzzy coat of tau paired helical filaments. Due to the size of the aggregates of more than 1 MDa, the use of *high resolution magic angle spinning* was imperative. The presence and assignment of residues $^1\text{M-T}^{212}$ and $^{399}\text{V-L}^{441}$ indicate a highly dynamic N-terminal and C-terminal fuzzy coat, that transiently contact the PHF-core. In particular the very N-terminus shows a tightened electrostatic interaction with the PHF-core. These findings rationalize the affinity of conformation-specific PHF-tau antibodies that exhibit two discontinuous epitopes, one at the very N-terminus and the second within the PHF-core.

Contents

Affidavit	v
Acknowledgements	vii
Abstract	ix
1 Introduction	1
1.1 From Rigid To Dynamic Proteins	2
1.2 Intrinsically Disordered Proteins (IDPs)	5
1.2.1 Characteristics of IDPs	5
1.2.2 Functions of IDPs	5
1.2.3 IDPs in Disease	8
1.3 Tau As An Intrinsically Disordered Protein	9
1.3.1 Location and amino acid composition	9
1.3.2 Translocation of tau	11
1.3.3 The physiological role of tau	11
1.3.4 Phosphorylation of tau	14
1.3.5 The pathological role of tau	18
1.3.6 The interplay of tau and $A\beta$ in Alzheimer's disease . .	21
2 Nuclear Magnetic Resonance	23
2.1 Introduction	23
2.2 Nuclear Magnetism	24
2.3 ${}^3J_{HNH\alpha}$ Scalar Couplings	27
2.4 Secondary Chemical Shifts	30
2.5 Paramagnetic Relaxation Enhancement	31
2.6 Residual Dipolar Couplings	33
2.7 High-Resolution Magic Angle Spinning	34

3	Materials and Methods	39
3.1	Biochemical Methods	39
3.1.1	Production of htau40	39
3.1.2	Site-specific labeling of htau40 with MTSL	40
3.1.3	Aggregation of tau	40
3.2	NMR Spectroscopic Methods	41
3.2.1	2D experiments	41
3.2.2	Chemical shift difference	41
3.2.3	Secondary chemical shifts	41
3.2.4	Structure calculation	42
3.2.5	$^3J_{HNH\alpha}$ couplings	42
3.2.6	PRE experiments	43
3.2.7	Diffusion experiments	43
3.2.8	Relaxation experiments	45
3.2.9	3D experiments	46
4	Results and Discussion	51
4.1	Global Structure of 441-Residue Tau at Single Residue Resolution	51
4.1.1	Assignment procedure	51
4.1.2	PRE experiments reveal a network of long-range interactions	52
4.1.3	Discussion and Conclusion	55
4.2	Predicting the structural characteristics of tau	58
4.2.1	Introduction	58
4.2.2	Predicting intra-residue contacts with IUPred	59
4.2.3	Predicting residual secondary structure	65
4.2.4	Predicting tau-binding sites with ANCHOR	69
4.3	Structural impact of pseudophosphorylation on the global and local structure of 441-residue tau	73
4.3.1	Introduction	73
4.3.2	Global structural changes	73
4.3.3	Local structural changes	78
4.3.4	Conclusion	79
4.4	Molecular basis of the recognition of tau aggregates by conformation-specific antibodies	82
4.4.1	Introduction	82

4.4.2	Assignment of PHF-tau	82
4.4.3	Transient long-range interactions of the fuzzy coat . . .	85
4.4.4	Nature of the long-range interactions	87
4.4.5	Discussion	89
Appendices		111
A Preliminary results of phosphorylated tau		115
B Pulseprograms		117
C Curriculum Vitae		127

Chapter 1

Introduction

Proteins play a crucial role in all three branches of life. The search for structure and function of these biomolecules fascinated scientists not only since Gerardus Johannes Mulder named these "animal substances" proteins [1]. Though, nearly 120 years passed until a first detailed atomic structure of Myoglobin was presented [2]. Nowadays, a little more than 60 years later, structure-elucidation with atomic resolution advanced in such a tremendous manner that currently more than 66.000 atomic structures of proteins and other biomolecules are known. This enormous amount of data, together with the analysis of many genomes, provides a detailed insight into protein structures and their functions:

- The typical size of a protein is between 50-2000 residues [3] (but exceptional large proteins, like titin with more than 35.000 residues [4], are also present).
- Despite the astonishing combinatorial possibilities of the 20 common amino acids, the existence of only a couple of thousand folds is estimated [5].
- more than 30% of eukaryotic proteins have unstructured elements longer than 50 consecutive residues [6].
- 10% of all proteins are completely unstructured [7].

The occurrence of fully functional but intrinsically disordered proteins forced scientists to re-evaluate the structure-function paradigm of proteins.

1.1 From Rigid To Dynamic Proteins

That the amino acid sequence contains the key information for folding into an ordered state was demonstrated by Anfinsen [8]. Already at this time it was clear that the folding process towards the global energy minimum can not be reached by random search, since a random sampling of all possibilities within the conformational space would take an enormous amount of time [9] (Figure 1.1 A). To solve the problem, called "Levinthal's paradox", it was implicated that the amino acid sequence must contain sequential pathway informations on how the protein should fold, starting from one specific conformation [10] (Figure 1.1 B). However, this view ignores that denatured proteins, as well as intrinsically disordered proteins, exists as an ensemble, represented by many different conformations and hence many different starting points. The solution to this problem is to view the energy landscape as a funnel [11]. In this "new view" the energy landscape has the shape of a funnel, specific for the amino acid sequence, and different conformations can be represented by starting the folding process from different positions on the funnel edge, which all eventually lead to the energy minimum (Figure 1.1 C).

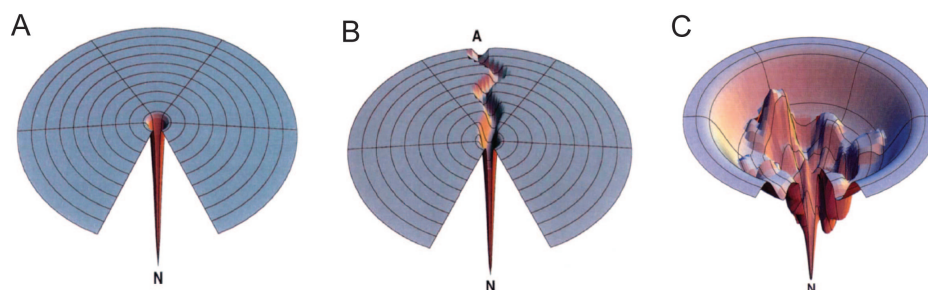


Figure 1.1: Energy landscape representation of different folding-models. A) The "golf-course" landscape representing the problem pointed out by Levinthal. The molecule would need an enormous amount of time to reach the global minimum by sampling the conformational space randomly. B) The "pathway" solution to the random search problem. A well-defined sequence of folding events coded in the amino acid sequence drives the protein from the unfolded random coil to a uniquely folded state. C) The new view proposes "funnels, not tunnels". The extension of the two-dimensional pathway to a three-dimensional funnel represents the current view (from ref. [11]).

The better understanding of protein folding was accompanied by a better understanding of protein-protein interaction. The classical structure-

function paradigm has been the dominant view of proteins for a long time. Therein, function relies on a well-defined 3-dimensional structure. The reason of success lies in the evidence gathered, for example the stereospecificity of α -amylase and emulsin, specifically only hydrolyzing α -glycosidic bonds and β -glycosidic bonds, respectively. Another example are Anfinsen's denaturing-renaturing experiments of RNase, whereas the function of RNase was restored after renaturation, explicitly demanding the need for a well-defined 3D structure [8]. The interaction of biomolecules were initially explained by the lock-and-key model [12], which assumes a strict complementarity of the interaction sites (Figure 1.2 A). However, the static lock-and-key model failed to describe the capability of certain enzymes to catalyze several different substrates. One counterexample was serum albumin, able to bind with high-affinity several small hydrophobic molecules [13]. Accordingly, Koshland suggested a modification of Fischer's model where the substrate binds to an active site, which structurally adopts to the substrate to perform its catalytic function [14] (Figure 1.2 B). This new model was termed "induced fit" and was able to explain the adaptability of an enzyme/protein to several different ligands. Nevertheless, even with the expanded model the protein is treated as if it exists in a single, stable conformation and not as a dynamic entity, which samples many conformations as an ensemble. The alternative model of "conformational selection" or "conformational sampling" takes into account this conformational heterogeneity and argues that weakly populated, higher energy conformations are responsible for recognizing and binding to partners with subsequent population shifts towards these conformers [15, 16]. Therein, a protein's free energy landscape consists of different conformations or substates in dynamic equilibrium (Figure 1.2 C). The populations of the substates follow statistical thermodynamic distributions, and the heights of the energy barriers separating the substates define the timescale of conformational exchange. If the free energy barriers are low relative to the Boltzmann energy ($k_B T$), thermal fluctuations can lead to significant populations of more than one conformational state in solution [16]. Support for this model comes from X-ray crystallography, where a monoclonal antibody was crystallized in two different conformations in the absence of an antigen [17] and from nuclear magnetic resonance for ubiquitin [18].

The key difference between induced fit and conformational selection is that in the induced-fit model the bound-protein-conformation forms only after interaction with a binding partner. In the conformational selection model,

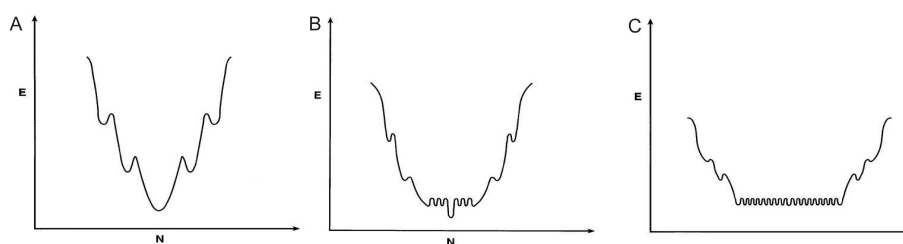


Figure 1.2: Defining the bottom of the funnel. A) In the lock-and-key model only one energy minimum at the bottom of the funnel is present, accounting for the one and only rigid and stable structure. B) The bottom of the funnel for the "induced fit" model possesses again only one global minimum (the bound state), but is surrounded by several local minima (representing the flexibility in the unbound state). C) The "conformational sampling" model with several equal minima separated by low energy barriers, representing the conformational heterogeneity with several substates in dynamic equilibrium (from [19]).

all conformations pre-exists in the ensemble and are dynamically sampled by the free protein in the absence of any ligand. Especially NMR is a powerful tool for studying conformational heterogeneity, due to its sensitivity to dynamics over a wide range of time scales and availability of residue-specific informations [20,21]. With newly developed NMR methods, sub-ensembles populated with as little as 0.5% can be characterized [22]. In addition, residual dipolar couplings (RDC) can be used to calculate the conformational ensemble of ubiquitin up to the microsecond timescale [18]. It was shown, that the conformational ensemble of free ubiquitin samples all 46 X-Ray structures of ubiquitin in complex with various binding partners [18].

Intrinsically disordered proteins can be considered as on the extreme end of these dynamic mechanisms, with a large degree of motion leading to rapid interconversion between highly heterogeneous conformers [23]. One can probably say that the development from rigid to flexible and dynamic proteins culminated in the discovery and acceptance of biological functional intrinsically disordered proteins.

1.2 Intrinsically Disordered Proteins (IDPs)

1.2.1 Characteristics of IDPs

The most remarkable feature of intrinsically disordered proteins (IDPs) is the absence of rigid secondary and tertiary structure. Just as the amino acid sequence codes for the presence of secondary structure in folded proteins, it was speculated that the amino acid sequence in IDPs codes for the absence of rigid secondary structure elements [24]. Sequence studies on folded and intrinsically disordered proteins showed a significantly different amino acid composition for disordered proteins (Figure 1.3). IDPs are depleted in 'order-promoting' amino acids (tryptophan, isoleucine, phenylalanine, leucine, cysteine, tyrosine, threonine, asparagine, histidine) and enriched in 'disorder-promoting' amino acids (valine, lysine, glutamic acid, proline, serine, glutamine and occasionally glycine) [25–27].

This composition results in a high net charge and a low overall hydrophobicity, precluding the formation of hydrophobic clusters and maintaining an extended conformation by electrostatic repulsion [6]. Proline for example is known to disfavor secondary structure but strongly favors an open and extended conformation. Therefore, a high content of prolines is linked with a lack of regular β -sheet or α -helical structure. Moreover, because proline and glutamine are poor in β -strand formation [28], Tompa et al argued that evolutionary selection against β -strand forming residues must be of high priority, because the ease of transformation from an openly structured IDP to a β -strand possesses a high risk of amyloidosis [26]. Another parameter accounting for the absence of rigid structure in IDPs is the low number of average contacts per residue. The protein folding process is directed and stabilized by the formation of intrachain residue-residue contacts to compensate for the conformational entropy [29]. Expectedly, residues with a low contact number are enriched in IDPs and residues with a high contact number are reduced, when compared to folded proteins [30].

1.2.2 Functions of IDPs

Biological processes and biological functions of IDPs were investigated and it was found that disorder is significantly enriched in five large functional areas [7]:

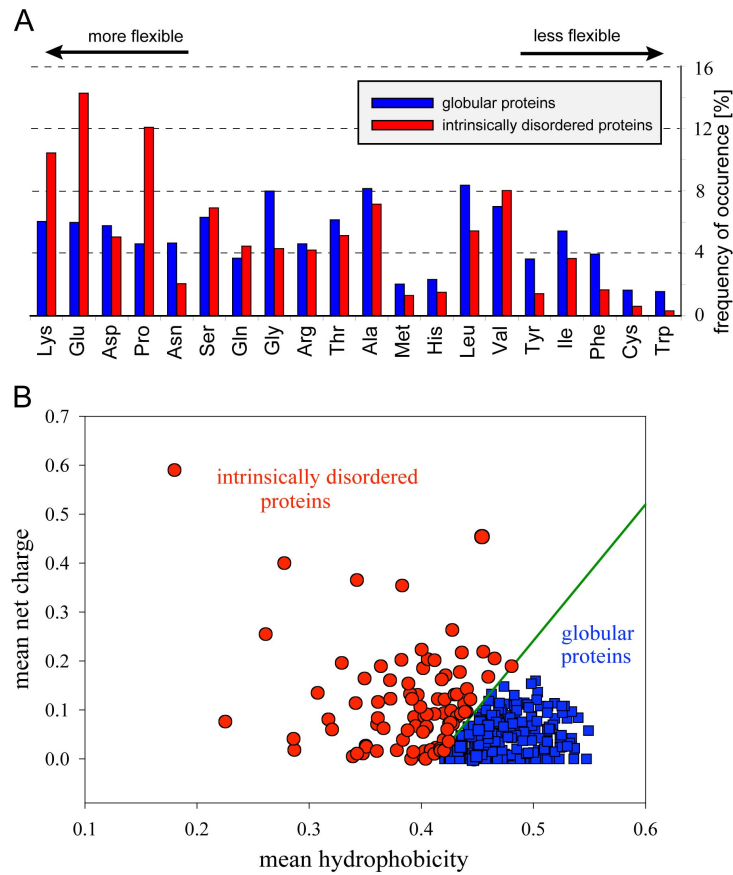


Figure 1.3: Amino acid composition. A) Amino acid frequencies of ordered and disordered proteins (adapted from [26]) B) Mean net charge versus mean hydrophobicity plot for a set of 275 folded (blue squares) and 91 natively unfolded proteins (red circles) (adapted from [27]).

1. Transcription and transcription regulation, also apparent in molecular functions such as DNA binding and nucleotide binding.
2. Signal transduction and the regulation of the cell cycle.
3. The biogenesis and functioning of ribosomes and chromatin.
4. mRNA processing, including splicing reactions.
5. The biogenesis and organization of the cytoskeleton.

What is the advantage of being intrinsically disordered? IDPs are capable of promoting simultaneous interactions with several binding partners

with fast molecular association and dissociation rates [24]. For processes where the association step is rate limiting, initial, relatively weak and non-specific interaction, as it is the case for IDPs, increases the association rate significantly [31]. The high flexibility of IDPs also results in a highly versatile interaction with various targets termed one-to-many signaling [24].

Another advantage can be seen in maintaining the cytoskeletal architecture. Microtubules, microfilaments and intermediate filaments are the three basic cytoskeletal components. Microfilaments are the most versatile elements of the cytoskeleton, involved in cellular processes like muscle contraction, cell adhesion, cell migration, vesicle and organelle transport, signaling and cell division [3]. They are assembled from the globular protein G-actin, also termed F-actin in its filamentous form, to form two intertwined fibers with diameter of 7 nm. The polymerization is controlled via the IDP T β 4 with 45 amino acids.

Intermediate filaments constitute a large and heterogeneous family of proteins. They are classified in 5 classes, 4 located in the cytoplasm and one located in the nucleus. The intermediate filaments in neurons, also called neurofilaments, are composed of three subunits, NF-L (70kDa), NF-M (150kDa), NF-H (200kDa). The variation in the molecular weight comes from the different length of the C-terminal tail. The C-terminus is disordered and highly repetitive and controls the spacing between neighboring filaments probably via an entropic exclusion and not by electrostatic interactions [32]. A similar assignment is proposed for the projection domains of the microtubule associated proteins MAP2 and tau, whereas in this case the N-terminal tail presumably regulates the spacing between microtubules [33]. The high intrinsic flexibility of IDPs can be translated into conformational entropy. In fact, entropic chain functions, termed entropic clocks, entropic springs and entropic brushes, are used to explain the complex functions of IDPs. As was already mentioned above, the spacing of neurofilaments and microtubules is explained via an entropic exclusion principle, also called entropic brush. Another example for an entropic brush are the nucleoporins in the nuclear pore complex (NPC). The exchange of molecular cargo between the nucleus and the cytoplasm functions across numerous perforations in the nuclear envelope, the NPCs. However, the NPC forms a barrier to larger molecular cargo (more than 40 kDa) that do not harbor a nuclear localization signal [34]. The nucleoporins contain disordered phenylalanine-glycine (FG)-repeat domains, projecting into the pore and are responsible for the restricted traffic [35] due

to an entropic exclusion mechanism [36]. One example for the entropic clock is the voltage-gated potassium channel in axons. It exist in three main states: closed, open and inactive. The closed state converts to the open state in response to an appropriate voltage gradient, but within 1 ms the potassium channel becomes inactivated, even if the voltage gradient is maintained. The channel is inactivated by a short helical region at the end of a disordered linker, and it was shown that the timing of the inactivation depends on the length and disorder of the linker. The linker acts as an entropic clock and provides the potassium channel with a timer function [37]. Titin and elastin are two examples for entropic springs. Titin is a huge protein with more than 35.000 residues and disordered regions greater than 1,000 amino acids. It is suggested to fix the length of the muscle cell and may function as a rubberlike entropic spring that helps to restore overstretched muscle cells to their natural, relaxed length [24]. Elastin, a basic component in skin, blood vessels and elastic ligaments, also features the ability to contract reversibly after stretching [38].

1.2.3 IDPs in Disease

The abundance of IDPs in the proteome raises the question, how strongly intrinsic disorder is involved in diseases. In a recent review it was pointed out that several IDPs are involved in cancer, cardiovascular diseases and diabetes [39]. In addition, IDPs are involved in numerous neurodegenerative disorders and amyloidoses, originating from the conversion of specific proteins from their soluble functional states into stable, highly ordered, filamentous protein aggregates, known as amyloid fibrils. Initially, it was assumed that only certain proteins can undergo these structural changes. However, it was shown that many, even disease-unrelated, proteins can aggregate into amyloid-fibrils under appropriate conditions [?]. Fibril-formation implies a considerable structural rearrangement. For folded proteins, it requires the destabilization of its rigid native structure, leading to a partially unfolded state and a subsequently refolding into a toxic state. For IDPs the open structure as a prerequisite for fibril-formation is already fulfilled despite their transient interactions. Therefore, it is tempting to speculate if IDPs or proteins with long intrinsically disordered regions are overrepresented in amyloidogenic diseases [27]. As it is pointed out by Pechman et al [41] the logical consequence of protein-protein interface regions is the higher affinity to ag-

gregation compared to other surface regions. One well-studied example is the aggregation of tau into paired helical filaments. The aggregation depends on two hexapeptide motifs located at the beginning of repeat 2 and 3. These hexapeptide motifs, transiently populating β -sheet conformation, are responsible for both, regular attachment points to microtubules and aggregation initiators [42].

1.3 Tau As An Intrinsically Disordered Protein

1.3.1 Location and amino acid composition

The tau gene is located on the long arm of chromosome 17 (17q21.1), spans multiple introns and is alternatively spliced to produce 6 isoforms. In mammals, the family consists of the neuronal proteins MAP2 and tau and the non-neuronal protein MAP4 [43]. The 6 isoforms of tau, ranging between 352 to 441 residues, differ by two inserts near the N-terminal end, Insert I1 and I2, and the presence of either four or three imperfect repeat sequences in the C-terminal half of the protein [44, 45]. The projection domain, contains the first 150 N-terminal residues, many of them acidic. They project away from microtubules (MT) when tau is bound to, probably maintain inter-filament distances [33]. The projection domain contains the inserts I1 and I2, which are alternatively spliced, so that either two, one or zero inserts are present, depending on the isoform. The projection domain is followed by the proline rich regions P1 and P2 containing 22 prolines out of 93 residues. After the proline-rich region the microtubule binding region (MTBR) follows containing either 3 or 4 imperfect repeat regions (R1-R4), whereas R2 is alternatively spliced. This region is essential for MT binding. The MTBR has a basic character and each repeat ends with a PGGG motif. At the beginning of R2 and R3 two hexapeptide motifs are present which are essential for aggregation. Already early studies recognized that tau has a unusual resistance to heat and acid treatment, without losing its function and a very low content of secondary structure [46]. These properties are a result of the high fraction of basic and hydrophilic amino acid residues, typical for an IDP, which resist a compact folding. Several biophysical studies corroborated that tau is a prototypical intrinsically disordered protein [47]. In fetal

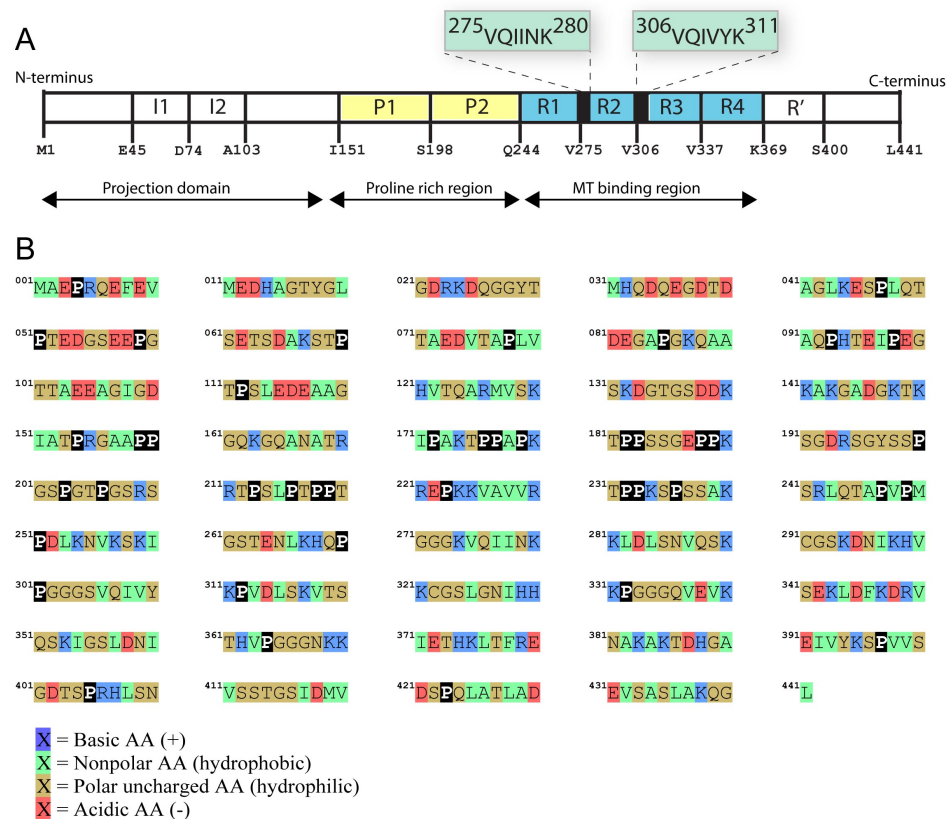


Figure 1.4: Domain organization of tau and sequence. A) The domain organization is shown with I1 and I2 as insert 1 and insert 2 in the projection domain, P1 and P2 as the two proline-rich regions and R1-R4 as the repeat regions, which are essential for MT-binding. At the beginning of R2 and R3 two hexapeptide motifs are essential for PHF formation. Repeat 2 is not present in all isoforms, because it is alternatively spliced. B) The sequence of full-length tau with 441 residues is shown. The projection domain, as well as the C-terminal part are negatively charged, whereas the repeat regions and the proline rich regions contain positively charged residues. In addition, the proline-rich regions contain 22 prolines, hence the naming.

brains only the shortest isoform without I1 and I2 and without R2 is present. In adult brains all isoforms are present with a ratio between the 3- and 4-repeat tau of 1:1 [48]. The fraction of the isoforms with two, one or zero inserts are 9%, 54% and 37%, respectively [48]. The conformation of tau in solution has been unknown but presumably highly variable, as it would be expected for a natively unfolded protein. FRET studies revealed a global

(average) paperclip-like folding of tau, wherein the repeat domains are in close proximity to the C-terminal and N-terminal ends of the molecule [49].

1.3.2 Translocation of tau

The neuron is a highly structured and specialized cell, with the dendrites and cell body receiving and processing information and the axon delivering information [50]. Cytoplasmic proteins destined for dendrites or axons carry nucleotide sequences, referred to as cis-acting elements, located in the 3'- or 5'-untranslated region (UTR) of their mRNAs. Trans-acting factor proteins have been found to bind cis-acting regions of mRNAs, thereby forming ribonucleoprotein particles (RNPs). They have been found associated with molecular motor proteins, which suggest that the latter may translocate them over cytoskeletal transport pathways [51]. Since tau is predominantly located in the axons of neurons in the central nervous system [52], whereas MAP2 is located in dendrites [53], a mechanism is held responsible for the non-uniformly distribution of tau and MAP2 [54, 55]. They found that a 240 bp sequence at the 3'-UTR of tau mRNA codes for an axonal localization signal, whereas a 3'-UTR MAP2 mRNA codes an dendritic localization signal [54, 55]. Swapping between the 3'-UTR tau mRNA axonal localization signal and the 3'-UTR MAP2 mRNA dendritic targeting signal resulted in the localization of tau in dendrites and MAP2 in axons [55]. However, these results were not confirmed by Utton et al [56] and they concluded that the mechanism found by Aronov et al might be restricted to the used embryonic carcinoma cells. Instead Utton et al argue that axonal transport is the source of differential localization of tau. In contrast, Hirokawa et al [53] argues that a differential turnover in axons and dendrites might be the responsible mechanism for the selective accumulation.

1.3.3 The physiological role of tau

Already in the early 1970s tau was implicated in the microtubule self-assembly process [57]. Microtubules are the biggest cytoskeletal components with a diameter of about 25 nm. They are assembled from heterodimers, α -tubulin and β -tubulin, and form a hollow cylindrical structure. The assembly generates two types of protein-protein contacts. Along the longitudinal axis of the microtubule, the "top" of one β -tubulin forms an interface with the "bot-

tom” of the α -tubulin in the adjacent heterodimer. This interface is very similar to the interface holding the α and β monomers together in the dimer subunit. Perpendicular to the longitudinal axis, neighboring protofilaments form lateral contacts between monomers of the same type (α - α and β - β). Together the longitudinal and lateral contacts are repeated in the regular helical lattice of the microtubule [3]. The head-to-tail arrangement of the tubulin-dimers results in a polarization of MTs with a plus-end (α -tubulin is at the end) and a minus-end (β -tubulin is at the end), which is also an important prerequisite for neuronal cell polarization. It is possible for purified tubulin to assemble with a wide range of diameters containing 9-18 protofilaments (Figure 1.5) [58].

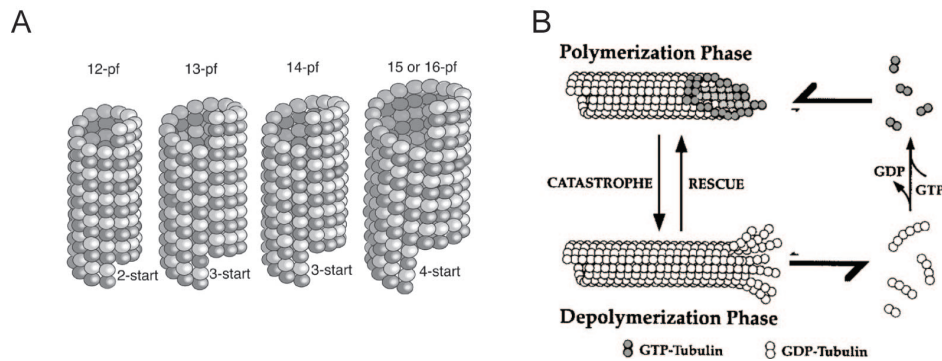


Figure 1.5: Behavior of Microtubules. A) Possible assembly of tubulin dimers to microtubules (from [58]). B) Microtubules are in a dynamic equilibrium between polymerization (called ”Rescue”) and depolymerization (called ”Catastrophe”). A GTP-tubulin cap stabilizes MT growth (from [59]).

MTs are very dynamic and the assembly and disassembly is determined by three factors: the critical concentration, the GTP hydrolysis and the binding of MAPs. In cells, MTs grow from the microtubule organization centre (MTOC) so that usually only the plus-end contributes to the dynamic of growing and shrinking. Is the tubulin-concentration below a critical level, MTs stop to grow and eventually start to shrink. Both, α -tubulin and β -tubulin, contain a guanosine triphosphate (GTP) molecule. The GTP bound to α -tubulin is non-exchangeable, but trapped between the 2 monomers of the heterodimer [58]. However, when heterodimers associate to form protofilaments, GTP on β -tubulin is hydrolyzed to guanosine diphosphate (GDP) as a consequence of the interaction with α -tubulin in the next dimer. The hydrolysis of GTP is usually very low, but will heavily increase during the

polymerization process. The hydrolysis of GTP to GDP releases free energy, making β -tubulin with GDP much more prone for dissociation. However, if the concentration is well above the critical concentration and the assembly therefore fast, the addition of another subunit occurs before the hydrolyzation and the MT grows (termed "rescue") [59]. MAPs, like tau, are the third factor. MAPs regulate MT-stability, because they effectively reduce the critical tubulin concentration. Where tau binds on MTs and whether tau compete or cooperate with other molecules (like the MT-stabilizing agent taxol) for binding site(s) is currently under debate. However, many studies revealed which parts of tau are essential to achieve this function. The MTBR contain either 3 or 4 imperfect repeats (R1-R4), essential for microtubule assembly. Moreover, in the absence of the two flanking regions P2 and R', the MTBR binds only weakly to microtubules. The flanking domains, on the other hand, bind to microtubules even in the absence of the repeats. This has led to the proposition of the jaws model of tau, wherein the regions flanking the repeats are responsible for positioning tau on the microtubule surface and the repeats are responsible for the catalytic activity, the microtubule assembly [60]. This is a typical behavior of an IDP, where several sites act synergistically to promote protein-protein binding. The NMR study of Mukrasch et al [61] showed the previous result on a per-residue level. The NMR experiments revealed that the residues $^{275}\text{VQIINKKLDLS}^{285}$ (comprising the hexapeptide in the beginning of repeat 2) show large chemical shift changes. In addition, clusters of positive charges close to the PGGG motifs at the end of repeats 1, 2, and 3 contributed to the MT binding, but not repeat 4. Moreover, $^{225}\text{KVAVVRT}^{232}$ and $^{240}\text{KSRLQTAPV}^{248}$ in the proline-rich domain P2 and $^{370}\text{KIETHKLTFRN}^{381}$ in R' are binding sites for MTs. In addition, site-directed mutagenesis at Lys224, Lys225, and Arg230 confirmed the role for P2 [62]. Interestingly, the tau protein contains several stretches of residual structure. Several NMR studies showed that at the beginning of repeats R2, R3 and R4 β -sheet propensities were found for residues $^{273}\text{Gly-Leu}^{284}$, $^{304}\text{Gly-Leu}^{315}$, and $^{333}\text{Gly-Asp}^{345}$, respectively [42, 61, 63]. If these preformed structural elements precede a rigid secondary structure upon binding, is unknown.

In vitro MT-tau binding assays often use preassembled MTs that are stabilized with taxol [64–67]. In a study, where MTs are co-assembled with tau, it was pointed out that preformed taxol-stabilized MTs may not fully meet physiological conditions, because taxol compete for a second binding

site of tau, located inside of MTs [68]. A model, based on cryo-electron microscopy (cryo-EM), proposes tau's binding on the inner and outer surface of MTs. It suggests that the repeat regions R1-R4 bind to the inner surface of MTs laterally whereas the N-terminal projection domain connects different heterodimers longitudinally [58]. In another cryo-EM study, tau was tagged with green-fluorescence protein (GFP) either at the C-terminus or N-terminus. The combination of metal-shadowing, cryo-EM and the globular protein-domain of GFP revealed, that GFP showed up only on the outer surfaces of MTs, no matter whether MTs were preformed or polymerized together with GFP-tau [69]. In addition, the authors found that tau binds along the protofilaments, as well as across. This is consistent with previous studies [66, 70] but contradicts the study by Kar et al. [68]. That there is evidence for a second tau-microtubule binding site, which might not coincide with taxol, was revealed by kinetic studies using fluorescence resonance energy transfer (FRET) [71]. Acrylodan-labeled tau (at residues C291 and C322) served as an acceptor of energy transfer from donor tryptophans from MTs. Interestingly, tau showed two binding behaviors. If tau was added to preassembled MT, with or without taxol, weak and exchangeable tau binding to MTs was observed. If tau and MTs were co-assembled, tau showed a strong and non-exchangeable binding behavior in addition to the weak and exchangeable behavior. The authors speculated that during co-assembling, tau may become irreversibly incorporated into the MT wall, whereas tau added to pre-assembled MTs is bound to the surface of MTs and remains exchangeable with the medium [71].

1.3.4 Phosphorylation of tau

The binding and release of tau from MTs is regulated via phosphorylation. Although glycosylation, glycation, ubiquitination, sumoylation and nitration have been reported for tau, phosphorylation plays the primary role [72]. Tau in its longest isoform with 441 residues contains 85 potential phosphorylation sites (45 serines, 35 threonines and 5 tyrosines) (Figure 1.6). Interestingly, the phosphorylation pattern is very dynamic and, in addition, also varies during the maturation from a neurite to a neuron. In a report from 1994 the dynamic changes in the phosphorylation state of tau in developing rat brains were investigated. They show that fetal tau, the shortest isoform, is strongly phosphorylated at Thr172, Ser193, Thr222, Ser387 and Ser395, and that the

phosphorylation of Ser193, Thr222, Ser387 persisted until postnatal day 19. Thereafter, these sites were dephosphorylated with a concomitant increase in the levels of the other five mature tau isoforms. However, phosphorylation at Ser387, Ser395 and Thr172 persisted through to the adult rat brain [73].

Numerous protein kinases have been found phosphorylating tau *in vitro*, like glycogen synthase kinase-3 β (GSK-3 β), cyclin-dependent kinase 2 and 5 (cdk2 and cdk5), cAMP-dependent protein kinase A (PKA) and protein kinase C (PKC), Ca²⁺/ Calmodulin-dependent kinase (CaMK), casein kinase 1, mitogen-activated protein kinase (MAPK), microtubule affinity regulating kinase (MARK), fyn and several others [74]. Especially phosphoryla-

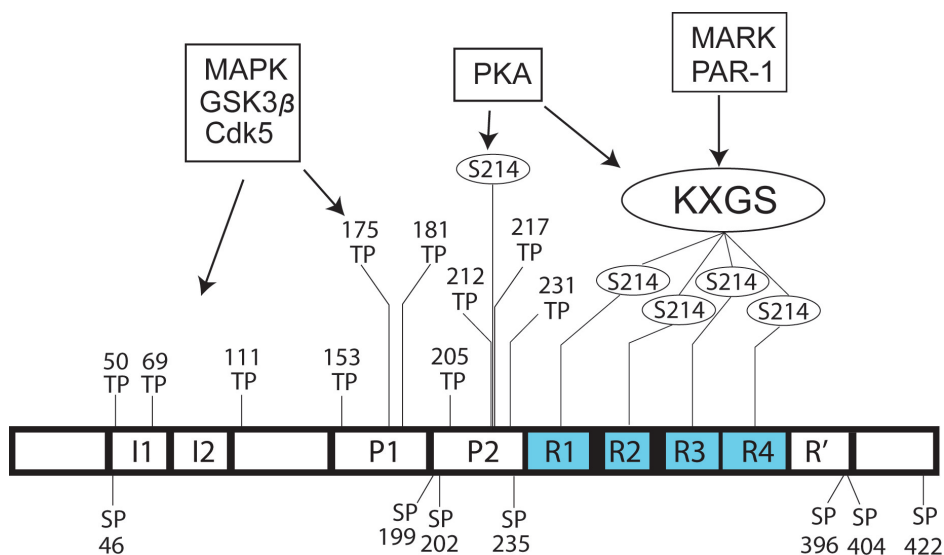


Figure 1.6: Phosphorylation sites of tau. Tau contains numerous Ser and Thr residues, many of which show abnormally high phosphorylation in Alzheimers disease and are diagnostic of Alzheimer tau. Phosphorylation within the repeats (at KXGS motifs by the kinase MARK/Par-1) efficiently detaches tau from microtubules (taken from [75]).

tions within the MTBR have dramatic effects, since they detaches tau from MTs [76]. Experiments also suggested the inter-dependence of phosphorylation sites, meaning that phosphorylation of tau by one kinase modulates the subsequent phosphorylation by other kinases [77]. However, the number of protein kinases that actually phosphorylate tau *in vivo* is probably much lower [78]. MARK [79], PKA [80], cdk5/p35[p25] [81] and GSK-3 β [82] have been shown to phosphorylate tau *in vivo*. GSK-3 β is expressed ubiq-

uitously and found at high levels in brain cells where it localizes predominantly in neurons. GSK-3 β is a proline-directed kinase. It phosphorylates unprimed sites at Ser/Thr-Pro motifs. In other cases the substrate must be phosphorylated first at a site that is four amino acids downstream to the target site, to prime the substrate for phosphorylation by GSK-3 β . This primed motif ((S/T)XXX(S/ T)[P]) often occurs in proline-rich regions of the substrate [83]. Interestingly, GSK-3 β phosphorylates tau at both primed (Thr231) and unprimed sites (Ser396 and Ser404), whereas only the primed sites decreases tau-microtubule interactions [84]. In another study additional phosphorylation sites were identified using different antibodies, whereas the most prominent are Ser46 and Thr50 in insert 1, Ser202 and Thr205 in the proline rich region 1 of tau [85]. Minor phosphorylation sites were Thr181, Ser199, Thr205, Thr212, Thr205 and Ser235. Although the majority of tau phosphorylation sites are Ser/Thr-Pro sites, Ser and Thr sites not followed by Pro residues are also phosphorylated by non-proline directed kinases. Data indicate that tau is phosphorylated by cAMP-dependent protein kinase (PKA) *in vivo*. In fact, it was shown that PKA primes tau for phosphorylation by GSK-3 β [86]. An *in vitro* study showed that pre-phosphorylation with PKA at Ser214, Ser262 and Ser409 resulted in an enhanced GSK-3 β -phosphorylation at Thr181, Ser199, Ser202, Thr205, Thr217, Thr231, Ser396, and Ser422, but inhibited the phosphorylation at Ser404 [87]. Subsequent phosphorylation with cdk5 revealed a completely different behavior. The PKA-prephosphorylation showed no detectable effect on tau phosphorylation by cdk5 at Thr181, Ser199, Thr205, Thr231 and Ser422, inhibited it at Ser202, Thr212, Thr217 and Ser404, and slightly promoted it at Ser396 [87]. Interestingly, Ser202 and Ser396 are phosphorylated in fetal human brain, but not in adult brains [88,89]. Besides the priming-effect PKA exhibit for GSK-3 β , PKA is able to phosphorylate the KXGS motifs, located in the MTBR. These are Ser262, Ser293, Ser324, and Ser356, leading to a strongly reduced tau-MT binding [90]. Actually, the phosphorylation of Ser262 alone is sufficient to reduce MT-affinity dramatically [91]. Interestingly, Landrieu et al., who investigated the PKA-phosphorylation of tau with NMR, found Ser208, Ser214, Ser324, Ser356, Ser409 and Ser416 to be phosphorylated, but not Ser262 and Ser293 [92], presumably due to an incomplete assignment. Like PKA, the microtubule affinity regulating kinase (MARK) phosphorylates the KXGS motifs. MARK, a calcium/calmodulin-dependent protein kinase, consist of an N-terminal catalytic domain, followed by the ubiquitin-associated

domain, spacer, and tail domains (kinase-associated, KA, domain). The size of MARK and its multidomain composition suggests that regulation could take place on several levels. The interest in the KXGS motifs comes from the observation that the detachment of tau from MTs is achieved most efficiently by phosphorylating the serines in these motifs. It is believed that this highly dynamic behavior is crucial for neurite outgrowth and establishing neuronal polarity [79]. Point mutations in the KXGS motif to a non-phosphorylatable KXGA motif showed no stimulation for neurite outgrowth, in contrast to wildtype tau [79]. Taken together, these results suggest that unstable MTs are necessary for neurite outgrowth. Another interesting aspect is the interplay of tau with MAP1b, another MAP found in axons. MAP1b and tau can substitute for one another during axonal growth [93]. A knock off tau and MAP1b transgenic mouse model resulted in severe defects in axonogenesis [94,95]. Interestingly, whereas phosphorylation of tau at the KXGS sites with MARK detaches tau from microtubules, phosphorylation at certain sites of MAP1b with GSK-3 β promotes the binding to microtubules. In other words, the inhibition of MARK reduces microtubule dynamics, whereas the inhibition of GSK-3 β enhances dynamics [79].

It is worth mentioning that the complexity of the topic produced a picture which is not always fully consistent. These contradictions might reflect the limitations inherent to certain methods, which for example require prior knowledge in great detail, for example antibody (AB) recognition essays. Many experiments determine a phospho-epitope with an AB specifically raised against this phosphorylation site, whereas the antibodies often vary in affinity and specificity [79]. The analysis of phosphorylation patterns might also remain incomplete due to the absence of AB against certain unknown phospho-epitopes. For example the study of Davies et al. detected only two phosphorylation sites in tau when phosphorylated with PKA, because only 2 AB against S214 and S409 were used [96]. In another study, Liu et al. could only speculate if after PKA phosphorylation, GSK-3 β might also phosphorylate Ser210 and Ser258, due to the non-availability of AB against these phospho-epitopes [87]. Moreover, in a recent NMR study it was shown, that the CDK5/p25 is not catalytically more active compared to CDK5/p35, in contrast to previous studies. Instead, both enzymes display virtually identical kinetics towards individual phosphorylation sites in tau. Therefore, the authors of the study concluded, that cellular toxicity of CDK5/p25 in vivo can not be ascribed to significant differences in the catalytic parameters [97].

These ambiguities show, that the study of post-translational modifications demand high-resolution methods, like NMR or mass spectrometry.

1.3.5 The pathological role of tau

The reason why phosphorylation of tau is so extensively investigated is not only due to its regulatory role for MT-binding. Hyperphosphorylated tau, aggregated into paired (or straight) helical filaments (PHFs) is the hallmark of a wide variety of neurodegenerative diseases, termed tauopathies [98]. The aggregation of tau seems rather counterintuitive due to its high solubility to at least 60 mg/ml. In addition, the amino acid composition should prevent the formation of a rigid secondary or tertiary structure. In tauopathies, like Alzheimer's disease (AD), tau exhibits an unusual degree of phosphorylation which is 3-4 times higher [76]. The reason, despite a large number of publications, is still not fully understood. *Per se*, phosphorylation is not necessarily a pathogenic phenomenon. The highest phosphorylation of tau can be found in fetal brains, maintaining a highly dynamic behavior. Even hyperphosphorylation seems to serve a physiological purpose in hibernating animals [99, 100]. This further complicates the relationship of hyperphosphorylation and aggregation. In addition, the relative proportions of tau isoforms observed in AD PHFs are similar to those that are characteristic of the six soluble tau isoforms observed in healthy adult human brain [48].

However, the way how tau finally forms neurofibrillary tangles (NFTs), which are found post-mortem in AD brains, is quite well characterized. This multistep-process involves a structural transition from random coil to β -sheet for several hydrophobic stretches [102]. The assembly of hydrophobic patches forms prefibrillar oligomers, which evolve to PHFs. PHFs assemble further to large lightmicroscopically visible neurofibrillary tangles (NFTs).

Two hexapeptides, $^{275}\text{VQIINK}^{280}$ and $^{306}\text{VQIVYK}^{311}$, at the beginning of R2 and R3 are crucial for the initial transformation. The hydrophobicity and β -sheet propensity [103] serves usually for MT-binding, but can also initiate β -sheet formation, supporting the theory of Pechmann et al. [41]. Importantly, $^{275}\text{VQIINK}^{280}$ and $^{306}\text{VQIVYK}^{311}$ alone are able to form aggregates [104, 105], whereas proline-mutations (I277P and I308P) within the hexapeptides are able to inhibit aggregation [106]. An increasing body of evidence suggest that the fibril-core forms a cross- β -structure, indicated by a 0.47nm ring in X-ray diffraction patterns [75, 105, 107]. Previous reports of

Frontotemporal dementia and parkinsonism linked to chromosome 17
Alzheimers disease
Progressive supranuclear palsy
Picks disease
Argyrophilic grain dementia
Corticobasal degeneration
Progressive subcortical gliosis
Amyotrophic lateral sclerosis/parkinsonism-dementia complex
Diffuse neurofibrillary tangles with calcification
Dementia pugilistica
Tangle-only dementia
Downs syndrome
Gerstmann-Straussler-Scheinker disease
Hallervorden-Spatz disease
Creutzfeldt-Jakob disease
Multiple system atrophy
Niemann-Pick disease type C
Prion protein cerebral amyloid angiopathy
Subacute sclerosing panencephalitis
Myotonic dystrophy
Non-guanamian motor neuron disease with neurofibrillary tangles
Postencephalitic parkinsonism

Table 1.1: Neurodegenerative diseases caused by tau. Tau is involved in numerous neurodegenerative diseases, with Alzheimers disease as the most prominent one. Hyperphosphorylated tau incorporated into insoluble aggregates is a common feature in all diseases. However, the composition and morphologies of aggregates is different (from [101]).

an α -helical fibril-core [108, 109] could not be verified.

The β -sheet propensity of the hexapeptide motifs is a result of an alternating pattern of polar and apolar amino acids, separated by pattern-breaking sequences [106]. The impact of the alternating pattern is underscored by mutation studies, where the pattern-breaking sequences are mutated to prolong the alternating pattern. The expression of this structurally enforced protein in cells led to an almost instantaneous aggregation of the protein associated with cytotoxicity [110]. This also explains the enhanced aggregation of the familial tau mutation Δ K280 found in frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). The deletion Δ K280 causes a flip of the second part of the motif so that almost all of the hydrophobic side chains point to the bottom, thus increasing the hydrophobicity and the tendency for aggregation on this face of the sheet [106].

Protease digestion and solvent accessibility studies demonstrated that the core of PHFs is mainly built from the repeat sequences in the C-terminal half of tau, containing the two hexapeptides that are crucial for PHF aggregation [111]. Furthermore, this and other studies showed that the core is surrounded by a fuzzy coat. Protease digestion indicated that the fuzzy coat comprises the N-terminal half of the protein as well as the C-terminus with more than 200 residues in total. The fuzziness in EM pictures, together with EPR experiments suggest a rather flexible nature of the N- and C-terminal fuzzy coat [112]. The fuzzy coat seems to play an important role in aggregation and promoting cell death. It is shown, that the C-terminus of tau has an inhibitory effect on polymerization [113]. The N-terminus holds an ambiguous role. On the one hand it seems to inhibit fibril formation [114, 115], on the other hand it might promote cell-death after a successful formation of fibrils [116].

What drives the conformational transformation where no mutations are involved? Filamentous tau from AD brains are extensively phosphorylated and there is an ongoing debate on how phosphorylation and aggregation are related [117]. The question is, if hyperphosphorylation precedes or follows aggregation. Many investigations pointed out, that a high degree of phosphorylation detaches tau from MTs, thereby increasing free tau which is potentially prone for structural transitions [98, 101]. Especially the known FTDP-17 mutations either disrupt MT binding or leading to a tau isoform imbalance, result in an increase in free unbound tau [118]. Also overexpression was suggested as a mechanism for an enhanced fraction of free tau in AD-

affected brains. However, considering the concentration of tau in neurons, which is typically between 1-4 μM [119], and the high solubility of tau *in vitro* to more than 1000 μM , it is still unclear which role increased levels of free tau play for PHF formation. Due to the high solubility of tau, it was generally difficult to study the formation of PHFs without introducing aggregation-enhancing mutations. The discovery of so called "aggregation enhancer" greatly facilitated PHF-formation studies [75]. Heparin [120], RNA [121] and arachidonic acid [122] are three potent tau-aggregation enhancing molecules. Interestingly, the only common feature of all three molecules is a high negative charge, indicating that not a distinct structure but a modified electrostatic environment causes tau-aggregation. This might be similar to *in vivo* processes, where hyperphosphorylation results in a high number of negative charges (depending on the pH, one phosphate adds 1-2 negative charges), possibly able to overcompensate basic repulsion of the proline-rich region and repeat-regions. This might enable close tau-tau interactions [123], eventually leading to interactions of intermolecular hydrophobic patches (hexapeptides). The role of the C-terminus in tau-aggregation is further elucidated in publications investigating unusual protease activity. It was shown that the cleavage of the C-terminus of tau at Asp421 enhances the polymerization of tau [101, 113, 124].

1.3.6 The interplay of tau and $A\beta$ in Alzheimer's disease

Alzheimer's disease is characterized by two hallmarks, the extracellular amyloid plaques, mainly composed of a protein called beta-amyloid ($A\beta$), and the already discussed intracellular neurofibrillary tangles (NFTs), composed of hyperphosphorylated tau. The question arises what is the connection of tau and $A\beta$? That there is a connection in AD pathology was demonstrated in mouse models [125, 126], where signs of degeneration occurred when tau was co-expressed with $A\beta$, but not in tau knockout neurons.

Hyperphosphorylation of tau was explained by either up-regulated kinases, down-regulated phosphatases, or both [48]. However, another possibility might be a mislocation of tau. Already early investigations suggested that tau might not necessarily remain in axons in AD [127]. Recently, this idea was picked up again and it was speculated that relocation of tau from the axonal to the somatodendritic compartment, where neurofibrillary tan-

gles become apparent first, might be a key event in tau pathology [128]. It is tempting to speculate that "molecular fences", which restrict the movement of membrane proteins in neurons to maintain cell polarity [50], have cytosolic counterparts which restrict cytosolic protein movements to maintain cytosolic polarity. In fact, microtubules and microfilaments are considered to undertake the task.

Their intrinsic polarity translates asymmetry in a developing neuron. In a study from 2008, Witte et al showed, that local MT stabilization with taxol in one neurite in undifferentiated neurons causes axon formation [129]. Therefore, the destabilization of MTs, for example due to the detachment of tau, might be sufficient to induce the breakdown of a cytosolic cell polarity. This seems to be the case in AD, where the predominant localization of tau in axons is disrupted [130]. In a recent paper it was shown that the mislocation of endogenous tau in the somatodendritic compartment was triggered by A β -oligomers, transfected into hippocampal cells, establishing a connection between A β and tau in AD [131]. They showed that tau mislocation resulted in severe neuronal damage due to the disappearance of dendritic spines from affected dendrites and a higher phosphorylation of tau. Interestingly, tau missorting also occurred as a response to several types of stressors, like H_2O_2 , serum deprivation, glutamate and extracellular ATP [131]. Another mechanism was proposed from Oddo et al [132]. They found that A β oligomerisation starts intraneuronally, linking A β and tau pathology. Another study cleared the details of this linkage by demonstrating with Western blotting, ELISA and surface plasmon resonance the formation of a stable tau-A β complex with nanomolar affinity [133]. A different mode of interaction is suggested by Ittner et al [134]. They propose a dendritic function of tau, the targeting of the tyrosine kinase Fyn to the dendritic spines. The authors argue that the breakdown of cytosolic cell polarity increases the concentration of tau in dendrites, which is already present, even in healthy neurons. The increased concentration of Fyn in spines due to an increased concentration of tau, stimulates a cascade of events which mediates excitotoxicity stimulated by extracellular A β -oligomers (reviewed in [135]).

Chapter 2

Nuclear Magnetic Resonance

2.1 Introduction

The resistance to heat, a wide variation in pH, ionic strength, solvent polarity and denaturation [136] are unique characteristics of the intrinsically disordered protein tau. Several biophysical techniques were used to investigate and classify this disordered protein. For example differential scanning calorimetry showed no denaturation of secondary structure upon heating, but found some poly-proline II helical conformations with Raman spectroscopy and were not able to exclude residual β -structures [137]. Circular dichroism spectroscopy (CD spectroscopy) also probes globally for the presence of rigid secondary structure elements and was used to illustrate the absence of rigid secondary structure [136]. Small angle x-ray scattering (SAXS) showed a unusual high radius of gyration for tau, indicating a high level of disorder [47]. Fluorescence resonance energy transfer is another method used to investigate the global folding of tau [49].

Unfortunately, the aforementioned methods are limited to a global investigation, making the interpretation of data from IDPs in general difficult. Since many apparently fully disordered proteins possess stretches of residual secondary structure, the "low resolution" methods can often not fully resolve all the "hidden" structural and dynamical characteristics of an IDP. Currently, the only two methods capable of structure determination on an atomic level are X-ray crystallography and nuclear magnetic resonance (NMR). However, the high flexibility prevents crystallization of the protein and even if crystals of IDPs can be formed, the conformational variability of the backbone would make data interpretation very challenging. Therefore, NMR is currently the

only method which possesses the power to resolve and analyze IDPs with atomic resolution.

2.2 Nuclear Magnetism

A fundamental quantity associated with neutrons, protons and electrons is the spin angular momentum. The introduction of an additional "spin" angular momentum, not arising from orbital motion, like the orbital angular momentum, was demanded from Stern-Gerlach experiments [138]. Whereas in electron paramagnetic (or spin) resonance (EPR or ESR) spectroscopy the spin angular momentum of the unpaired electron is of interest, NMR deals with the spin angular momentum of the nuclear proton. The angular momentum, represented by the operator \hat{I} , is linked to its magnetic moment μ via the gyromagnetic ratio γ [139]:

$$\mu = \gamma \hbar \hat{I}, \quad (2.1)$$

with $\hbar = \frac{h}{2\pi}$ as the Planck constant. The gyromagnetic ratio is a constant and characteristic for each nucleus, resulting from the gyromagnetic ratios of protons and neutrons. The angular momentum operator is a vector quantity specified by the three cartesian components \hat{I}_x , \hat{I}_y , \hat{I}_z and a magnitude \mathbf{I}^2 . However, due to quantum mechanical restrictions, the magnitude \mathbf{I}^2 and only one cartesian component, by convention \hat{I}_z , can be specified simultaneously [140]. When placed into a magnetic field, \mathbf{B} , the magnetic moment interacts with the \mathbf{B} -field and was revealed that the spin angular momentum is quantized into $(2I+1)$ states [139]. For the proton, a spin- $\frac{1}{2}$ particle ($\hat{I}=\frac{1}{2}$), two states occur:

$$\begin{aligned} E(\alpha) &= -\frac{1}{2} \hbar \gamma B_0 \\ E(\beta) &= +\frac{1}{2} \hbar \gamma B_0 \end{aligned} \quad (2.2)$$

Whereas B_0 is the static external magnetic field along the z-direction. The energetically different α and β states are split according to the Zeeman effect and the associated energy difference ΔE is [141]:

$$\Delta E = \hbar \gamma B_0 \quad (2.3)$$

With Bohr's frequency condition, $\Delta E = h\nu = \hbar\omega_0$, the so called Larmor frequency ω_0 can be derived:

$$\omega_0 = \gamma B_0. \quad (2.4)$$

Irradiation of the spins with radio-frequency pulses matching the energy difference between the α and β states allows the manipulation of the nuclear angular momentum (spins) in many different ways. This is described by the product operator formalism [142, 143]. Generally, spins can be rotated by applying radio-frequency pulses, spins can evolve a coupling to neighboring spins and spins can transfer their magnetization to other spins either via the intervening space or via bonds. Since the sample contains a very large number of spins the macroscopic magnetization can be seen and described as an ensemble average of all spins. The use of the density operator $\hat{\rho}$ provides a powerful tool to describe the time evolution of the spin ensemble [141, 144, 145]. The solution to the *Liouville-von Neumann* equation:

$$\frac{d\hat{\rho}(t)}{dt} = i [\hat{\rho}(t), \hat{\mathcal{H}}] \quad (2.5)$$

gives the evolution in time of the density operator:

$$\hat{\rho}(t) = \exp(-i\hat{\mathcal{H}}t)\hat{\rho}(0)\exp(i\hat{\mathcal{H}}t), \quad (2.6)$$

with $\hat{\rho}(t)$ as the density operator at time t , $\hat{\rho}(0)$ as the density operator at zero time and $\hat{\mathcal{H}}$ is the hamilton operator (Hamiltonian). The Hamiltonian adopts different forms, depending if a pulse, a delay or other actions with neighboring spins alter the density operator.

In the following, the effect of a pulse on the density operator is described which finally leads to the product operator rules used in NMR [146]. The cartesian spin angular momentum operators (spin operators) can be represented in matrix form by using the Pauli spin matrices [141]:

$$\hat{I}_x = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{I}_y = \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{I}_z = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (2.7)$$

They obey cyclic commutation rules and are orthonormal. A pulse rotates the macroscopic magnetization, therefore the $\exp(-i\hat{\mathcal{H}}t)$ terms from eq. 2.6 are represented by the following pulse operators [141, 145]:

$$\hat{R}_x(\theta) = \begin{bmatrix} c & -is \\ -is & c \end{bmatrix} \quad \hat{R}_y(\theta) = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \quad \hat{R}_z(\theta) = \begin{bmatrix} c - is & 0 \\ 0 & c + is \end{bmatrix} \quad (2.8)$$

Where $c = \cos(\frac{\theta}{2})$ and $s = \sin(\frac{\theta}{2})$.

The initial magnetization at equilibrium is $\hat{p}(0) = \hat{I}_z$. A $\theta = 90^\circ$ pulse with y-phase, applied to the initial magnetization results in:

$$\begin{aligned} \hat{p}(t) &= \hat{R}_y(\theta) \hat{I}_z \hat{R}_y^{-1}(\theta) = \frac{1}{2} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned} \quad (2.9)$$

which is \hat{I}_x . That demonstrates that a 90° pulse with y-phase, acting on z-magnetization \hat{I}_z , results in x-magnetization \hat{I}_x .

The evolution of the \hat{I}_x magnetization under the free precession Hamiltonian:

$$\hat{\mathcal{H}} = \Omega \hat{I}_z, \quad (2.10)$$

with Ω as the chemical shift in the rotating frame, is (with $\hat{p}(0) = \hat{I}_x$):

$$\begin{aligned} \hat{p}(t) &= \exp(-i\Omega \hat{I}_z t) \hat{I}_x \exp(i\Omega \hat{I}_z t) \\ &= \frac{1}{2} \begin{bmatrix} 0 & \cos\Omega t \\ \cos\Omega t & 0 \end{bmatrix} + \frac{i}{2} \begin{bmatrix} 0 & -\sin\Omega t \\ \sin\Omega t & 0 \end{bmatrix} \end{aligned} \quad (2.11)$$

This equals:

$$\hat{I}_x = \hat{I}_x \cos(\Omega t) + \hat{I}_y \sin(\Omega t). \quad (2.12)$$

This result represents that the spin polarization, originally along the x-axes, evolves a y-component in time, similar to a precession in the x-y plane. For a one- and two-spin system, the evolution of the product operators are summarized (see e.g. [141]) :

- Free precession
one-spin system

$$\begin{aligned} \hat{I}_z &\xrightarrow{\Omega_I \hat{I}_z t} \hat{I}_z \\ \hat{I}_x &\xrightarrow{\Omega_I \hat{I}_z t} \hat{I}_x \cos(\Omega_I t) + \hat{I}_y \sin(\Omega_I t) \\ \hat{I}_y &\xrightarrow{\Omega_I \hat{I}_z t} \hat{I}_y \cos(\Omega_I t) - \hat{I}_x \sin(\Omega_I t) \end{aligned}$$

two-spin system (with coupling)

$$\begin{aligned}
\hat{I}_z &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} \hat{I}_z \\
\hat{I}_x &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} \hat{I}_x \cos(\pi J_{IS}t) + 2\hat{I}_y\hat{S}_z \sin(\pi J_{IS}t) \\
\hat{I}_y &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} \hat{I}_y \cos(\pi J_{IS}t) - 2\hat{I}_x\hat{S}_z \sin(\pi J_{IS}t) \\
2\hat{I}_z\hat{S}_z &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} 2\hat{I}_z\hat{S}_z \\
2\hat{I}_y\hat{S}_z &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} 2\hat{I}_y\hat{S}_z \cos(\pi J_{IS}t) - \hat{I}_x \sin(\pi J_{IS}t) \\
2\hat{I}_x\hat{S}_z &\xrightarrow{2\pi J_{IS}\hat{I}_z\hat{S}_z t} 2\hat{I}_x\hat{S}_z \cos(\pi J_{IS}t) + \hat{I}_y \sin(\pi J_{IS}t)
\end{aligned}$$

- Pulses

phase $\pm x$

$$\begin{aligned}
\hat{I}_z &\xrightarrow{\theta\hat{I}_{\pm x}t} \hat{I}_z \cos(\theta) \mp \hat{I}_y \sin(\theta) \\
\hat{I}_y &\xrightarrow{\theta\hat{I}_{\pm x}t} \hat{I}_y \cos(\theta) \mp \hat{I}_z \sin(\theta) \\
\hat{I}_x &\xrightarrow{\theta\hat{I}_{\pm x}t} \hat{I}_x
\end{aligned}$$

phase $\pm y$

$$\begin{aligned}
\hat{I}_z &\xrightarrow{\theta\hat{I}_{\pm y}t} \hat{I}_z \cos(\theta) \pm \hat{I}_x \sin(\theta) \\
\hat{I}_y &\xrightarrow{\theta\hat{I}_{\pm y}t} \hat{I}_y \\
\hat{I}_x &\xrightarrow{\theta\hat{I}_{\pm y}t} \hat{I}_x \cos(\theta) \mp \hat{I}_z \sin(\theta)
\end{aligned}$$

2.3 ${}^3J_{H^N H^\alpha}$ Scalar Couplings

Two nuclear spins are able to exchange information either directly via the intervening space or indirectly over the bonds via the electrons. The first coupling interaction is called the (direct) dipolar coupling, whereas the second mechanism is called scalar coupling. The scalar coupling can be transmitted to up to 9 bonds (9J) in small molecules via the Fermi contact interaction between the magnetic dipole moment of the nucleus and the magnetic dipole moment of the s-electrons [139, 147]. In proteins the angular dependence

of the scalar couplings is used to probe for backbone conformations. For example the one-bond coupling between C^α and its attached H^α ($^1J_{C^\alpha H^\alpha}$) primarily depends on the dihedral angle Ψ and can probe reliably for residues in α -helical conformation [148]. Another scalar coupling, the homonuclear three-bond $^3J_{H^N H^\alpha}$ scalar coupling, contains valuable secondary structure information due to the intervening dihedral backbone-angle Φ [149]. The vicinal proton-proton coupling is related to the Karplus equation, making it possible to relate the magnitude of $^3J_{H^N H^\alpha}$ couplings to structural parameters [150]:

$$^3J_{H^N H^\alpha} = A \cos^2(\phi - 60) - B \cos(\phi - 60) + C. \quad (2.10)$$

For measuring the couplings the intensity-modulated heteronuclear single-

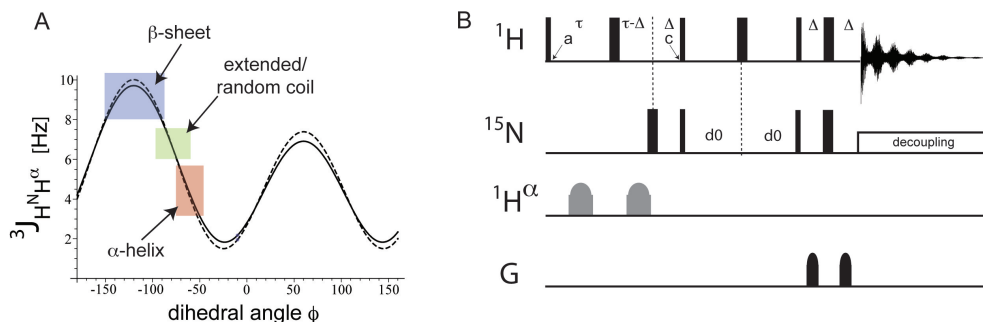


Figure 2.1: Vicinal proton-proton coupling A) Karplus equation for the dihedral angle ϕ with two different sets of parameters (dashed line: $A=6.4$, $B=-1.4$, $C=1.9$ from [151]; solid line: $A=7.13$, $B=-1.31$, $C=1.56$, from [152]). Also indicated are the regions of the secondary structure elements: 8-11 Hz found in β -sheet structures, 4-5 Hz for α -helical structures, according to [153]. B) Pulse sequence of the intensity-modulated HSQC (IM-HSQC). The grey shaped pulses for H^α are only applied for the decoupled experiment.

quantum coherence (IM-HSQC) experiment can be used. The advantage of using this experiment lies in the fast recording time due to its 2-dimensionality [153]. In addition, it is sufficient to use a single-labeled (^{15}N) sample. The product operator formalism for the IM-HSQC experiment will be discussed now. The first 90° pulse on protons with phase x (a) results in transverse proton magnetization

$$H_z^N \xrightarrow{90_x^\circ} -H_y^N \quad (2.11)$$

The 180° (broadband) pulse on protons separating the two τ evolution times (τ and $\tau-\Delta+\Delta=\tau$) results in refocusing the chemical shift evolution for H_y^N

and H_y^α . Therefore, only coupling to H_z^α occurs during 2τ :

$$-H_y^N \xrightarrow{\text{coupled}} H_y^N \cos({}^3J_{HNH^\alpha}2\tau\pi) + -2H_x^N H_z^\alpha \sin({}^3J_{HNH^\alpha}2\tau\pi). \quad (2.12)$$

The in-phase and anti-phase terms are modulated due to the ${}^3J_{HNH^\alpha}$ scalar coupling with a cosine and sine function, respectively. In the second delay τ , a 180° nitrogen pulse between the two delays Δ allows the coupling of the amide protons to the attached nitrogen:

$$\begin{aligned} &\xrightarrow{\text{coupling}} H_y^N \cos({}^3J_{HNH^\alpha}2\tau\pi) \cos(\pi J_{HN}2\tau) \\ &\quad - 2H_x^N N_z \cos({}^3J_{HNH^\alpha}2\tau\pi) \sin(\pi J_{HN}2\tau) \\ &\quad - 2H_x^N H_z^\alpha \sin({}^3J_{HNH^\alpha}2\tau\pi) \cos(\pi J_{HN}2\tau) \\ &\quad - 2H_y^N H_z^\alpha N_z \sin({}^3J_{HNH^\alpha}2\tau\pi) \sin(\pi J_{HN}2\tau) \end{aligned} \quad (2.13)$$

The delays Δ are optimized to transfer magnetization to the heteronucleus, therefore only terms 2 and 4 are retained from eq. 2.13. In addition, only the second term will give observable single-quantum coherence during the acquisition. Therefore, at time point (c) the only term we have to consider is:

$$-2H_x^N N_z \cos({}^3J_{HNH^\alpha}2\tau\pi) \sin(\pi J_{HN}2\tau) \quad (2.14)$$

In principle, the part described above can be seen as a slightly modified INEPT building block in a normal HSQC experiment [154]. In the decoupled experiment, no coupling to the H^α can evolve:

$$\begin{aligned} &\xrightarrow{\text{decoupled}} H_y^N \cos(\pi J_{HN}2\tau) \\ &\quad - 2H_x^N N_z \sin(\pi J_{HN}2\tau) \end{aligned} \quad (2.15)$$

Since the product operator formalism is the same for the coupled and decoupled experiment, the term of interest at time point (c) is:

$$-2H_x^N N_z \sin(\pi J_{HN}2\tau) \quad (2.16)$$

After time-point (c) the coupled and decoupled experiments are similar. As it can be seen from eq. 2.14 and 2.16, the two terms of interest differ by a cosine-modulation of the coupled experiment, which results in a lower signal intensity. Therefore, the intensity difference is related to the ${}^3J_{HNH^\alpha}$ scalar coupling via:

$$\frac{I_{\text{coupled}}}{I_{\text{decoupled}}} = \cos({}^3J_{HNH^\alpha}2\tau\pi), \quad (2.17)$$

or:

$${}^3J_{HNH^\alpha} = \frac{\arccos\left(\frac{I_{coupled}}{I_{decoupled}}\right)}{\pi 2\tau} \cdot c \quad (2.18)$$

with c as a correction factor due to different relaxation contributions in the coupled and decoupled experiment [155].

2.4 Secondary Chemical Shifts

The chemical shift of NMR active nuclei depends on their direct and therefore unique chemical environment. More precisely, the chemical shifts are a result of the shielding properties from surrounding electrons (intra-atomic) and from magnetic and electrostatic interactions with other nuclei (inter-atomic) [156]. The observation of an empirical correlation between the chemical shifts and the peptide backbone conformation [157,158] caused theoretical research of the factors determining the chemical shifts [159]. The theoretical analysis of the structural and environmental dependencies of ${}^1\text{H}$, ${}^{13}\text{C}$ and ${}^{15}\text{N}$, chemical shifts in peptides revealed that:

- ${}^1\text{H}^\alpha$ chemical shifts are influenced by the local backbone conformation and ring currents of nearby aromatic groups [160].
- the amide proton (H^N) chemical shifts are dominated by hydrogen bonding [159], but in addition are influenced by the backbone conformation [160].
- for nitrogen, dihedral and side chain angles, hydrogen bonding, next neighbors and electrostatic interactions contribute to backbone-nitrogen chemical shifts [160].
- the ${}^{13}\text{C}^\alpha$ and ${}^{13}\text{C}^\beta$ are sensitive to the amino acid side chain and are strongly influenced by the backbone dihedral angles [156].

Especially the ${}^{13}\text{C}^\alpha$ and ${}^{13}\text{C}^\beta$ shifts have become a very powerful addition to structure determination. First of all, the high chemical shift dispersion (due to unique side chains) reduces signal overlap, making them usable even for larger proteins. For example, the ${}^{13}\text{C}^\alpha$ of Glycine resonates at around 45 ppm, whereas ${}^{13}\text{C}^\alpha$ of Proline resonate at around 63 ppm. Secondly, the significant influence of the backbone dihedral angles clearly can indicate alpha-helical

or beta-sheet arrangements. Excluding glycines, $^{13}\text{C}^\alpha$ chemical shifts are shifted downfield by 2.2 (Ala) to 4.5 ppm (Thr) in helices, and upfield by -0.4 (Thr) to -1.8 (Arg) ppm in β -sheets, compared to the random-coil chemical shift [160].

These values show for C^α a deshielding character in α -helices and stronger shielding in β -sheets. Compared to C^β , C^α show a more unique separation between α -helix and β -sheet [158], therefore making especially this nuclei more reliable when looking at residual secondary structure elements in intrinsically disordered proteins. The secondary chemical shifts are calculated as follows:

$$\Delta\delta = C_{obs}^\alpha - C_{RC}^\alpha. \quad (2.19)$$

With C_{RC}^α as the random coil value and C_{obs}^α as the observed value. Currently used random coil values taking into account chemical shift changes from neighboring residues [161] and the pH [162]. Notably, prolines induce an upfield shift of about -2 ppm for the preceding residue [163].

2.5 Paramagnetic Relaxation Enhancement

Scalar couplings and chemical shifts are powerful tools to investigate local structural changes. However, NMR offers additional possibilities for elucidating global structural parameters, for example by using diffusion experiments to get the hydrodynamic radius [164, 165] or paramagnetic relaxation enhancement (PRE) experiments. For PRE experiments, metalloproteins [166], lanthanides [167] or stable radicals [168] have been used, which possess unpaired electrons. As mentioned, an electron has associated with it an orbital angular momentum and a spin angular momentum, which are usually canceled out by the second electron in the orbital. However, for an unpaired electron this is not the case and since an angular momentum is always associated with a magnetic moment, it is:

$$\mu_e = g\mu_B \frac{\mathbf{J}}{\hbar}, \quad (2.20)$$

with μ_e as the electronic magnetic moment, g as the Landé factor, μ_B as the Bohr magneton and \mathbf{J} as the sum of the orbital and spin angular momentum of the electron. The Bohr magneton is the electronic equivalent to the nuclear magneton μ_N , therefore also inversely proportional to the mass of the

particle:

$$\mu_B = \frac{e\hbar}{2m_e}, \quad (2.21)$$

with m_e as the electron mass and e as the electron charge. As for the nuclear spin, a gyromagnetic ratio γ_e for the unpaired electron spin can be defined:

$$\frac{\mu_e}{\mathbf{J}} = \frac{g\mu_B}{\hbar} = \gamma_e \quad (2.22)$$

By using the values $9.274 \cdot 10^{-24}$ [J/T] for μ_B , -2.0023 for the electronic g-factor, $5.050 \cdot 10^{-27}$ [J/T] for μ_N and 5.585 for the nuclear g-factor, it turns out that the electronic magnetic moment is 657 times the nuclear magnetic moment [169]. Due to that fact, interactions between an electron magnetic dipole and nuclear magnetic dipoles can range up to 25 Å, whereas dipolar interactions between two nuclear dipolar moments range usually up to 5–6 Å. Not only the range but also the strength of interaction is increased, evident in enhanced longitudinal and transverse relaxation rates of nearby nuclear spins. The relaxation rates are affected by a r^{-6} distance-dependence. Therefore, enhanced relaxation rates can be converted into distances [168–170]. As it was pointed out by Clore et al. [171], longitudinal relaxation rates are very sensitive to the internal correlation time, that is the flexibility within the molecular frame, to cross relaxation processes and to exchange processes (e.g. with water in the case of amide protons). Transverse relaxation, R_2 , is largely insensitive to these effects, providing therefore a more reliable source. The additional paramagnetic relaxation contribution originating from one unpaired electron with spin $s=\frac{1}{2}$ is defined as [172]:

$$R_2^{para} = \frac{1}{20} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\gamma_H^2 \gamma_e^2 \hbar^2}{r^6} \left[4\tau_c + \frac{3\tau_c}{1 + (\omega_H \tau_c)^2} \right], \quad (2.23)$$

whereas the terms:

$$\left[\frac{1\tau_c}{1 + (\omega_H - \omega_e)^2 \tau_e^2} + \frac{6\tau_c}{1 + (\omega_H + \omega_e)^2 \tau_e^2} + \frac{6\tau_c}{1 + (\omega_e \tau_c)^2} \right], \quad (2.24)$$

are neglected, because with $\omega_H = \gamma_H B_0$ and $\omega_e = \gamma_e B_0 = 657\omega_H$ these terms are extremely small and therefore not influencing the relaxation rate. τ_c is the correlation time. Since the peak intensity is related to the relaxation rate, it is possible to separate the relaxation contribution of the unpaired electron by recording two experiments, one with the attached spin label I_{para} , the second one with either a cleaved or reduced spin label I_{dia} . By taking the intensity

ratio I_{para}/I_{dia} , sources of relaxation present in both experiments will cancel, leaving the paramagnetic relaxation contribution as the only source for the reduced signal intensity.

2.6 Residual Dipolar Couplings

In addition to scalar couplings, another interaction exist which can couple two remote spins. In contrast to the electron-mediated scalar coupling, dipolar couplings are mediated over space, were the dipolar field of one spin directly influences the dipolar field of another spin in close proximity. The equation for a dipole-dipole interaction, derived from classical electrodynamics is [173]:

$$\mathbf{E}_{dd}(\mathbf{r}) = \frac{\mu_0}{4\pi r^3} \left[\boldsymbol{\mu}_1 \boldsymbol{\mu}_2 - \frac{3(\boldsymbol{\mu}_1 \mathbf{r})(\boldsymbol{\mu}_2 \mathbf{r})}{r^2} \right] \quad (2.25)$$

with μ_0 as the permeability (magnetic) constant and r as the distance between dipole 1 and dipole 2. It was pointed out previously that quantum mechanically the magnetic dipoles are represented by spin operators:

$$\boldsymbol{\mu} = \gamma \hbar \hat{\mathbf{I}}, \quad (2.26)$$

with γ as the previously mentioned gyromagnetic ratio, \hbar as Planck's constant divided by 2π and $\hat{\mathbf{I}}$ as the spin operator. According to the correspondence principle, which states that the behavior of systems described by the theory of quantum mechanics reproduces classical physics in the limit of large quantum numbers [139], we can rewrite eq. 2.25:

$$\hat{\mathcal{H}}_{dd}(\mathbf{r}) = \frac{\mu_0 \hbar^2 \gamma_1 \gamma_2}{4\pi r^3} \left[\hat{\mathbf{I}} \cdot \hat{\mathbf{S}} - \frac{3(\hat{\mathbf{I}} \cdot \mathbf{r})(\hat{\mathbf{S}} \cdot \mathbf{r})}{r^2} \right] \quad (2.27)$$

This is the interaction hamiltonian for dipolar couplings between two spins $\hat{\mathbf{I}}$ and $\hat{\mathbf{S}}$. Due to the spherical symmetry of the dipolar field, the spherical coordinate system is used to describe the dipolar interaction in the homonuclear and heteronuclear case [174]. After considerable rearrangements of terms (with discarding those, who have insignificant contribution to the dipolar coupling), it can be shown, that the dipolar coupling interaction is represented by [174]:

$$\hat{\mathcal{H}}_{dd}(\mathbf{r}) = \frac{\mu_0 \hbar^2 \gamma_1 \gamma_2}{4\pi r^3} (3 \cos^2 \theta - 1) \left[\hat{I}_z \hat{S}_z - \frac{1}{2} (\hat{I}_x \hat{S}_x + \hat{I}_y \hat{S}_y) \right] \quad (2.28)$$

in spherical coordinates with θ as the angle between the internuclear vector connecting the two spins and the external magnetic field. In the homonuclear case, the longitudinal as well as the transverse component of spin $\hat{\mathbf{I}}$ affects spin $\hat{\mathbf{S}}$. For heteronuclear residual dipolar couplings, like one-bond $^1H - ^{15}N$ RDCs, only the longitudinal \hat{I}_z and \hat{S}_z magnetization contributes significantly to the dipolar interaction. In the heteronuclear case, eq. 2.28 changes to:

$$\mathcal{H}_{hetero,dd}(\mathbf{r}) = \frac{\mu_0 \hbar^2 \gamma_1 \gamma_2}{4\pi r^3} (3 \cos^2 \theta - 1) \hat{I}_z \hat{S}_z. \quad (2.29)$$

Dipolar couplings are a rich source of structural information, because they contain information on the orientation of internuclear vectors relative to the magnetic field, independent of its location within the protein. Besides probing local geometry, dipolar couplings also have a global ordering character relative to a common frame [175]. However, in isotropic solution, rotational Brownian motion causes the internuclear vector to sample all possible internuclear angles, effectively averaging out the term $(3 \cos^2 \theta - 1)$. As a result, solution NMR spectra do not contain the valuable orientational information anymore.

In 1982 it was shown, that residual dipolar couplings can be re-introduced in liquid samples by a weakly alignment of the molecule itself in the \mathbf{B} -field [176]. Nowadays, usually external alignment media are used, like bicelles, phages, purple membrane fragments, polyethylene glycol compounds or polyacrylamide gels which also allows a scaling of the dipolar splitting to around 10^{-3} compared to complete dipolar coupling [177]. The dipolar splitting, the measured dipolar coupling, is defined as follows:

$$D_{IS}(\mathbf{r}) = \frac{\mu_0 \hbar^2 \gamma_1 \gamma_2}{4\pi r^3} \langle (3 \cos^2 \theta - 1) \rangle, \quad (2.30)$$

whereas $\langle (3 \cos^2 \theta - 1) \rangle$ represents the restriction to only the allowed angles (subset of all angles).

The dipolar coupling D adds to or subtracts from the J -coupling so that one measures $J \pm D$ instead of J . An experiment to measure RDC's is the inphase-antiphase HSQC (IPAP-HSQC) [178].

2.7 High-Resolution Magic Angle Spinning

As it was pointed out in the previous sections, NMR is a powerful tool in gaining information about biomolecules with atomic resolution. However,

proteins larger than 100 kDa are a challenge for traditional multidimensional backbone experiments. An increasing molecular weight is accompanied with:

- an increasing number of resonances, leading to crowded/overlapped spectra, difficult/impossible to analyze
- broad peaks with a low signal-to-noise ratio due fast transverse relaxation
- ineffective averaging of anisotropic spin-spin interactions, again leading to broad signals

Crowded spectra can be partly avoided by increasing the dimensionality of the experiments to currently up to 7 dimensions [179, 180]. Unfavorable relaxation properties can be reduced by using deuterated proteins [141] and applying transverse relaxation optimized (TROSY) [181] pulse sequences. Another approach effectively reducing the line width of signals is the magic angle spinning (MAS), heavily used in solid state NMR. Andrew et al [182] and Lowe [183] first showed that rapid spinning of a solid sample at a certain angle relative to the applied magnetic field reduces the NMR peak widths dramatically. The effect can be explained by looking at the formula for the dipolar coupling interaction [174]:

$$\hat{\mathcal{H}}_{dd} = \frac{\mu_0 \hbar^2 \gamma_I \gamma_S}{4\pi r^3} (3 \cos^2 \theta - 1) \left[\hat{I}_z \hat{S}_z - \frac{1}{2} (\hat{I}_x \hat{S}_x + \hat{I}_y \hat{S}_y) \right], \quad (2.31)$$

with γ_I and γ_S as the gyromagnetic ratio for spin $\hat{\mathbf{I}}$ and $\hat{\mathbf{S}}$, θ as the angle between the internuclear vector connecting these two spins and the external magnetic field, with μ_0 as permeability (magnetic) constant and r as the distance between spin $\hat{\mathbf{I}}$ and $\hat{\mathbf{S}}$.

As it can be seen, the dipolar coupling interaction has an angular dependence. More precisely, the dipolar coupling will disappear if the term $(3 \cos^2 \theta - 1)$ is zero. Exactly this happens when internuclear vector θ between the dipoles and the external magnetic field takes an angle of 54,7 degree. Since rotating the sample at this angle produces such a spectacular effect, it is called the magic angle. The effect is based on the fact that no dipolar spin-spin coupling additional to a scalar coupling can occur, when the interacting dipolar field lacks the longitudinal component, since only longitudinal components alter the effective magnetic field felt by a spin and therefore the resonance frequency (Figure 2.2). The MAS is also used

for samples which are not completely solid but retain some degrees molecular motion, such as large protein aggregates [184], tissues [185], or cellular environments [186] and is then called high resolution magic angle spinning (HRMAS). HRMAS can not only average out dipolar couplings but also

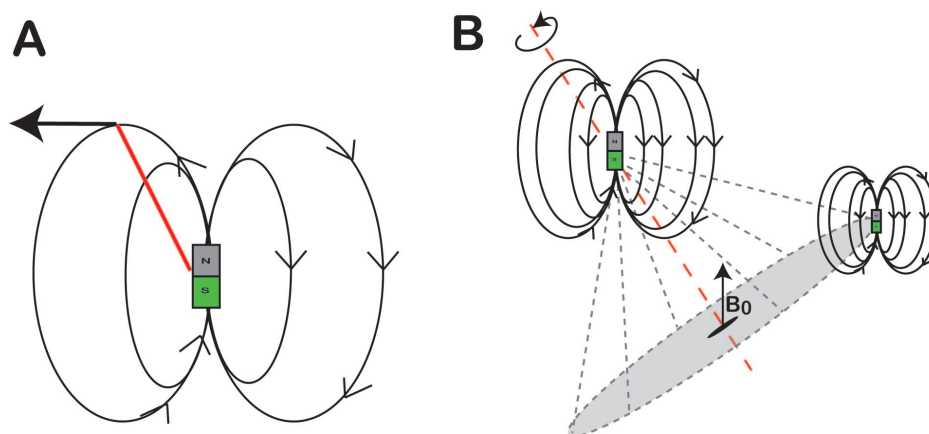


Figure 2.2: Averaging of dipolar couplings with MAS A) Conceptually, the magnetic field lines are closed and range from parallel to anti-parallel to the external magnetic field. That means, that at a specific angle, the longitudinal component of the magnetic field lines must be zero. This is the case at an angle of 54.7 degrees, where the dipolar field has no longitudinal component [187]. Therefore, dipolar coupling or CSA interactions can not influence the energy (chemical shift) of neighboring spins anymore. B) Spinning the sample at the magic angle results in a time-averaged angle of 54.7 degrees between the internuclear vector of the spins and the static magnetic field.

chemical shift anisotropies (Figure 2.3). The magnetic field experienced by the nuclear spin depends on the shielding properties of the electron spins. The surrounding electron cloud rarely possess a spherical symmetry, except for s-orbital electrons. Instead, the spatial anisotropic electron density, and therefore the spatial anisotropic chemical shift shielding, can be represented mathematically as a tensor and geometrically as an ellipsoid, elongated along the bond. That in turn means, that depending on the orientation of the bond with respect to the external magnetic field, the chemical shift (shielding) and therefore the resonance frequency varies, giving rise to broad signals. It is possible to separate the isotropic and anisotropic component of the chemical shift according to [188]:

$$\hat{\mathcal{H}}_{CS} = \delta_{iso} B_0 \hat{\mathbf{I}}_z + \delta_{CSA} B_0 \hat{\mathbf{I}}_z (3 \cos^2 \theta - 1) \quad (2.32)$$

With δ_{iso} as the isotropic component of the chemical shift, θ as the angle between the longest axis of the shielding ellipsoid and the external magnetic field and δ_{CSA} defined as:

$$\delta_{CSA} = \delta_{33} - \delta_{iso}, \quad (2.33)$$

with δ_{33} as the longest principal axis of the shielding tensor with the strongest shielding effect. In solution state NMR, fast molecular tumbling leads to

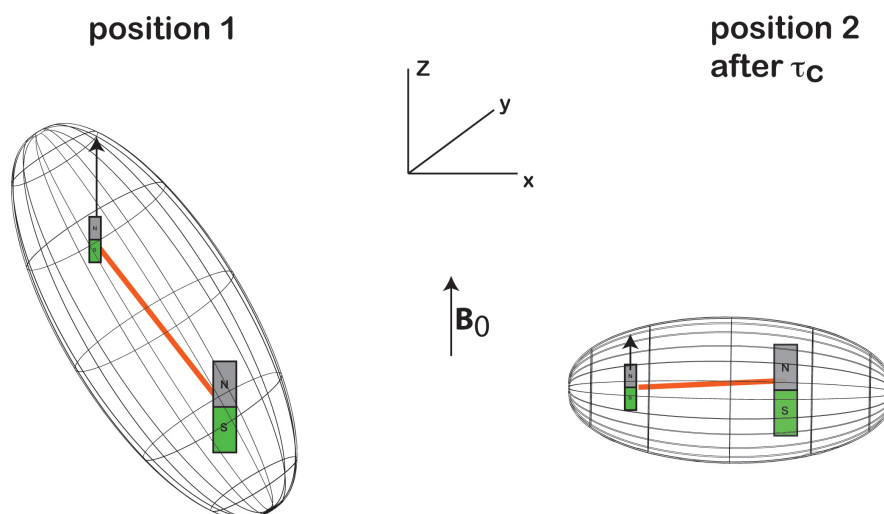


Figure 2.3: Chemical shift anisotropy The chemical shielding, represented as an ellipsoid, defines the chemical shift of the nuclear spins. The strength of the shielding is represented by the length of the arrows (pointing along the B_0 -field). The anisotropic chemical shielding tensor has an angular dependence, so has the chemical shift. Fast molecular tumbling (small τ_c) or spinning at the magic angle averages the anisotropic component to zero.

rapid sampling of all orientations, naturally averaging the anisotropic component to zero, resulting in narrow signals. However, this is not the case in gels/solids with slow or no molecular tumbling. Therefore, spinning at the magic angle needs to be done to average the anisotropic component of the chemical shift to zero. When working with large molecular aggregates like fibrils, another phenomenon contributes to the NMR linewidth in HRMAS samples, namely magnetic susceptibility differences due to sample heterogeneity. Investigations of fibrillar states of proteins inevitably results in heterogeneous samples. In fact, when working with tau-fibrils, the sample can roughly be divided into three phases. Phase 1 consists of areas free of protein

only filled with solvent. Phase 2 consists of the fuzzy coat, accounting for an area partly filled with protein and partly filled with the solvent. Phase 3 consists of areas which are more or less completely filled with the protein aggregate, accounting for the rigid fibril core. Every phase exhibit a different spin density and is therefore different magnetizable, which is expressed via the magnetic susceptibility. The magnetizability, \mathbf{M} , a substance can acquire, is connected to the magnetic field strength, \mathbf{H} , via the dimensionless magnetic susceptibility χ :

$$\mathbf{M} = \chi\mathbf{H}. \quad (2.34)$$

The connection of the magnetic field strength \mathbf{H} and magnetic field \mathbf{B} is given by [189]:

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) = \mu_0(1 + \chi)\mathbf{H}, \quad (2.35)$$

with μ_0 as the previously defined permeability (magnetic) constant (Note that the relative permeability is $\mu_r = 1 + \chi$ and that $\mathbf{B} = \mu_0\mu_r\mathbf{H} = \mu\mathbf{H}$ with μ as the permeability). As it can be seen, the magnetic field is determined by the magnetic susceptibility. Therefore, in a sample with different magnetic susceptibilities the magnetic field within the sample will vary. Especially phase 2 will be subject of a magnetic field gradient, giving rise to a distribution of Larmor frequencies. It was pointed out by Lippens et al [187] that line broadening due to differences in magnetic susceptibility can be treated as a dipolar interaction and that the $(3 \cos^2 \theta - 1)$ dependence vanishes when rotated at the magic angle [187, 190]. Therefore, MAS is also capable of removing these differences to produce narrow lines in HRMAS spectra.

Chapter 3

Materials and Methods

3.1 Biochemical Methods

Expression, purification, mutation, labeling and aggregation of htau40 was conducted in the lab of Eckard Mandelkow. Therefore, only a brief description is given here.

3.1.1 Production of htau40

Human tau constructs were expressed in the vector pNG2 (a derivative of pET-3a, Merck-Novagen, Darmstadt) in *Escherichia coli* strain BL21(DE3) as described [191]. The expressed proteins were purified from bacterial extracts by making use of the heat stability of the protein and by FPLC SP-Sepharose chromatography (Amersham Biosciences). To label the tau proteins with ^{15}N and ^{13}C stable isotopes, the *E. coli* culture expressing htau40 was grown in a M9 minimal medium with $^{15}\text{NH}_4\text{Cl}$ (1 g liter $^{-1}$) and ^{13}C glucose (4 g liter $^{-1}$). Protein samples uniformly enriched in ^{15}N were prepared by growing *E. coli* bacteria in minimal medium containing 1 g liter $^{-1}$ $^{15}\text{NH}_4\text{Cl}$. The cell pellets were resuspended in boiling extraction buffer (50 mM MES, 500 mM NaCl, 1 mM MgCl_2 , 1 mM EGTA, 5 mM dithiothreitol, pH 6.8) complemented with a protease inhibitor mixture. The cells were disrupted with a French pressure cell and subsequently boiled for 20 min. The soluble extract was isolated by centrifugation, the supernatant was dialyzed against two changes of cation exchange chromatography buffer A (20 mM MES, 50 mM NaCl, 1 mM EGTA, 1 mM MgCl_2 , 2 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, pH 6.8) and loaded on an FPLC SP-Sepharose column. The

proteins were eluted by a linear gradient of cation exchange chromatography buffer B (20 mM MES, 1 M NaCl, 1 mM EGTA, 1 mM MgCl₂, 2 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, pH 6.8). NMR samples contained 0.5-1.3 mM ¹⁵N- or ¹⁵N/¹³C-labeled protein in 95% H₂O, 5% D₂O, 50 mM phosphate buffer, pH 6.8, with 1 mM dithiothreitol. Mutations in htau40 were created by site-directed mutagenesis using the Quick-change kit (Stratagene, Amsterdam, The Netherlands). The introduced modifications were verified by DNA sequencing.

3.1.2 Site-specific labeling of htau40 with MTSL

To destroy dimers and get free SH groups, the ¹⁵N labelled htau40 protein present in PBS buffer, usually at concentration of 100-150 μM, was incubated for 10 min at 37 °C in the presence of 4M GdnHCl and 5 mM DTT in PBS, pH 7.4. Immediately before spin labeling, dithiothreitol and GdnHCl were removed from the buffer using size exclusion chromatography (PD-10 columns, GE Healthcare, Freiburg, Germany). The buffer used for equilibration and spin labeling was PBS, pH 7.4. Sulfhydryl groups of the purified cysteine tau mutants were reacted with 10-15 fold excess of the spin label MTSL (MTSL = (1-Oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl) Methanethio-sulfonate) (Toronto Research Chemicals, Toronto, Ontario, Canada), at room temperature for 2.5 hours. Unreacted spin label was removed and tau protein was rebuffed into 50 mM Na-Phosphate, pH 6.8 using PD-10 columns. Spin-labeled tau derivatives were concentrated using Amicon Ultracel (Millipore, Cork, Ireland, molecular weight cut off: 5.000).

3.1.3 Aggregation of tau

PHFs of wild-type and mutant htau40 were formed by mixing ¹³C/¹⁵N or ¹⁵N-labeled protein (~1.5 mM) with heparin 5000 (heparin:tau 1:4) and incubated at 37 °C for 4 days. The reaction was then pelleted at 160000g for 40 min. The protein supernatant was complemented again with heparin 5000 and incubated for another 4 days. The reaction was then pelleted again and the pellets were combined.

Before the NMR measurements, NaN₃ was removed and the sample was resuspended in 50 mM sodium phosphate buffer (pH 6.8). The fibrils were pelleted three to four times by ultracentrifugation (40.000 rpm) for 30 min

at 4 °C to remove any residual monomeric protein. Finally, the pellet was transferred into a 4 mm rotor with 10 μ l D₂O and buffer.

3.2 NMR Spectroscopic Methods

3.2.1 2D experiments

All 2D ¹H-¹⁵N-HSQC experiments were recorded on a Bruker spectrometer (Bruker Karlsruhe, Germany) with either 800 or 900 MHz proton frequency. The spectrometers are equipped with a TCI cryoprobe (Z-gradient). The HSQCs for aggregated tau (PHF-tau) were recorded on the 900 MHz spectrometer, equipped with a 4mm TXI high-resolution MAS (HRMAS) probehead (Z-gradient). The HSQC pulse sequence contained a 3-9-19 watergate block for water suppression [192]. The experiments were recorded with 600 increments in the indirect dimension and 1k points (cryoprobe) or 2k points (HRMAS) in the direct dimension, zero-filled to 4k and 2k points in the direct and indirect dimension, respectively and processed with a sine-squared bell window function shifted either by 90 degree (SSB=2 in TopSpin), or by 60 degree (SSB=3 in TopSpin). The nitrogen carrier frequency was set to 118 ppm and the spectral width was 25 ppm. For protons, the carrier was set to the water resonance frequency. The data were processed using TopSpin 2.0 (Bruker).

3.2.2 Chemical shift difference

To compare the ¹H-¹⁵N-HSQC chemical shifts from different samples, the differences in the resonance position was calculated using the formula:

$$\Delta\delta = \sqrt{\left(\frac{\Delta\delta_N}{5}\right)^2 + (\Delta\delta_H)^2} \quad (3.1)$$

3.2.3 Secondary chemical shifts

The residual secondary chemical shifts were calculated as the difference between the observed *C*^α chemical shifts and the empirical derived random coil shifts [161]. Random coil values preceding prolines were corrected [163], as well as random coil values for histidines, glutamates, and aspartates [162], as

the chemical shifts of these residues are particularly sensitive to pH. To estimate the secondary structure propensity in contiguous segments of tau, the observed C^α shift was assigned a value in percentage, by using the random coil value and rigid secondary structure-value of this particular residue [162]. The percentage-values were summed and divided by the number of residues in this segment.

3.2.4 Structure calculation

Structure calculations were performed using XPLOR-NIH, version 2.9.7 [193]. An all-atom representation of httau40 was used. Structural energy terms from steric repulsion, bond length, bond angles, dihedral angles, and favored regions of the Ramachandran map were employed. For restraining a single molecule simultaneously by all PRE distance restraints, torsion angle dynamics were started at 3000 K with the temperature reduced to 20 K, followed by a short energy minimization. 50 structures were calculated starting from a random coil. The seven lowest-energy structures that satisfied the 2288 distance restraints with no violations greater than 1 Å were used for calculation of the average contact map. Single molecule calculations were followed by ensemble calculations, in which distance restraints do not have to be fulfilled by a single molecule, but collectively by the ensemble of molecules. Ensemble calculations were started from the lowest-energy structure obtained in the single molecule calculation (see above) and performed in two rounds. Initially, distance restraints were enforced onto an ensemble of 30 molecules [194]. Torsion angle dynamics were used with the temperature reduced from 10000 K to 5000 K. The lowest energy structure obtained from this first round of ensemble averaging was subjected to another round of structure calculation using an ensemble size of 5. Torsion angle dynamics was used and the temperature was reduced from 3000 K to 1000 K. A total of 100 structures were calculated. The average contact map obtained from the seven lowest-energy structures of the ensemble was very similar to the one obtained from single molecule calculations.

3.2.5 $^3J_{H^N H^\alpha}$ couplings

All $^3J_{H^N H^\alpha}$ scalar couplings were measured using an intensity modulated HSQC [153] on a Bruker 900 Avance spectrometer. Coupling values were

calculated from the intensities of the coupled and decoupled spectrum (see chapter 2):

$${}^3J_{HNH\alpha} = \frac{\arccos\left(\frac{I_{coupled}}{I_{decoupled}}\right)}{\pi 2\tau} \cdot c \quad (3.2)$$

The experiment was recorded in an interleaved way. For wildtype htau40, E-mutant and PHF-tau, 2 times 600 increments in the indirect dimension were recorded with 2k points in the direct dimension. The spectral width was 10 ppm and 23.5 ppm for proton and nitrogen, respectively. For nitrogen, the carrier was set to 118 ppm. All experiments were recorded at 278K. Each transient had 32 scans (wildtype htau40 and E-mutant) or 64 scans (PHF-tau) and always 16 dummy scans. The relaxation delay was set to 1.2 seconds. 2τ was set to 18 ms. The H₂O signal was suppressed by using the WET-sequence [195]. Secondary ${}^3J_{HNH\alpha}$ scalar couplings were calculated as the difference between experimental ${}^3J_{HNH\alpha}$ scalar couplings and random coil values [196].

3.2.6 PRE experiments

PRE broadening was investigated using ¹⁵N-labelled htau40 (E-mutant) at a concentration of 15 and 50 μM (15 μM) in 50 mM phosphate buffer at pH 6.8. For PHF-tau, the sample contained 30-70 mg protein. PRE effects were measured from the peak intensity ratios between two 2D ¹⁵N-¹H HSQC NMR spectra acquired in the presence of the nitroxide radical and after addition of 4 mM DTT (heated to 45 °C for 30 min before measurement) to the same sample. Addition of DTT will cleave the MTSL tag from the cysteine residue, resulting in the diamagnetic state of the protein. For the intensity profile plot, intensities were averaged over a three-residue window.

3.2.7 Diffusion experiments

For an accurate estimation of the diffusion coefficient the temperature on the respective spectrometer was calibrated by either using the Bruker au-program 'caltemp' (with 4% MeOH in MeOD), or by using a calibrated thermometer within a 5 mm NMR-tube filled with water. For the gradient strength the residual ¹H signal in a sample containing 99.9% D₂O was used. The value of $1.902 \pm 0.002 \frac{m^2}{s}$ was taken for the self-diffusion coefficient of

HDO at 25 °C [197]. By using the decay of the HDO signal, the gradient strength was calculated according to Chou et al [198].

For determining the hydrodynamic radius of monomeric wildtype htau40, the protein was dissolved in 99.9 % D2O, 50 mM phosphate buffer (pH 6.8) at 278K. The samples contained dioxane (concentration 2%) as an internal radius standard and viscosity probe [199]. 1D ^1H spectra were collected as a pseudo-2D experiment employing the standard Bruker pulse sequence ledbpgp2s [200]. The gradient strength was linearly increased from 2% to 95% of the maximum gradient strength in 16 steps, with 100% gradient strength corresponding to 56.9 G/cm. For each ^1H spectrum 128 scans and 16k complex data points with a spectral width of 7.200 Hz were acquired. Signal intensities corresponding to the aliphatic region of the ^1H spectra (3.3-0.5 ppm) were read out with the TopSpin T1/T2 Relaxation module (Bruker Instruments). The diffusion data (signal intensity versus gradient strength) were fitted to exponential functions using Igor Pro 5.01 (WaveMetrics). From the apparent diffusion coefficients of htau40 and dioxane and the known hydrodynamic radius of dioxane (2.12 Å) [199], hydrodynamic radii of monomeric htau40 were calculated [201]:

$$R_h^{prot} = \frac{D^{ref}}{D^{prot}} R_h^{ref} \quad (3.3)$$

For PHF-tau, eight one-dimensional ^1H spectra were collected as a function of gradient amplitude employing the BPP-LED pulse sequence [200]. The gradient strength was increased from 1.0 to 47.5 Gauss/cm in a linear manner. Each ^1H spectrum comprised 128 scans plus four steady-state scans. 4096 complex points were acquired with a spectral width of 10800 Hz. Signal intensities of the aliphatic region of ^1H spectra (3.3-0.5 ppm) were read out with the TopSpin T1/T2 relaxation module and fitted to an exponential function. The size of the fibrils was estimated from electron micrographs as 1000 nm in length and 15 nm in diameter. Therefore, an additional Perrin factor was introduced in the Stokes-Einstein equation, accounting for the highly anisotropic PHFs [202]. Based on these calculations the molecular weight was estimated to be larger than 1 MDa.

For the E-mutant a new approach was employed by using the new pulse sequence PGSTE-WATERGATE [203] and by using the diffusion coefficient of another protein (lysozyme) instead of a small molecule. In addition, this pulse sequence allowed the diffusion measurements in H_2O . Three ex-

periments for wildtype htau40, E-mutant and lysozyme were recorded on a Bruker 600 Avance spectrometer equipped with a cryogenic probe for each sample, containing a concentration of 100 μM . By using the same spectrometer, we assure that the same temperature and the same gradient strength was present for all samples. By using the same buffer and the same concentration, we assumed similar viscosities for all samples. The gradient strength was linearly incremented in 16 steps from 25% to 95% of the maximum gradient strength, corresponding to 5% to 75% of the maximum gradient strength felt by the spins (due to the double utilization of the gradients in the PGSTE-WATERGATE). 32 scans for every increment were recorded. The intensities were read out with the TopSpin T1/T2 relaxation module and fitted in Igor Pro. The measured diffusion coefficient and the well known hydrodynamic radius of lysozyme [199, 204, 205] was used in equation (3.3) to calculate the hydrodynamic radii of tau.

3.2.8 Relaxation experiments

All PHF-relaxation experiments were conducted on a Bruker 900 MHz spectrometer equipped with an HRMAS probehead (Bruker, Karlsruhe). The temperature was set to 278K and the spinning frequency was 6 kHz. The sample was in 50 mM phosphate buffer, pH 6.8 and centrifuged with 40.000 rpm. The pellet was diluted again in buffer and centrifuged. This step was repeated four times to remove any residual monomeric tau. Initially, the sample contained 66 mg of PHF-tau, whereas after the washing steps roughly 30-40 mg remained. The R2 experiment was recorded as an interleaved experiment with relaxation times of 8, 30, 50, 70, 100, 120 and 150 ms. The on-resonance R1 ρ experiment was measured with a field strength of 2.5 kHz and with relaxation delays of 8, 30, 50, 75, 100, 125 and 150 ms in an interleaved fashion. Both experiments contained a heat-compensating element at the beginning of the pulse sequences. For both experiments, the last relaxation delay was set to 8 ms to check the stability of the sample. The ^{15}N carrier frequency was set to 118 ppm and the sweep width was 25 ppm. The number of scans for both experiments was 16 with 128 dummy scans. For each R2 and R1 ρ relaxation delay 440 increments in the indirect dimension were recorded. D1 was set to 1.5 seconds, resulting in an experimental time of 24h per experiment. The data were analyzed with NMRPipe [206] and Sparky (T.D. Goddard and D.G. Kneller, <http://www.cgl.ucsf.edu/home/sparky>). The relaxation

curves were fitted to a single exponential decaying function and the standard deviation was calculated.

3.2.9 3D experiments

The HNN experiment [207] connects three consecutive residues and therefore allows an unambiguous assignment of the backbone resonances. The experiments were recorded on a Bruker spectrometer (Bruker Karlsruhe, Germany) with either 800 (MARK-phosphorylated tau) or 900 (E-mutant) MHz proton frequency, equipped with a TCI cryoprobe (Z-gradient). 60 increments were recorded in both indirect nitrogen dimensions. The spectral width was set to 25.5 ppm for nitrogen and the carrier-frequency was set to 118 ppm. For protons, the spectral width was 10 ppm with the carrier on the water resonance (4.69 ppm). 8 and 12 scans per transient were recorded for the E-mutant and MARK-phosphorylated tau, respectively. The experimental time was 40h for the 8-scan experiment and 60h for the 12-scan experiment. The data were processed with NMRPipe [206].

The HACANNH experiments [208] were recorded on a Bruker spectrometer (Bruker Karlsruhe, Germany) with either 800 (MARK-phosphorylated tau) or 900 (E-mutant) MHz proton frequency. The in-phase (IP) and anti-phase (AP) components were recorded in an interleaved-fashion. The spectrometers were equipped with a TCI cryoprobe (Z-gradient). For the E-mutant (number of scans: 8) and MARK-phosphorylated tau (number of scans: 16) 50 increments were recorded for the indirect ^{15}N -dimension and 40 increments in the indirect ^{13}C -dimension (duration was 24h and 48h for the E-mutant and MARK-phosphorylated tau, respectively).

For PHF-tau, a 3D HNCA experiment was recorded under HRMAS conditions using a watergate for water suppression [211], a constant time period for ^{13}C evolution, DIPSI-2 for proton decoupling and GARP for nitrogen decoupling during acquisition. 2048 x 68 x 88 points (^1H x ^{13}C x ^{15}N) with 50 scans were recorded, resulting in a total experimental time of 4 days and 3 hours. The spinning frequency was set to 6 kHz so that the spinning band is outside of the amide proton frequency region. For the 3D experiment, roughly 30 mg of aggregated tau was filled into a 50 μL HRMAS rotor.

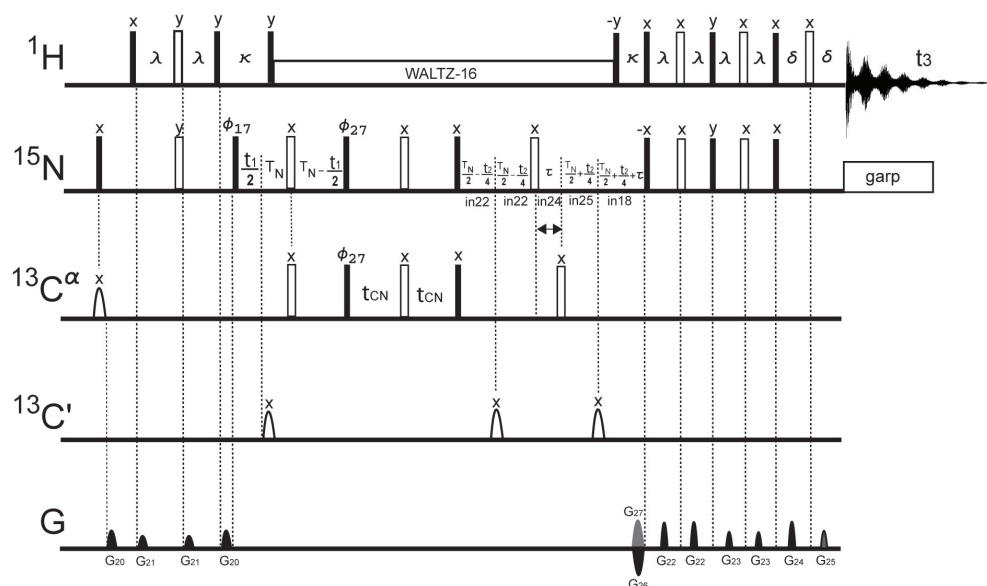


Figure 3.1: The HNN experiment. Filled bars and hollow bars (arcs) indicate 90 degree and 180 degree hard (shaped) pulses, respectively. Protons are decoupled during the t_1 and t_2 evolution periods by using WALTZ-16, whereas the Garp sequence was applied during acquisition for ^{15}N decoupling. The delays are $\lambda = 2.7$ ms, $\kappa = 5.4$ ms and $\delta = 0.25$ ms. T_N was set to 15 ms. τ_{CN} must be optimized and was set to 12 ms. Numbering of phases and gradients are according to the experiment. The phases are $\phi_{17} = x, x, -x, -x$; $\phi_{27} = x, -x, x, -x$ and the receiver phase $\phi_{31} = x, x, -x, -x$. The gradient strength was set to $G_{20} = 51\%$, $G_{21} = 31\%$, $G_{22} = -23\%$, $G_{23} = 17\%$, $G_{24} = 50\%$, $G_{25} = -50\%$, $G_{26} = -50\%$ and $G_{27} = 50\%$ of the maximum gradient strength (56.9 G/cm). G_{26} and G_{27} are the frequency encoding gradients, whereas the combination of G_{24} and G_{25} decode the P- and N-type coherences [209] for t_2 . The gradient duration was set to $G_{20} = 2\text{ms}$, $G_{21} = 2\text{ms}$, $G_{22} = 2\text{ms}$, $G_{23} = 2\text{ms}$, $G_{24} = 0.2\text{ms}$, $G_{25} = 0.075\text{ms}$, $G_{26} = 2.705\text{ms}$ and $G_{27} = 2.705\text{ms}$. Frequency discrimination in t_1 is achieved using States-TPPI phase cycling of ϕ_{17} along with the receiver phase. In addition, the experiment uses the PEP-sensitivity enhancement [210]. For the constant-time evolution, the incrementation must satisfy $\text{in}22 = \text{in}25$ and $\text{in}18 = \text{in}24 + \text{in}25$ (adapted from [207]).

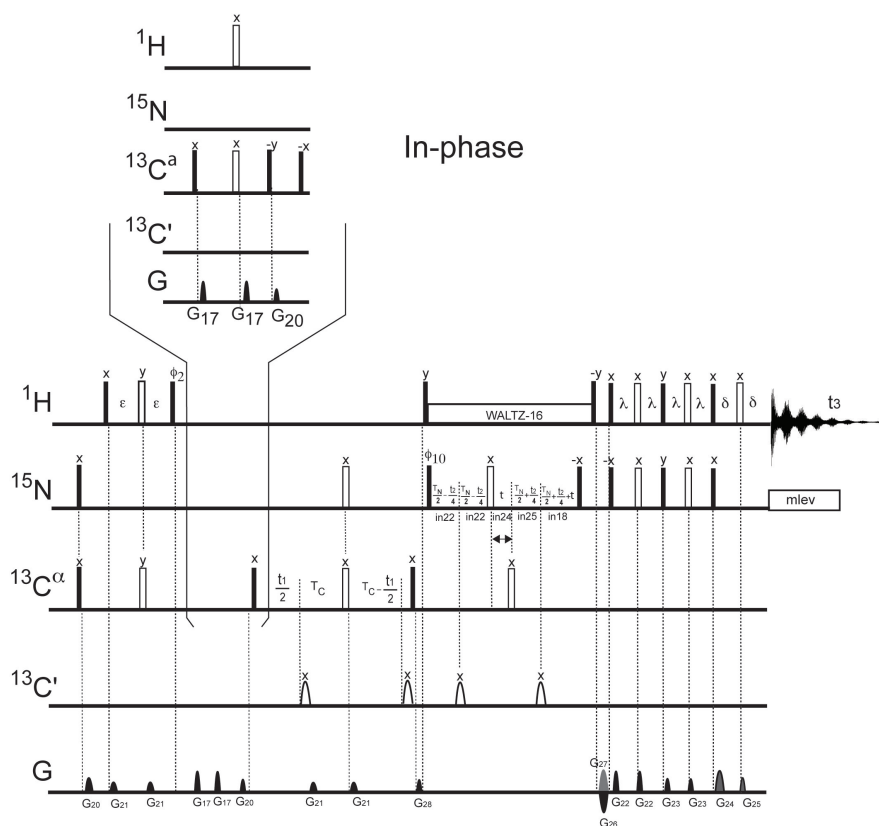


Figure 3.2: The (IPAP)-HACANNH experiment. Filled bars and hollow bars (arcs) indicate 90 degree and 180 degree (shaped) pulses, respectively. Protons are decoupled during the t_2 evolution period by using WALTZ-16, whereas the mlev sequence was applied during acquisition for ^{15}N decoupling. The delays are $\epsilon = 1.5$ ms $\lambda = 2.65$ ms and $\delta = 0.25$ ms. T_C was set to 14ms. τ_N must be optimized and was set to 14ms. Numbering of phases and gradients are according to the experiment. The phases are $\phi_2 = y, -y$; $\phi_{10} = x, x, -x, -x$ and the receiver phase $\phi_{31} = x, -x, -x, x$. The gradient strength was set to $G_{17} = 30\%$, $G_{20} = 61\%$, $G_{21} = 41\%$, $G_{22} = -33\%$, $G_{23} = 27\%$, $G_{24} = 50\%$, $G_{25} = -50\%$, $G_{26} = -50\%$, $G_{27} = 50\%$ and $G_{28} = 43\%$ of the maximum gradient strength (56.9 G/cm). G_{26} and G_{27} are the frequency encoding gradients, whereas the combination of G_{24} and G_{25} decode the P- and N-type coherences [209] for t_2 . The gradient duration was set to $G_{17} = 0.75$ ms, $G_{20} = 1.5$ ms, $G_{21} = 1.1$ ms, $G_{22} = 1.0$ ms, $G_{23} = 1.0$ ms, $G_{24} = 0.2$ ms, $G_{25} = 0.068$ ms, $G_{26} = 2.705$ ms, $G_{27} = 2.705$ ms and $G_{28} = 2.0$ ms. Frequency discrimination in t_1 is achieved using States-TPPI phase cycling of ϕ_2 . In addition, the experiment uses the PEP-sensitivity enhancement [210]. For the ^{15}N constant-time evolution, the incrementation must satisfy $\text{in22} = \text{in25}$ and $\text{in18} = \text{in24} + \text{in25}$ (adapted from [208]).

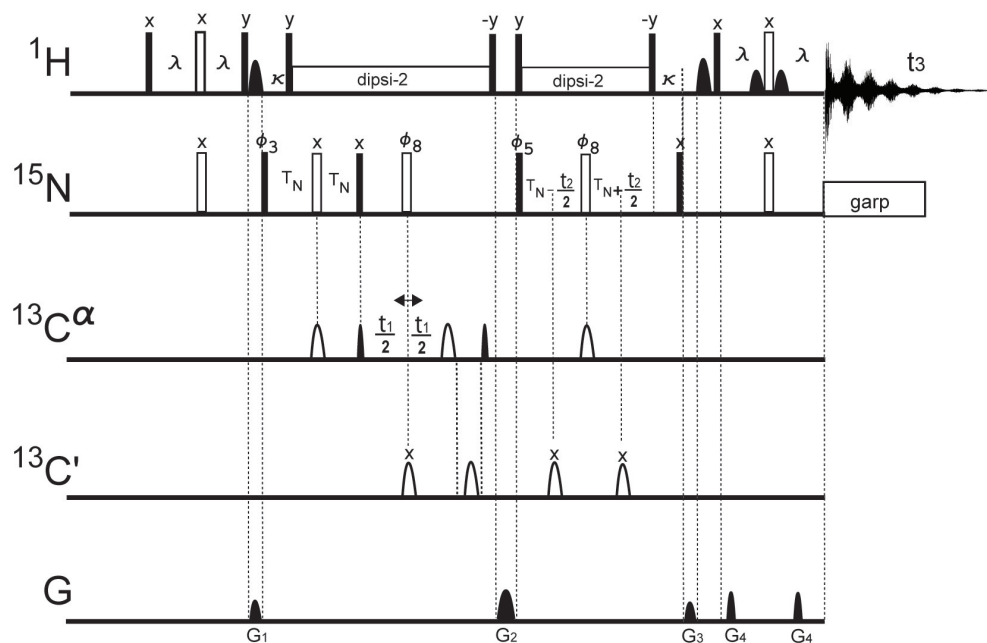


Figure 3.3: The HNCA experiment. Filled bars and hollow bars (arcs) indicate 90 degree and 180 degree (shaped) pulses, respectively. Protons are decoupled during the t_1 and t_2 evolution periods by using DIPSI-2, whereas for ^{15}N decoupling the garp sequence was applied during acquisition. The delays are $\lambda = 2.3\text{ms}$ and $\kappa = 5.5\text{ms}$. T_N was set to 12ms. Numbering of phases and gradients are according to the experiment. The phases are $\phi_3 = (8)x, (8)-x$; $\phi_5 = x, x, -x, -x$; $\phi_8 = (4)x, (4)-x$ and the receiver phase $\phi_{31} = 2(x, -x, -x, x), 2(-x, x, x, -x)$. The gradient strength was set to $G_1 = 50\%$, $G_2 = 40\%$, $G_3 = 60\%$ and $G_4 = 30\%$ of the maximum gradient strength (55 G/cm). The gradient duration was set to 1.0ms. Frequency discrimination in t_1 and t_2 is achieved using States-TPPI. For CO decoupling, gaussian pulse cascades (Q5 and Q3) were used.

Chapter 4

Results and Discussion

4.1 Global Structure of 441-Residue Tau at Single Residue Resolution

4.1.1 Assignment procedure

A prerequisite to every in-depth NMR analysis is the assignment of the resonances to the corresponding amino acid in the protein's sequence. As mentioned, NMR is currently the only biophysical method providing data of intrinsically disordered proteins on a per-residue level. Nevertheless, NMR of intrinsically disordered proteins, especially of tau, is considered to be challenging for several reasons:

- The very large number of amino acids results in crowded spectra with many resonances.
- the lack of an ordered structure results in a low dispersion of signals in the proton dimension and heavy resonance overlap, additionally complicating data analysis.
- The degeneracy of amino acids. Gly (11.1%), Ser (10.2%), Lys (10%) and Pro (9.3%) account for $\approx 40\%$ of all amino acids in full length tau (htau40) with 441 residues.
- The repeat regions R1-R4 and the pseudo-repeat R' exhibit very similar amino acid sequences.

- The proline-rich regions P1 and P2 contain a very large number of prolines, making it difficult to connect adjacent residues in the backbone with standard amide-proton based pulse sequences.

These reasons limited previous NMR studies to fragments of tau [42, 61, 63, 212–214] or to an incomplete assignment [65, 92, 215–217]. Despite the fact that the data recording, analysis and assignment was done by Marco Mukrasch, it is mentioned here, because it firstly exemplifies a successful strategy to overcome the obstacles [218]. Secondly, a similar approach was used for the assignment of the tau-mutants/tau-aggregates in the subsequent sections. To assign full-length htau40 three overlapping fragments were produced (Figure 4.1D):

1. A 185-residue fragment comprising the N-terminal half up to the repeat region, but excluding the two inserts that are affected by alternative splicing named K25).
2. A 198-residue fragment containing the repeat region and the two proline-rich flanking regions (named K32).
3. A 168-residue fragment covering most of the C-terminal half except for the second repeat R2 (named K10).

The two 29-residue inserts in the N-terminal half were only present in htau40 (Figure 1.4). The use of 3D (HA)CANNH [208] and 3D HNN [207] experiments in conjunction with high-field spectrometers allowed the assignment of more than 98% of non-proline backbone resonances for the full length htau40 protein (Figure 4.1E). htau40 is currently by far the largest assigned IDP. Superposition of 2D HSQCs of the three fragments with that of htau40 showed that many resonances observed in the three fragments were found at identical positions as in htau40 (Figure 4.1D).

With the assignment further investigations are possible, like residue-specific flexibility studies, residual secondary structure propensities, hotspots of MT-binding and the global folding of tau.

4.1.2 PRE experiments reveal a network of long-range interactions

To study the global folding of htau40, we employed paramagnetic relaxation enhancement (PRE) experiments [169]. The primary sequence of htau40 con-

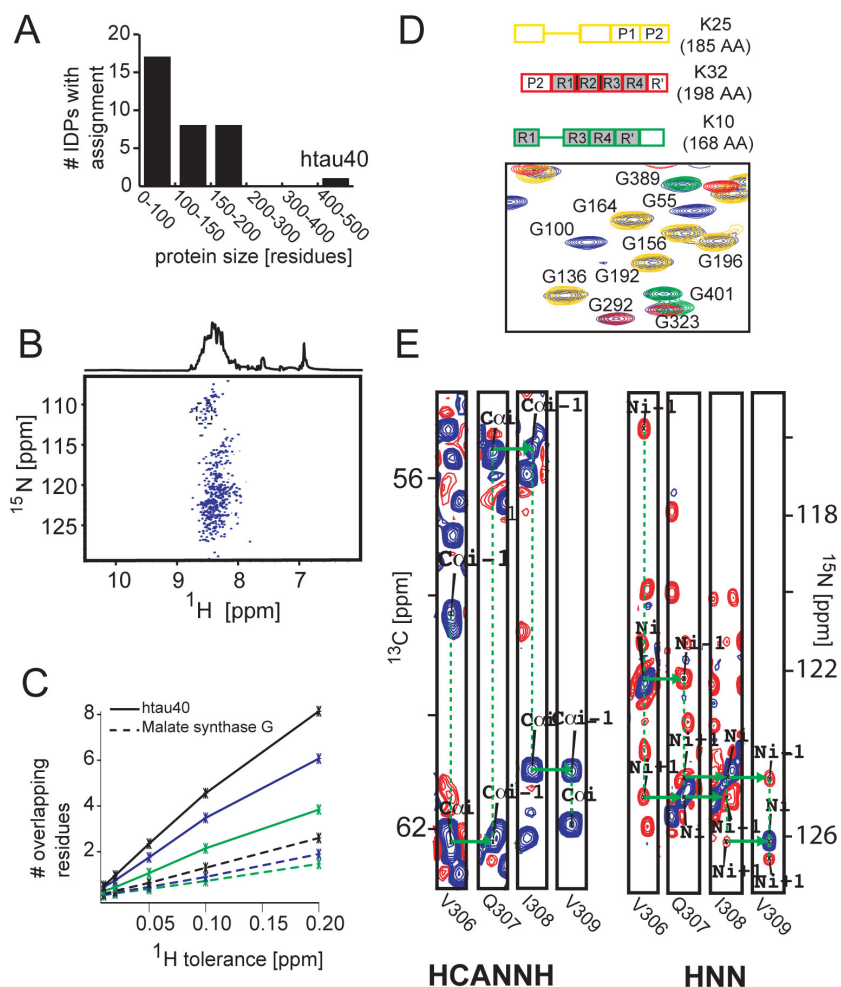


Figure 4.1: Assignment of 441-residue tau A) Number of intrinsically disordered proteins with NMR backbone assignment. B) ^1H - ^{15}N HSQC spectra of httau40 with ^1H projection on top. C) Comparison of spectral overlap observed in HSQC spectra of httau40 (solid line) and 731-residue malate synthase G (dashed line). Black, blue, and green indicate ^{15}N chemical shift tolerances of 0.2, 0.15, and 0.1 ppm, respectively. D) Superposition of a selected portion of the ^1H - ^{15}N HSQCs of the three tau fragments K25 (yellow), K32 (red), and K10 (green), and of 441-residue httau40 (blue). The domain organization of the three tau constructs is indicated. E) Assignment strategy for httau40. 2D strips of high-resolution 3D HCANNH (left) and HNN spectra (right). The connectivity path linking residues V306 to V309 is marked in green (taken from [218]).

tains two cysteines (C291 and C322) that provide natural attachment points

for the nitroxide radical (1-oxy-2,2,5,5-tetramethyl-D-pyrroline-3-methyl) - methanethiosulfonate (MTSL). In addition, five different mutants were constructed and labelled with MTSL which harbor only a single cysteine in the projection domain (A15C or A72C), in the proline-rich region (A239C), or near the C terminus (A384C or A416C). Figure 4.2 shows the PRE profiles (ratio of NMR signal intensities in the paramagnetic and diamagnetic state) of the amide protons of httau40 for six different MTSL-labelled httau40 samples. For a fully extended chain, NMR signal intensities in the paramagnetic and diamagnetic state should be identical for residues that are more than ten to 15 residues away from the position of the spin label. Thus, if httau40 would be a fully extended chain most residues will show a PRE intensity ratio of one. In clear contrast, we observed for many residues far in sequence from the spin-labelled residue PRE intensity ratios below one, indicating transient long-range contacts between the spin-label and distant areas of sequence (Figure 4.2). When the spin-label was attached to position

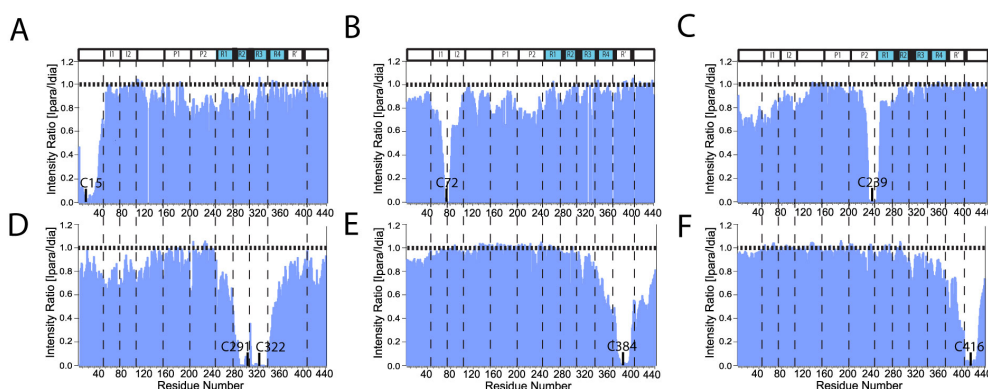


Figure 4.2: PRE of Amide Protons in Spin-Labelled Tau (A-F) PRE profiles of amide protons in spin-labelled httau40. wt httau40 and five single-cysteine mutants (A15C, A72C, A239C, A384C, and A416C) were labelled with MTSL at residue (A) 15, (B) 72, (C) 239, (D) 291 and 322, (E) 384, and (F) 416. Intensity ratios were averaged over a three-residue window. Decreases in peak intensity ratios that occur far from the site of spin-labeling (≥ 10 residues) are indicative of long-range contacts (≤ 25 Å) between the spin-label and distant areas of sequence (taken from [218]).

15, intensity ratios of approximately 0.8 were observed in the proline-rich domain P2 (residues 198-244) as well as residues downstream of P2 (residues 130-200) and residues in the repeat region up to residue 330 (Figure 4.2A). A

transient long-range interaction between the N-terminal region and its central domain is further supported by the PRE profile of C239-MTSL labelled htau40, for which intensity ratios of 0.6-0.8 were observed for residues 1-80 and weaker broadening extended up to residue 150 (Figure 4.2C). On the other hand, signal intensity ratios in case of C291/C322- and C384-MTSL labelled htau40 indicate that the C-terminal domain (residues 360-441) transiently contacts the repeat region and the 40 N-terminal residues (Figure 4.2D and E).

4.1.3 Discussion and Conclusion

The appearance of tau as an elongated molecule by some EM methods [219,220] suggested that the conformation in solution was extended in agreement with the natively unfolded state of tau and the accessibility to multiple kinases throughout the chain. However, evidence for global folding began to emerge from several antibodies that had discontinuous epitopes comprising residues near the N terminus and within the repeat domain (e.g., Alz-50, MC-1 [221]). This evidence was further confirmed by FRET studies showing that tau was able to form a double hairpin, leading to a paperclip structure whereby both N- and C terminus were folded into the vicinity of the repeat domain [49]. This concept is now substantially expanded and refined by the NMR analysis (Figure 4.2). Whereas fluorescence resonance energy transfer combined with electron paramagnetic resonance requires two labels, one label acting as donor and the other one as acceptor, PRE monitored by NMR requires only a single paramagnetic centre such as MTSL attached to a free cysteine. Even more important, whereas in FRET only a single distance can be measured for each donor-acceptor pair, all nuclei in the protein serve as acceptors. Thus, a large number of intra-residual distances (≥ 100) can be measured from a single MTSL-labelled sample. In this study, six uniformly distributed MTSL positions provided a total of 2,288 distance restraints.

To obtain more direct insight into the ensemble of structures populated by htau40 in solution, all NMR signal intensity ratios ≤ 0.9 were converted into distance restraints using the r^{-6} dependence of the PRE effect on the electron-proton distance [169]. In this way we obtained from the six PRE profiles 1,224 intramolecular long-range contacts ranging between 0 Å and 25 Å. In addition, PRE intensity ratios close to one (here ≥ 0.9) indicate that the corresponding amide proton is on average more than 25 Å away from the spin

label, allowing lower distance boundaries of 25 Å for these residues. The to-

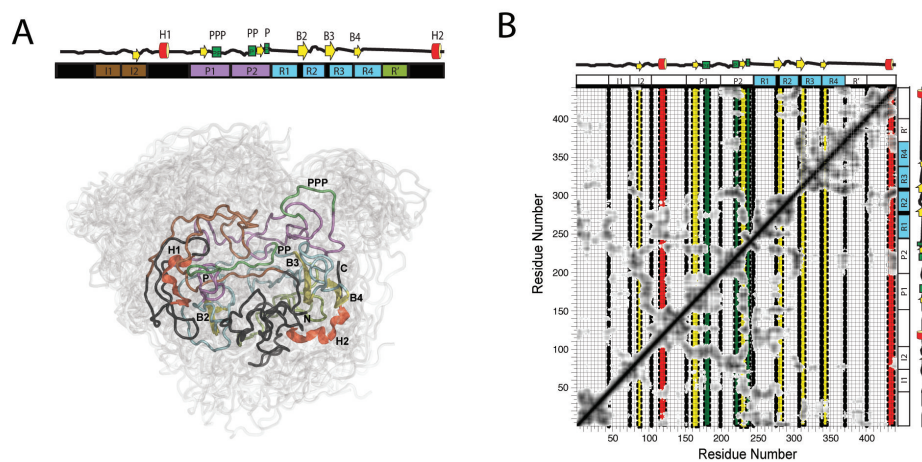


Figure 4.3: Native-State Conformations of 441-Residue Tau in Solution A) Representation of the conformations of htau40 calculated from PRE data. Left panel: Color coding follows the domain organization diagram shown above. Regions of transient α -helical structure (H1[114-123] and H2[428-437]) and β -structure (B2[274-284], B3[305-315] and B4[336-345]) are shown in red and yellow, respectively. Polyproline II stretches (PPP[175-184], PP[216-223], and P[232-239]) are colored green (see next section). In the background, an ensemble of 20 conformations is shown. B) Average contact map for the seven lowest-energy structures obtained from a calculation in which all distance restraints were enforced onto a single molecule. A continuous grey scale from 3 Å (black) to 22 Å (white) is used (taken from [218]).

tal of 2,288 PRE distance restraints was subjected to a structure calculation using simulated annealing [193]. Structure calculations of proteins are challenging when the protein exchanges rapidly between different conformations, such that only a single NMR signal is observed. Rapid exchange between multiple conformations is clearly present in the highly dynamic tau protein and the PRE intensity ratios shown in Figure 4.2 are values averaged over a large ensemble of conformations. Moreover, due to the r^{-6} dependence, conformations with short intramolecular distances contribute more strongly to the PRE broadening than more extended structures. To take into account the dynamic nature of htau40 we performed both single molecule calculations, in which all distance restraints were enforced simultaneously onto a

single molecule, as well as ensemble calculations in which the PRE distance restraints had to be fulfilled not by a single structure but collectively by an ensemble of 30 conformations, respectively [194, 222]. Clearly, even the 30 conformer calculations are a compromise and underestimate the number of conformations htau40 can assume in solution. Nevertheless they better reflect the ensemble nature of the PRE distances, i.e., the observed PRE broadening arises from a very large number of different conformations and each conformation may only fulfill a small subset of distance restraints at any given time. On the other hand, single molecule calculations can allow direct access to the more compact conformations that htau40 could potentially assume in solution. The structure calculations in which all distance restraints were enforced onto a single molecule resulted in an ensemble of compact conformations (Figure 4.3). The shown structures fulfill all 2,288 experimental distance constraints within 0.5 Å. It is readily apparent that many different conformations are in agreement with the experimental PRE distance restraints (structures shown in light grey). The conformation highlighted as ribbon diagram in Figure 4.3 has a radius of gyration R_g of 48 Å and is therefore at the lower end of the distribution of R_g values obtained from SAXS [223]. Subsequently, this conformation was used in the ensemble calculations. In agreement with the fact that a single compact structure could fulfill all distance restraints, the same was true for structures calculated by ensemble averaging. However, as distance restraints only had to be fulfilled by an ensemble of structures, more expanded conformations were obtained. The average radius of gyration of the ensemble of calculated structures was approximately 65 Å, in agreement with the average value obtained for htau40 from SAXS [223]. htau40 is a highly dynamic protein and many conformations are in agreement with the experimental PRE profiles (Figure 4.3A). To summarize, the PRE data reveal an intricate network of long-range interactions of monomeric tau, in agreement with the paper clip model proposed by FRET measurements [49].

4.2 Predicting the structural characteristics of tau

4.2.1 Introduction

Much effort has been put into predicting the secondary and tertiary structure of a protein. Yet, whereas the prediction of secondary structure for globular proteins is quite successful, the performance of tertiary structure predictors is not yet satisfying [224]. In 1974 Chou and Fasman presented a method for predicting secondary structures [225]. Therein, an algorithm was developed, based on the empirically found frequency of an amino acid to be in an α -helix, β -sheet or turn, to calculate the probability for a given sequence to form a secondary structure. A more advanced algorithm was published in 1978, termed GOR-method [226, 227]. The decision which conformation is assigned to a residue depends in addition on the conformational probability of the 8 neighboring residues (termed cooperativity). Nowadays, machine-learning algorithms like neural networks, hidden markov models, support vector machines or k-nearest-neighbor approaches, in combination with sequence alignment, increased the prediction accuracy significantly [224]. That structure is more conserved than sequence [228, 229] explains the application and success of predictors including sequence alignment. A further increase in prediction accuracy was achieved by implementing long-range interactions in the program Porter, yielding a correct prediction of 4 out of 5 residues on average [230].

Albeit several techniques are applied for predicting both globular and disordered proteins, structure prediction for disordered proteins is currently more challenging for several reasons. Whereas the protein data bank contains 73.656 structures (as of June 7th, 2011), the databank for disordered proteins, DisProt [231], contains only 643 proteins and 1375 disordered regions (Disprot release 5.7, June 7th, 2011). This imbalance limits the size of training-sets for machine-learning algorithms. For example the training-set for DISOPRED2 [232], a disorder predictor using a support vector machine, contained 176.550 ordered and 4.590 disordered residues [233]. To increase the prediction accuracy disorder-predictors take the amino acid composition into consideration as an additional aspect. IDPs are depleted in 'order-promoting' amino acids (tryptophan, isoleucine, phenylalanine, leucine, cysteine, tyrosine, threonine, asparagine, histidine) and enriched in 'disorder-

promoting' amino acids (valine, lysine, glutamic acid, proline, serine, glutamine and occasionally glycine) [25–27]. However, three different 'flavors' were found which show slight differences in the composition. Flavor C has more histidines, methionines, and alanines than usually found in ordered proteins or the other two flavors. Flavor S has less histidines than usually found in ordered proteins or the other two flavors, and Flavor V has more of the least flexible amino acids (cysteine, phenylalanine, isoleucine and tyrosine) than the other disorder flavors [234]. Additionally, long disordered regions (≥ 30 residues) differ in their amino acid composition compared to short (≤ 30 residues) disordered regions [235]. Therefore, the prediction accuracy is less significantly improved by using sequence alignment and a training set might be biased towards a certain flavor. Another distinctive feature used in disorder-predictors is the amino acid complexity. Disordered proteins usually have a low sequence complexity. For example in htau40 glycine, serine, lysine and proline account for $\approx 40\%$ of all amino acids. However, a simple amino acid composition not necessarily implicates a disordered region, since fibrous proteins exhibit a well folded structure despite their low amino acid complexity [236].

Here we test three different disorder-predictors to assess the capability of predicting three different aspects of tau: intra-residue contacts, residual secondary structure and protein-binding sites.

4.2.2 Predicting intra-residue contacts with IUPred

During folding, a reduction of conformational entropy must be overcompensated by an increasing number of stabilizing residue-residue contacts in order to reach an energetically more favorable state [29]. Accordingly it was theorized, that the amino acid composition of IDPs is unable to sufficiently compensate for the reduced conformational entropy due to a primary sequence unable to form the necessary number of residue-residue contacts [30]. This energy approach is the basis of the program IUPred [237], wherein an algorithm was developed that predicts pairwise interaction energies between two residues. Based on the known structure of 785 globular proteins a 20×20 interaction matrix M was generated, using the approach of Thomas and Dill [238]. The total energy E of a protein can be mathematically expressed as the sum of pairwise energies between residue i and j by using the

interaction matrix M:

$$E = \sum_{i,j=1}^{20} M_{ij}C_{ij} \quad (4.1)$$

with M_{ij} as the matrix element representing the interaction energy between residue i and j and C_{ij} as the number of interactions between residues of type i and j . The total energy E can be decomposed into residue-specific contributions:

$$e_i = \sum_{j=1}^{20} M_{ij}C_{ij} \quad (4.2)$$

with:

$$E = \sum_{i=1}^L e_i \quad (4.3)$$

whereas e_i represents all amino acids of type i that contributes to the total energy via contacts C_{ij} to their interaction partners j , within a certain length L (the width). Since these energy values are 'contaminated' due to secondary structure, the authors want to estimate 'clean' residue specific energy values based only on the amino acid composition. Therefore, the authors proposed that the estimated energy of all residues of type i (N_i) depends on the number of contacts to the residue type j , which in turn is proportional to the fraction of residue type j (n_j):

$$e_i(\text{estimate}) = N_i \sum_{j=1}^{20} P_{ij}n_j. \quad (4.4)$$

The proportionality factor is the 20×20 energy predictor matrix P . The matrix elements of P are derived by minimizing ($\partial Z_i / \partial P_{ij} \rightarrow 0$) the function:

$$Z_i(\text{estimate}) = \sum_k [e_i - N_i^k \sum_{j=1}^{20} P_{ij}n_j^k]^2. \quad (4.5)$$

k -times for k proteins.

Despite the fixed values of the predictor matrix, the energy contribution for every amino acid to the overall energy of the protein depends on the presence and frequency of the other 20 amino acid types as potential interaction partners. For example, the contribution of lysine to disorder or order largely depends upon the adjacent incorporation of more negative or positive charged residues. Since the matrix was derived from a database of globular

proteins only, the program does not rely on any kind of (flavor-biased) disordered dataset. A residue-resolved score for the energy contribution of every position is achieved by considering the potential interaction partners usually up to 100 residues apart. Finally, the inter-dependent score is smoothed over a 21-residue window. To account for very long-range interactions found in tau, a parameter was changed to include the 400 neighboring residues (the width) in the energy calculation. One has to keep in mind that changing the width changes the fraction of each residue type. In addition, the averaging was done over a 11-residue window so that weaker contributions are not averaged out.

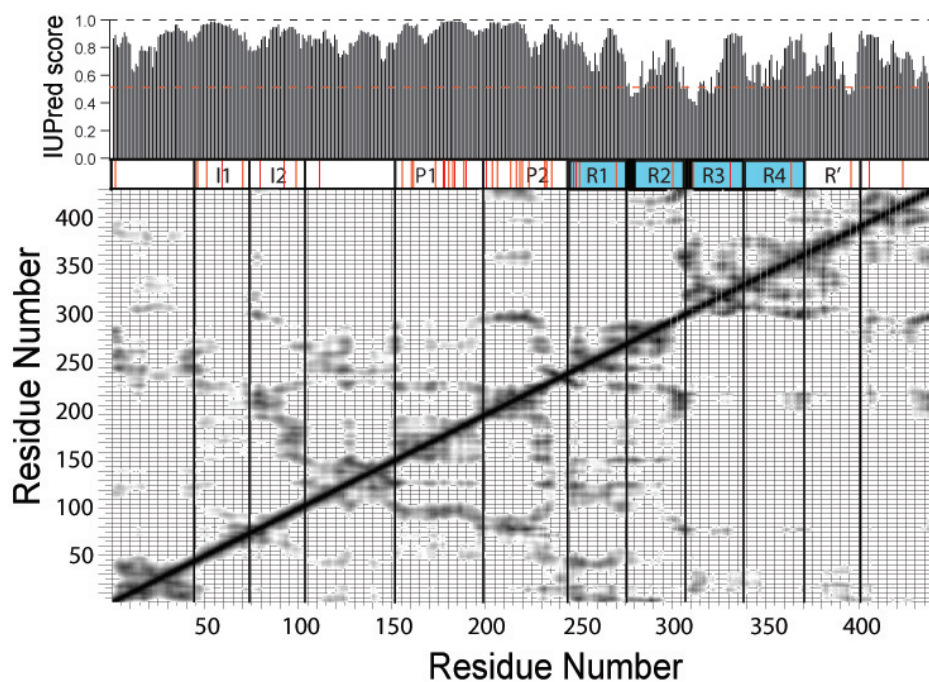


Figure 4.4: Residue-residue contacts. On top: IUPred calculation of the inter-residue contact energy to determine either order or disorder for a residue. The 400 neighboring residues are included in the calculation. The result is smoothed over a 11-residue window. A score value of 0 indicates complete order. A score value of 1 indicates complete disorder. Values below 0.5 are regarded as ordered. Lower part: Average contact map derived from 6 PRE measurements for the seven lowest-energy structures obtained from a calculation in which all distance restraints were enforced onto a single molecule. A continuous grey scale from 3 Å (black) to 22 Å (white) is used. The red lines in the domain organization indicate proline residues.

Structure calculations based on experimental PRE data were done to reveal the overall conformation of htau40. To extract long-range interactions that occur in many of the calculated structures, we additionally determined all C^α distances in each structure and averaged the resulting contact map over the ensemble of structures (Figure 4.4 lower part) [218]. Here we assess the performance of IUPred to predict transient residue-residue contacts by comparing the resulting score values with the residue-residue contacts derived from the contact map. For the first 29 N-terminal residues, the IUPred graph (Figure 4.4 upper part) shows score values between 0.6 and 0.9. When looking at the contact map of tau we found that the first 50 residues favor a compact conformation, also indicated by a wide PRE-broadening profile (Figure 4.2A). For the N-terminal residues of insert I1, the contact map shows a close proximity with repeat 1 (R1). Additionally, a close proximity of residues from I2 and the proline-rich regions P1 and P2 is found. Taken together, these transient residue-residue contacts can explain the IUPred score-drops found for I1 and I2. For the domain between I2 and P1, the IUPred score shows a local minimum for residue 123. This drop can be explained by a combination of short-range residue-contacts due to a α -helical propensity found for residues $^{114}\text{LEDEAAGHVT}^{123}$ and long-range interactions with P1, P2 and R1, as indicated with black spots on the contact map. For P1 the contact map shows an interesting pattern. Every residue of P1 and almost every residue of P2 is in close proximity to each other, indicated by the square of dark and semi dark spots at the diagonal. Additionally, P1 is in close proximity to P2 and I2. However, only three dips are visible with minima for A168, S195 and K225/V229, therefore not indicating the substantial network of residue-residue contacts. This might be a result of the special amino acid composition, since an increased occurrence of prolines is found in the proline-rich region, as indicated by the red lines in the domain organization in figure 4.4. A closer look at the sequence reveals a correlation between high IUPred scores and the presence prolines. Indeed, the predictor matrix used by IUPred assigns low and therefore disadvantageous pairwise interaction values for proline residues, resulting in high IUPred values.

For R1, the contact maps shows a close proximity with the first 40 N-terminal residues, I1 and the domain in between I2 and P1. The IUPred score shows minimum values for residues 256 and 257 and therefore for the region in between four prolines (P247, P249, P251 and P270). Since the contact map would suggest uniformly low IUPred values for R1, the predicted

valley might be again attributed to the flanking proline residues. The two hexapeptide motifs ($^{275}\text{VQIINK}^{280}$ and $^{306}\text{VQIVYK}^{311}$), prone to aggregation [103, 104, 239] at the beginning of R2 and R3 show the lowest IUPred values. The contact map reveals that R2 favors a compact conformation with many additional contacts to residues of the N-terminal part of tau, whereas R3 has contacts with C-terminal residues, indicated correctly by IUPred. High IUPred values of R2 and R3 correlate consistently with the positions of two prolines (P301 in R2 and P332 in R3). Repeat R4 assumes a compact conformation in solution, reflected by the contact map, that is also correctly predicted by IUPred. For residues next to P364, IUPred shows a high value above 0.8. Surprisingly, for R', the lowest IUPred values are found for Y394, only two residues away of P397. Despite the contact map does not indicate a compact conformation or many residue-residue contacts, the sequence $^{390}\text{AEIVYK}^{395}$ shows a remarkable similarity with the second hexapeptide motif, very likely explaining the low IUPred values. For the C-terminal domain, the IUPred score continuously drops towards the very C-terminus, reflecting the increasing residue-residue contacts indicated by the contact map.

We initially used a width of 400 residues, to account for very long-range residue-residue contacts seen in PRE measurements and the contact map. By varying the width, it should be possible to separate the influence of short-, medium- and long-range inter-residue contacts on the IUPred score. Therefore we plotted the IUPred score with a width of 20 (short-range), 100 (medium-range) and 400 (long-range) residues (Figure 4.5). The IUPred score drops for the first 40 N-terminal residues that have smallest values for the smallest width of 20 residues, whereas the highest IUPred-values are found for a width with 100 residues and intermediate values for a width with 400 residues. This interesting behavior can be explained by looking at the contact map. A compact conformation accounts for the low IUPred score with a width of 20 residues. Since the width determines the frequency of favorable as well as unfavorable amino acid types, increasing the width without additional favorable residue-contacts would increase the IUPred score. This is seen for the 40 N-terminal residues with a width of 100 residues. The contact map does not indicate any additional close proximities with residues that are 100 residues apart. In contrast, when increasing the width to 400 residues, the IUPred score drops, now accounting for additional contacts. When looking at the contact map, these additional contacts are due to a

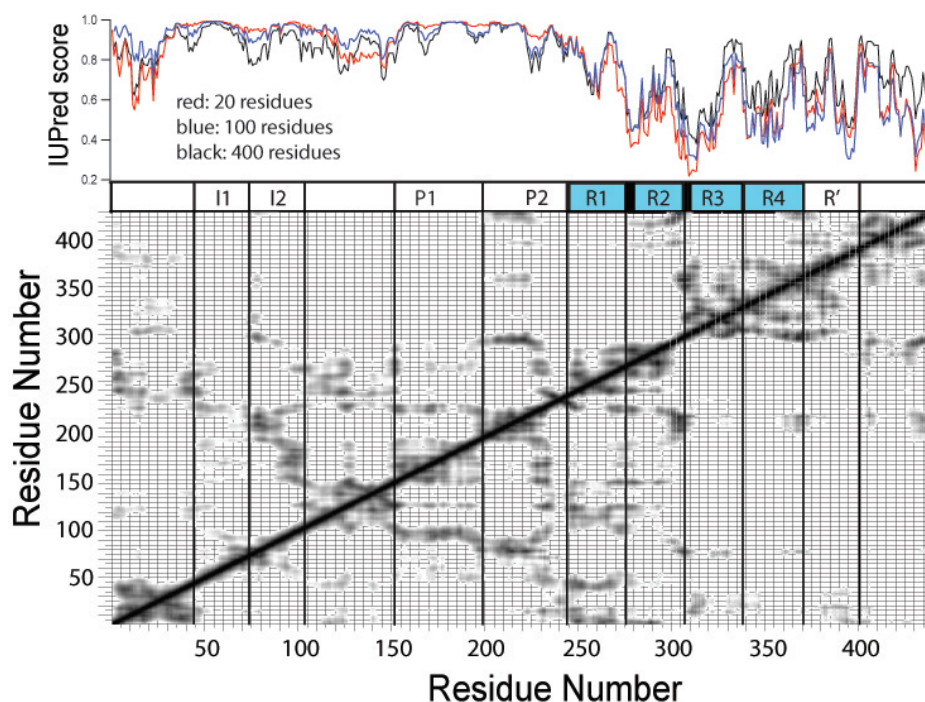


Figure 4.5: Range-dependence of residue-residue contacts. On top: IUPred calculation of the inter-residue contact energy with a width of 20 (red), 100 (blue) and 400 (black) residues. The results are smoothed over a 11-residue window. A reduced score value indicates a more favorable interaction energy for this residue. Note that the IUPred score ranges from 0.2 to 1.0 for a better representation. Lower part: Average contact map derived from 6 PRE measurements for the seven lowest-energy structures obtained from a calculation in which all distance restraints were enforced onto a single molecule. A continuous grey scale from 3 Å (black) to 22 Å (white) is used.

close proximity of the N-terminus with the repeat regions (200-250 residues away).

For the dip found around residue 80 the lowest and highest IUPred score is found for a width of 400 and 20 residues, respectively. The contact map reveals only minor short-range contacts and no additional contacts with the next 100 residues, explaining the low and intermediate values for the width of 20 and 100 residues. However, the region around residue 80 shows strong contact with the proline-rich region, roughly 120-150 residues away, explaining the lowest IUPred value for a width of 400.

In the proline-rich regions P1 and P2 three dips are visible with minima

for A168, S195 and K225/V229. All three dips depend strongly on the used width, corresponding to increasing residue-residue contacts, in general agreement with the contact map. However, a more pronounced dip for K225/V229 could be expected when a width of 400 residues is used. The same expectation holds true for R1, that an increasing width should correspond to a decreasing IUPred value. Though, almost no difference between the width of 20, 100 and 400 residues is seen, therefore not addressing the range-dependence. It might be that an increasing width counterbalances favorable and unfavorable residue-residue contacts. Interestingly, for the two hexapeptide motifs, the lowest IUPred value are found by using a width of 20 residues, explainable by the dominating presence of tight short-range contacts. For the C-terminal domain it can be seen that IUPred can not account for interactions with the repeat regions, since a residue-width of 20 residues shows the lowest score value.

In summary, IUPred shows for many residues an IUPred score below 0.8, indicating the intricate network of residue-residue interactions, in agreement with the PRE profiles. Moreover, a good correlation is established between the IUPred score and the contact map for regions lacking prolines. For proline-containing regions the IUPred score is biased towards high values. Especially for the proline-rich regions P1 and P2, which interact with many residues of tau, IUPred shows unexpectedly high values. By varying the width we assessed the range-dependence of the IUPred score and found a good correlation between the range of residue-residue contacts and the IUPred-variation.

4.2.3 Predicting residual secondary structure

PreDisorder is a predictor using neural networks [240]. It is based on MULTICOM-CMFR, which was, together with MULTICOM [241], among the best predictors in the eighth 'critical assessment of techniques for protein structure prediction' (CASP8) [242,243]. In PreDisorder, every residue is described by 25 values. The target protein sequence is first aligned against several template sequences using PSI-BLAST [244], creating an input profile. From the PSI-BLAST profile, each amino acid at the corresponding position from the target sequence is assigned with a frequency value (accounting for 20 values). In addition, a secondary structure prediction is made for each residue (coil, sheet, helix; accounting for 3 values). Furthermore, solvent accessibility pre-

diction is carried out, determining if the residue is more or less than 25% exposed, accounting for the last to values. These values were processed by using a recursive neuronal network, resulting in a residue-specific probability of disorder. The neuronal network was trained with a datasets and 10 times cross-validated. Residues with values above 0.5 are considered disordered.

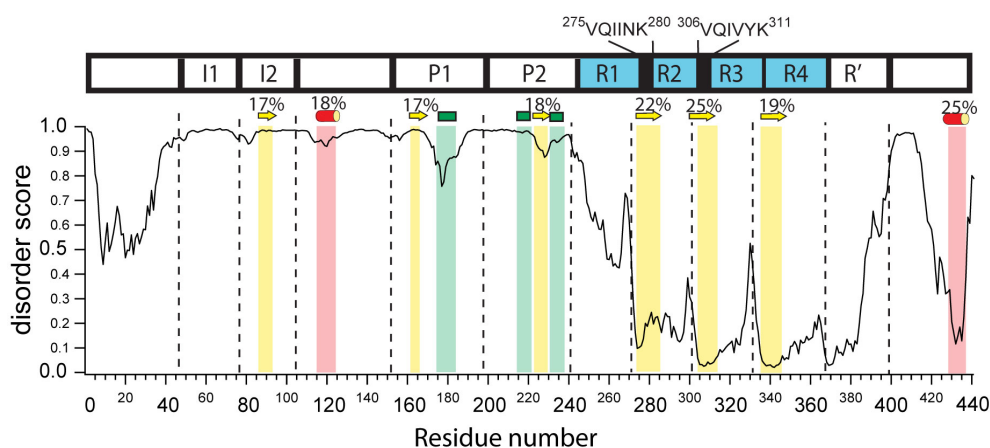


Figure 4.6: Predicting residual structure. PreDisorder assigns 25 values to every residue to predict order/disorder. A neuronal network is used to combine and weight the the values in a residue-specific manner. Residues with values above 0.5 are considered disordered. In the background regions with β -sheet (yellow), α -helical (red) and poly-proline II helical (green) propensities are indicated.

An obvious feature of PreDisorder compared to other disorder predictors is the high variability of the score values (Figure 4.6). The first 35 residues show a score below 0.5 for several residues (E9, E12, L20-D22, K24) indicating a strong tendency towards a relatively ordered structure. However, NMR data, like C^α secondary shifts and $^3J_{HNH\alpha}$ couplings, did not indicate any residual structure [218], despite the fact that the very N-terminus is known to maintain the spacing between MTs [33], to interact with the axonal plasma membrane [245] and to affect the fast axonal transport in PHF tau [116]. In addition, the first two β -sheet propensities ($^{86}\text{GKQAAQ}^{92}$, average score: 0.98 and $^{161}\text{GQKGQA}^{166}$, score: 0.98) where not detected. Even scaling the predicted probability by a lower threshold to increase the sensitivity of predictions does not reveal these secondary structure propensities. One might speculate that the combination oft shortness and low propensity prevents the prediction of these two secondary structure elements. However, a small but

significant dip for the first helical propensity (residues ¹¹⁴LEDEAAGHVT¹²³) is found. The first polyproline II helical propensity shows a clear dip with a minimum score of 0.77 for residue P177, whereas the other two polyproline II propensities were not detected. Nevertheless, the β -sheet propensity in P2 showed a clear dip (minimum of 0.87 for V228). When approaching the repeat region, four local minima are observed. The first minimum is located in the middle of repeat 1 with N265 exhibiting the lowest score (0.43), a region where no secondary structure propensity could be assigned. However, it was shown that this region, especially ³⁷²ETHKLTFR^{ENAK}³⁸³, interacts with the aggregation inducer heparin and also transiently contacts MTs when tau is bound to (Figure 4.8) [61,218]. The remaining three local minima coincide perfectly with the three β -sheet propensities found by NMR. Another minimum, located at the beginning of R' for the sequence ³⁶⁷GNKKIETH³⁷⁴ is found. Interestingly, the sequence similarity of ³⁹⁰AEIVYK³⁹⁵ with the second hexapeptide motif (³⁰⁶VQIVYK³¹¹) does not result in low PreDisorder values, but to a correct raising in the score, in agreement with experimental data. The score rises to an average value of 0.97 for residues ⁴⁰³T-V⁴¹¹. The last local minimum is found to coincide with the C-terminal α -helical propensity with an average score of 0.22 for ⁴²⁸L-A⁴³⁷.

A major drawback of PreDisorder is the lack to differentiate between types of secondary structure. We wanted to investigate if predictors for globular proteins, especially designed to discriminate regions containing α -helices and β -sheets, may solve that problem. For that, we first compared the results of AGADIR, a program specifically predicting the residue-specific helical content [246–249] (Figure 4.7). It can be seen, that AGADIR predicts for ¹³⁷S-A¹⁴⁵ and ⁴²²S-Q⁴³⁹ a helical content of 5.3% and 5.2%, respectively. Whereas NMR does not show α -helical propensity for ¹³⁷S-A¹⁴⁵, the C-terminal prediction coincides with the NMR data. For the N-terminal helical propensity (residues ¹¹⁴LEDEAAGHVT¹²³) AGADIR assigns a helical content of 2.5%, too low to be regarded as significant. Even the helical content found for ¹³⁷S-A¹⁴⁵ and ⁴²²S-Q⁴³⁹ is very close to the detection border, since the standard deviation is $\pm 6\%$. The output of the web-server based PredictProtein prediction distinguishes between α -helices, β -sheets and disorder [250]. The prediction with PredictProtein does not indicate any secondary structure until residue K274 (Figure 4.7). For the β -sheet propensity found in R2 PredictDisorder shows an increase in the score for the β -sheet prediction. These score values increase for R3 and R4, matching with the experimentally found

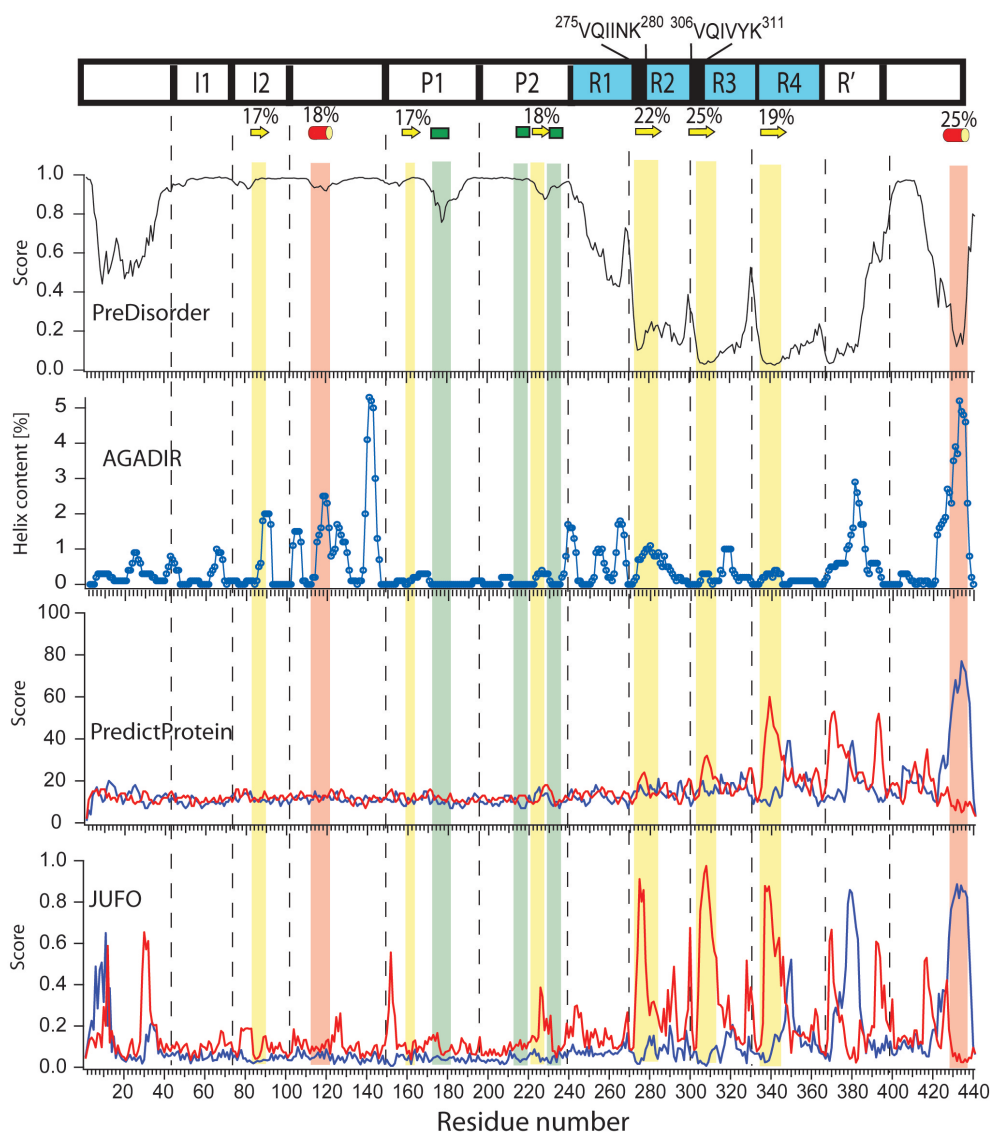


Figure 4.7: Predicting residual structure. For comparison, PreDisorder is plotted on top. Thereunder, a plot of AGADIR is shown. AGADIR specifically predicts the α -helical content. The PredictProtein output distinguishes between α -helices (blue line), β -sheets (red line) and disorder (output not shown). For the lowest panel, the prediction of Jufo is shown. The output usually distinguishes between α -helices (blue line), β -sheets (red line) and disorder (output not shown).

β -sheet propensities. Interestingly, α -helical content is found for R4, slightly shifted to the found β -sheet content. In R' PredictProtein shows two regions with an increased score for β -sheets and one region with an increased score

for α -helix in between, coinciding with helix content found from AGADIR. The two helical propensities found in R4 and R' are likely due to a found turn-propensity for these residues [61, 212]. Like AGADIR, PredictProtein predicts the C-terminal helical propensity correctly. Jufo [251], another web-server based predictor correctly identifies the three β -sheet propensities in R2-R4 and the C-terminal α -helical propensity. In addition, a similar behavior like PredictProtein is found for R'. N-terminal of R1 several higher scores for β -sheet propensities are found whereas only in P2 (for ²²⁴KKVAVVR²³⁰) one β -sheet propensity coincides with experimental NMR data.

To summarize, PreDisorder performs very good and reliably in predicting out-of-the-ordinary disordered regions. The high variability of the score values needs, however, a careful interpretation, since the impression of rigid secondary structure elements in tau may arise due to very low PreDisorder values. To actually assign a structure to these regions we tried three predictors for globular proteins. We found that the C-terminal helical propensity is reliably identified by PredictProtein, Jufo and AGADIR. In addition, three highest β -sheet propensities in R2-R4 could be detected by PredictProtein and Jufo. Whereas PredictProtein is too insensitive, Jufo shows too many false negative results, complicating a reliable prediction with Jufo. Erroneously, both programs also assigned a region in R' to have an α -helical propensity. For the weaker residual structure propensities N-terminal of R1 no reliable structure assignment was possible. Yet, predicting residual secondary structure remains highly challenging and prone to an erroneous interpretation.

4.2.4 Predicting tau-binding sides with ANCHOR

ANCHOR tries to predict segments of disordered proteins that undergoes a disorder-to-order transition upon binding [252, 253]. The prediction relies on the energy predictor matrix P of IUPred [237]. In total three score values are assigned to each residue. For the first value, the IUPred score of residue S at position k is averaged within a residue window w_1 :

$$S_k = \frac{1}{N(w_1)} \sum_{k-w_1}^{k+w_1} s_j \quad (4.6)$$

For the second value the pairwise interaction energy is estimated for a residue when forming intramolecular contacts. This estimation is similar to IUPred,

but with a window size w_2 optimized during training:

$$E_i^{int,k} = \sum_{j=1}^{20} P_{ij} n_j^k(w_2). \quad (4.7)$$

$n_j^k(w_2)$ represents the frequency of residue type j around residue k within a residue window w_2 , P_{ij} is the energy predictor matrix. The third value estimates the energy a residue may receive when interacting with a globular protein:

$$E_i^{glob,k} = \sum_{j=1}^{20} P_{ij} f_j^{glob}, \quad (4.8)$$

with f_j^{glob} as the frequency of residue j , found on average in globular proteins. The difference between $E_i^{int,k}$ and $E_i^{glob,k}$ estimates the energy a residue may gain upon binding:

$$E_i^{gain,k} = E_i^{int,k} - E_i^{glob,k}. \quad (4.9)$$

The final prediction score is a linear combination of these three values:

$$I_k = p_1 S_k + p_2 E_i^{gain,k} + p_3 E_i^{gain,k}, \quad (4.10)$$

with the coefficients p_1 , p_2 and p_3 determined during training. w_1 is fixed to 10 residues, accounting for a window size of 21 residues.

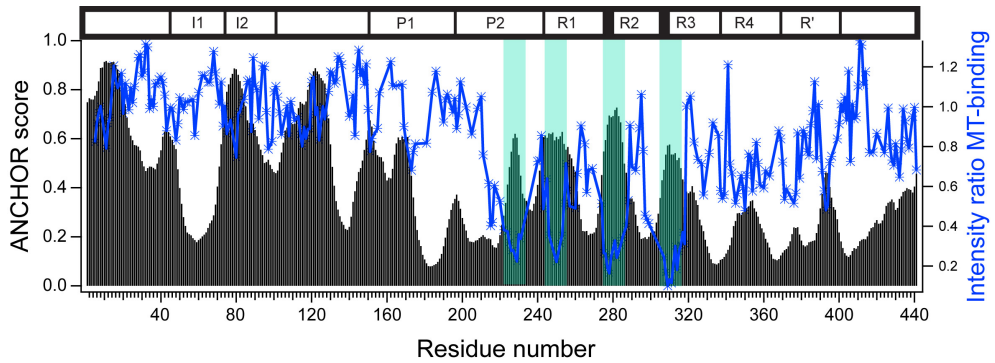


Figure 4.8: Prediction of tau-microtubule binding-sides. ANCHOR prediction (black) is overlaid with the intensity profile of tau when bound to microtubules (blue). High ANCHOR values indicate a high propensity to bind. Low intensity values indicate regions of tau that bind to microtubules. The hot spots of tau-microtubule binding are indicated (mint). On top, domain organization of tau. The window size is 21 residues.

The program reveals many regions that exhibit an ANCHOR score greater than 0.5, predicting potential binding regions (Figure 4.8, black bars). Local maxima occur throughout the whole sequence of tau, especially within the projection domain local maxima are observed for M11, L43, P78, H121, K150 and R170. Additional local maxima above 0.5 occur in P2 (V228), R1 (around M250), R2 (L282) and R3 (Y310). When looking at the intensity ratio between free and MT-bound tau [218] (Figure 4.8, blue line) it can be seen that the local ANCHOR-maxima in P2 and the repeat regions exactly coincide with the hot-spots of MT-binding (Figure 4.8, mint bars). Along with the hot-spots, N-terminal regions around residue 10, 42, 80, 120, 150 and 175 transiently bind to MTs, indicated by reduced intensity ratios. Exactly these regions exhibit a high ANCOR score. C-terminal of the MT-tau binding hot-spots, two regions show a distinct drop in the intensity ratio (around 376 and 393), again perfectly superimposed with local ANCHOR-maxima. Interestingly, when tau is incubated with tubulin dimers, strong signal attenuation is present for the regions around P78 and H121 (data not shown).

At first view, the result given by ANCHOR seems a little surprising. When considering only MT-binding, one expects only five to six clear peaks in the ANCHOR-prediction, corresponding to three to four MT-binding regions (in the repeat regions) and the 'jaws' regions (in P2 and R') [60, 61]. However, since the binding to and the assembly of MTs is highly regulated via kinases, additional binding must occur at different regions of tau. From 85 potential phosphorylation sites (45 serines, 35 threonines and 5 tyrosines) (Figure 1.6), many of them were shown to be phosphorylated in vivo [72, 75]. When plotting the phosphorylation sites onto the ANCHOR prediction results an ambiguous picture is presented (Figure 4.9). Whereas some phosphorylation sites are located on top of predicted maxima, many phosphorylation sites are located next to maxima or within valleys. Experimental data of tau phosphorylated with MARK showed that the serines in the KXGS motifs were phosphorylated to 100% (S262, S324 and S356) and 81% (S293), respectively. In addition, S305 (66%), S352 (65%), S413 (45%) and S416 (80 %) were phosphorylated (Figure A.1B and C, boxed peaks were assigned by Dr. Martin Schwalbe). By comparing experimental data and theoretical prediction, no obvious correlation between the degree of phosphorylation and the ANCHOR-landscape is found.

ANCHOR prediction for the pseudophosphorylated residues S199E, S203E,

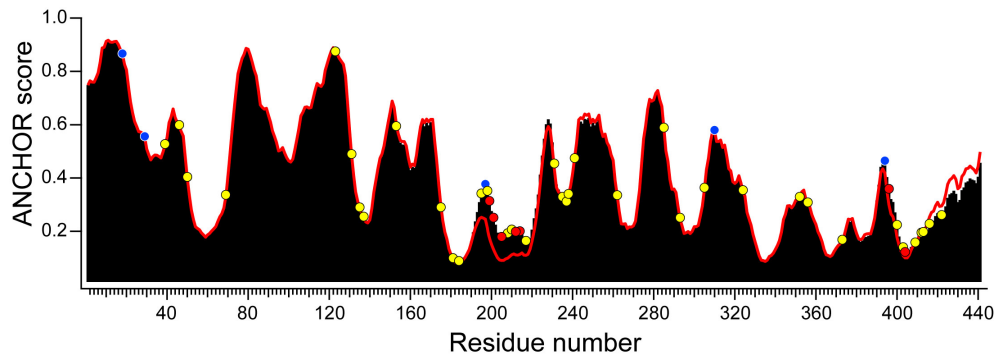


Figure 4.9: Potential binding-sides for kinases. ANCHOR prediction (black) superimposed with serines/threonines (yellow circles) and tyrosines (blue circles) that can be phosphorylated. The red line indicates the ANCHOR prediction for the pseudophosphorylated residues S199E, S203E, T205, T212E, S214E, S396E and S404E (red circles). High ANCHOR values indicate a high propensity to bind. The window size is 21 residues.

T205, T212E, S214E, S396E and S404E Figure (4.9, red circles) resulted only in local differences, not affecting the hot-spots of MT-binding (4.9, red line). A confirmation of this prediction can be done by incubating the tau E-mutant with microtubules. According to the prediction, only the MT-binding around residue 200 should be weakened, leaving the other binding regions unaffected. Taken together, a very good correlation between the ANCHOR prediction and MT binding is observed. Additional peaks in the ANCHOR-prediction might be attributed to kinase binding-sides. However, further elucidation supporting this interpretation is needed.

4.3 Structural impact of pseudophosphorylation on the global and local structure of 441-residue tau

4.3.1 Introduction

As it was pointed out in the introduction, phosphorylation plays an essential role for tau. Not only is the binding and release of tau from MTs modulated via a unique and complex phosphorylation pattern within tau's microtubule binding region (R1-R4). Additionally, phosphorylation seems to be closely linked to tau's pathological role. Phosphorylation, especially in the predominantly positive charged repeat regions, might be capable of opening up the paperclip conformation, therefore easing intermolecular interactions, especially between the aggregation-prone hexapeptide motifs at the beginning of R2 and R3 [49, 103, 104]. Evidence supporting this argumentation comes from truncation experiments where the presence or close proximity of the N- and the C-terminus reduces the aggregation rates [113–115]. Tau in its longest isoform with 441 residues possesses 85 potential phosphorylation sites (45 serines, 35 threonines and 5 tyrosines) (Figure 1.4). To circumvent the problem of low specificity of kinases and incomplete phosphorylation of residues, we choose the approach of phosphorylation-mimicking mutations. Therefore, serines and threonines were mutated into glutamates, mimicking the attachment of negatively charged phosphorus. Previous reports showed, that pseudophosphorylation of tau is capable of successfully mimicking phosphorylation [254, 255]. The work reports on the impact in the local and global structure of 441-residue tau with phosphorylation-mimicking mutations within the epitopes recognized by the AD diagnostic antibodies AT8 (S199E, S202E, T205E), AT100 (T212E and S214E) and PHF1 (S396E and S404E).

4.3.2 Global structural changes

96 % of the backbone resonances of mutated htau40 were assigned using 3D HACANH and HNN experiments at 5 °C, pH 6.8 [207, 208]. To investigate global structural changes, paramagnetic relaxation enhancement (PRE) experiments and diffusion experiments were employed and compared to wildtype tau (wt tau) [168, 203, 218]. Since the magnetic moment from

an unpaired electron spin is 657 times larger than that of a nuclear spin, the range of dipolar interaction between an unpaired electron from a paramagnetic molecule and nearby protons in the polypeptide chain extends to up to 25 Å [169,170]. We attached paramagnetic MTSL at 11 different positions (C15, C72, C125, C178, C239, C256, C291 and C322, C322, C352, C384, C416) evenly distributed over the protein. The pseudophosphorylated

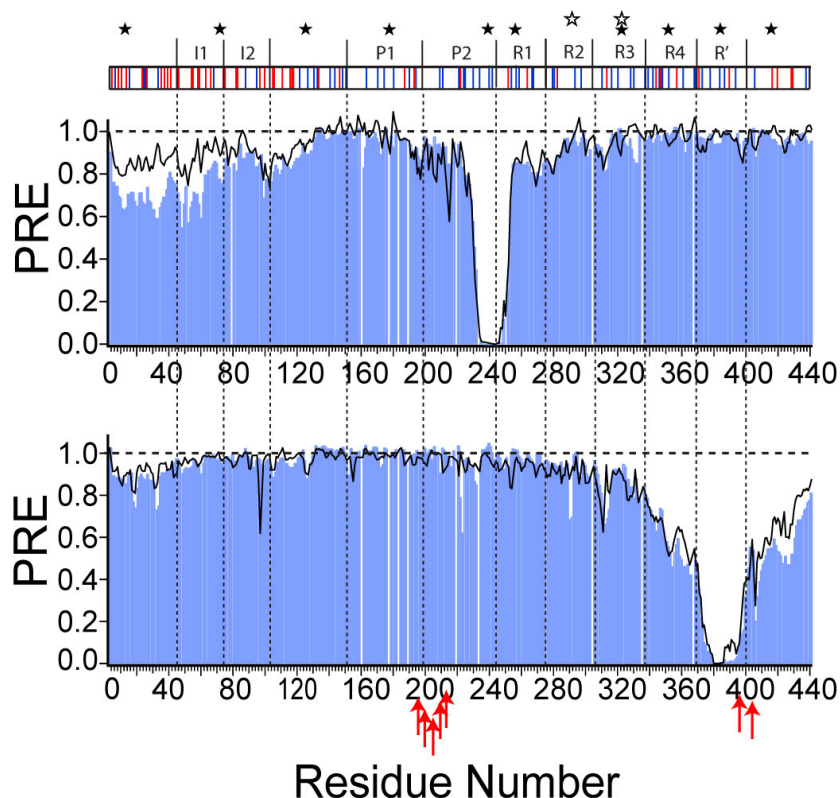


Figure 4.10: Comparison of the PRE broadening profiles. PRE broadening of the E-mutant (black line) and wt httau40 (blue bars) are shown. The broadening profile for the E-mutant is less pronounced at the N- and C-termini of the PRE profiles with MTSL at C239 and C384, respectively. This indicates a less compact conformation for the E-mutant. The domain organization with positively (blue) and negatively (red) charged residues is shown on top of the upper panel. In addition, the stars indicate the positions of the MTSL-label, whereas the open stars representing the double MTSL-labeled mutant with MTSL at positions C291 and C322. The red arrows below the lower panel indicating the mutation sites.

tau protein (E-mutant) showed a distinct broadening pattern compared to

wildtype httau40. The most dramatic change in the PRE profile is found for MTSL at residue 239 (Figure 4.10, upper panel). The introduction of MTSL at residue 239 resulted in significant less broadening for residues $^3\text{E-V}^{80}$. Less broadening indicates an increased distance between the spinlabel attached in the P2 region and the N-terminal residues. The increased distance of the N-terminus from the proline rich regions was confirmed by PRE profiles with MTSL attached to residue 15, 72 and 178 (Figure 4.11 A, B and D). Whereas MTSL@125 shows an increased distance of the N-terminus from residues lying in between the inserts and the proline rich region (Figure 4.11 C).

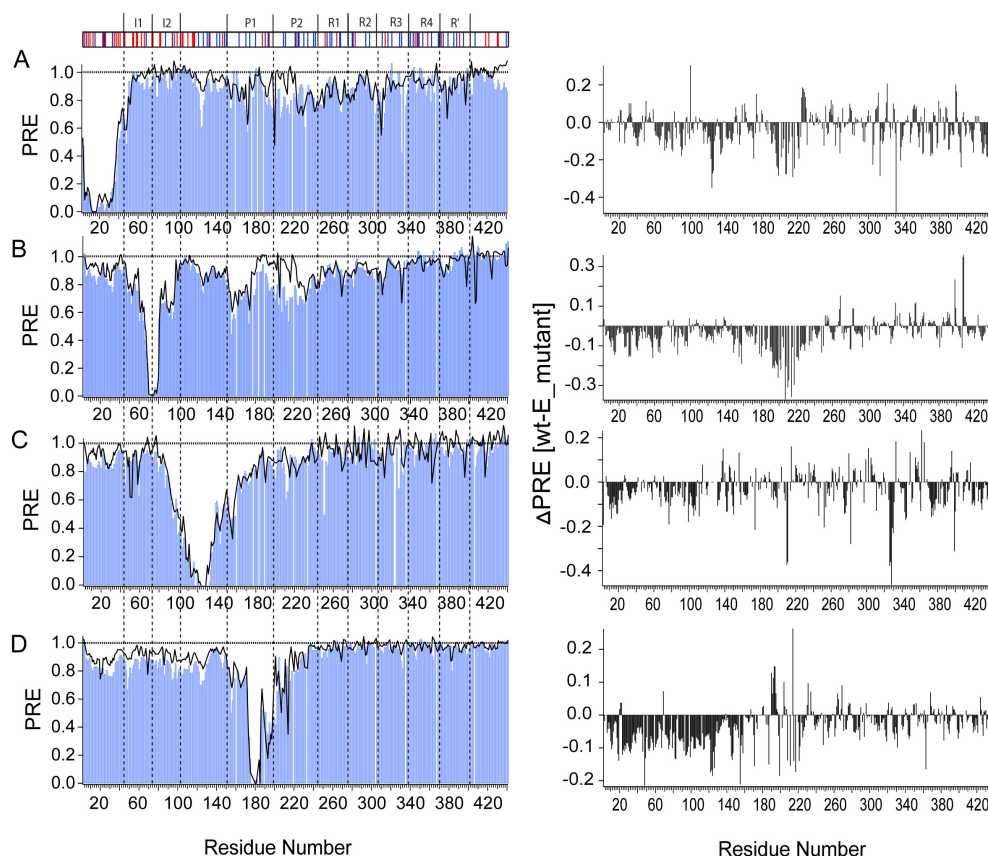


Figure 4.11: Comparison of the PRE broadening profiles. The broadening profiles for MTSL attached to residue 15 (A), 72 (B), 125 (C), 178 (D) is shown. Always on the right side the difference of the broadening profiles is shown (wt-E-mutant), whereas negative values indicating less broadening for the E-mutant. On top, the domain organization is shown, wherein positively charged (blue) and negatively charged (red) residues are indicated.

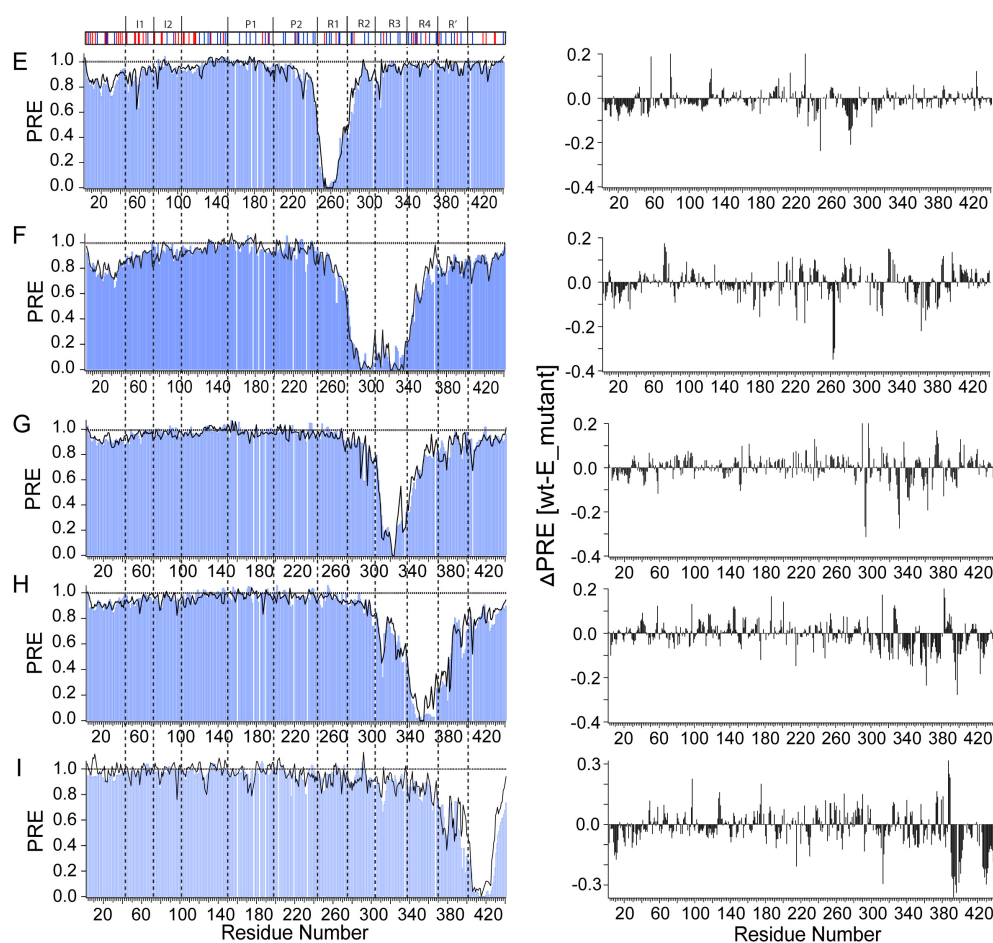


Figure 4.12: Comparison of the PRE broadening profiles (continued). The broadening profiles for MTSL attached to residue 256(E), 291 and 322 (F), 322 (G), 352 (H) and 416 (I) is shown. Always on the right side the difference of the broadening profiles is shown (wt-E-mutant), whereas negative values indicating less broadening for the E-mutant. On top, the domain organization is shown, wherein positively charged (blue) and negatively charged (red) residues are indicated.

The movement of the N-terminus away from the proline rich regions P1 and P2 can be seen as a consequence of changed electrostatic interactions. The introduction of negatively charged residues in the predominantly positively charged proline rich regions P1 and P2 caused a weaker electrostatic attraction of the predominantly negative charged N-terminus. The E-mutant PRE profile with MTSL attached to residue 72 showed an almost complete absence of signal broadening for residues near the mutations in the AT8 and

AT100 epitope ($^{185}\text{A-R}^{217}$ in Figure 4.11 B). Inspection of the primary sequence revealed the presence of 4 negatively charged residues close to the MTSL-tag (E73, D74, D81, E82) which electrostatically repel the region of the AT8 (S199E, S202E, T205E) and AT100 (T212E and S214E) epitope. One should note, that the mutations in the AT8 and AT100 epitope changed the net charge in the region $^{185}\text{A-R}^{217}$ from +2 to -3. The same characteristic broadening profile found for MTSL@72 is visible for MTSL@C15, which is in close vicinity to the negatively charged residues E12 and D13 (see the difference plots in Figure 4.11A and B). The absence of negative charges from the spinlabel MTSL@125 within a range of ± 7 residues resulted in the absence of such characteristic broadening for the region $^{185}\text{A-R}^{217}$, further substantiating the electrostatic importance for data interpretation. To complete the picture of how the AT8 and AT100 epitope mutations change long-range interactions, a tau mutant was produced harboring the MTSL-tag at residue 178, very close to the first mutation region. As it can be seen from Figure 4.11 D, the complete N-terminus until residue 172 showed an overall picture of reduced signal broadening for the E-mutant, indicating that the N-terminus moves away from the proline-rich region. On the other hand, MTSL@178 does not indicate any distance differences C-terminal of the AT8 and AT100 mutation (Figure 4.11 D), indicating no influence to the C-terminal global structure. Less dramatic results are found for samples with MTSL-tags at residues 256 291/322, 322 and 352 (Figure 4.12 E-G). However, the difference-plot of wt httau40 and the E-mutant revealed an N-terminal opening of the paper-clip conformation.

Interestingly, MTSL@C15 showed that the N-terminus and the very C-terminus are on average farther apart in the E-mutant (Figure 4.11 A). Therefore we wanted to know, if the C-terminus might also be affected by the pseudophosphorylation. Attaching the MTSL-tag to residue 352, 384 and 416 revealed an opening of the C-terminus (Figure 4.12 H and I), which is most clearly seen for MTSL attached to residue 384 (Figure 4.10, lower panel). Again, we placed one MTSL-tag (MTSL@416) very close to the second mutation region. The broadening profile for the C-terminal region clearly shows a less pronounced broadening profile for the E-mutant (Figure 4.12 I). Similar to MTSL@15, MTSL@416 revealed an increasing distance between the N- and C-terminus, indicated by negative peaks in the difference plot at the very N-terminus, corroborating the results from MTSL@15. However, attaching the MTSL-tag to residue 352 and 384 revealed no broadening differences

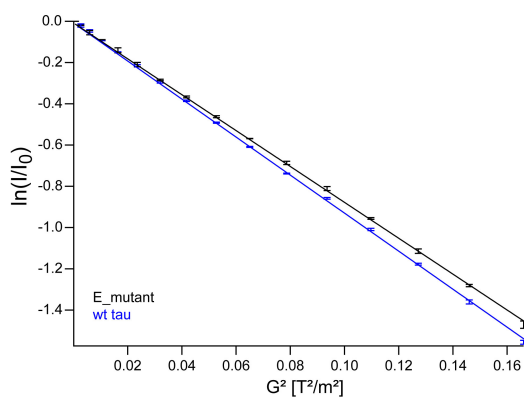


Figure 4.13: Determination of the diffusion constant. By using the pulseprogram PGSTE-WATERGATE [203], the diffusion coefficient of wt ht40 (blue line) and E-mutant (black line) was determined. The well known R_h of lysozyme [199, 204] was used to calculate the R_h of tau.

N-terminal to the mutation. These results, together with analysis of the N-terminal MTSL-mutants, show that the N- and C-terminal 'arms' swing away from the proline-rich and repeat region, whose long-range interactions are unaffected.

To support and quantify the amount of opening of the paperclip conformation diffusion experiments were conducted which directly measures the hydrodynamic radius (R_h), allowing a comparison of the average difference of R_h for the two structure ensembles. We measured for wt tau an ensemble averaged R_h of 61,6 (± 0.2) and for the E-mutant 65,2 (± 0.3), corroborating the PRE results (Figure 4.13)

4.3.3 Local structural changes

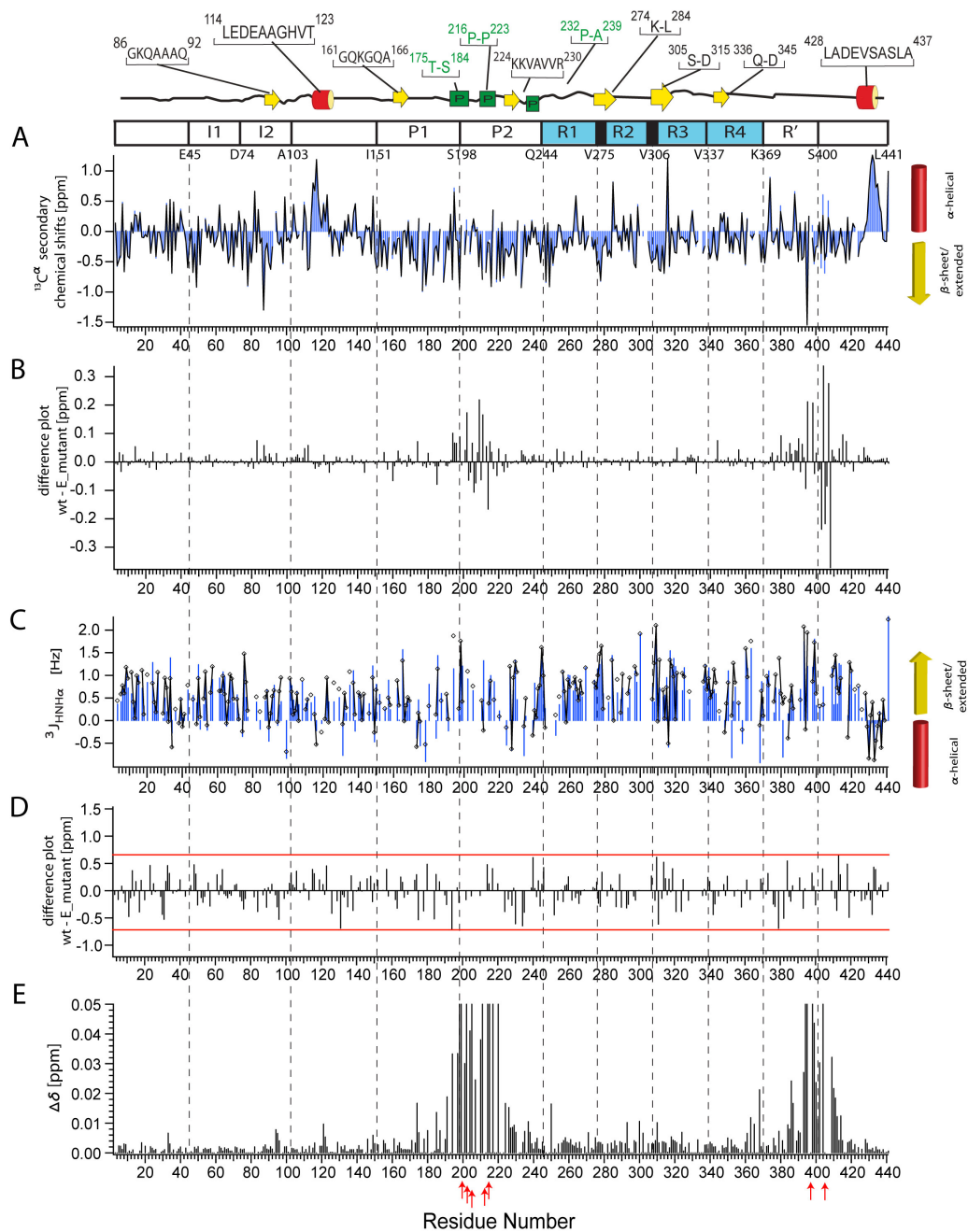
To see if global structural changes are accompanied by local structural changes, we investigated C^α secondary chemical shifts and $^3J_{HNH^\alpha}$ couplings [153, 256]. The deviation from random coil values for C^α chemical shifts [162, 163, 257] and $^3J_{HNH^\alpha}$ couplings [196] are common indicators for residual secondary structure, which can be compared to residual secondary structure found for wt ht40 [218]. As expected for a disordered protein, possessing an extended structure, many residues display small negative and positive deviations from random coil values for the C^α secondary chemical shifts and $^3J_{HNH^\alpha}$ secondary couplings, respectively (Figure 4.14A and C). The dif-

ference plots more clearly indicate the similarity of C^α secondary chemical shifts and $^3J_{HNH^\alpha}$ secondary couplings between the E-mutant and wt tau (Figure 4.14B and D), confirming that pseudophosphorylation in the AT8, AT100 and PHF1 epitope is not capable of inducing secondary structure changes [258]. Typical for a highly flexible backbone, no chemical shift perturbations for residues more than 5 amino acids away from the mutation sites in the E-mutant were observed (Figure 4.14E).

4.3.4 Conclusion

To summarize, we demonstrated that pseudophosphorylation at the AT8, AT100 and PHF1 epitopes changes the electrostatic interactions between the repeat and proline-rich regions and the N- and C-terminus. The evenly distributed MTSL-tags revealed, that pseudophosphorylation of S199E, S202E, T205E T212E and S214E resulted in a movement of the N-terminus away from the proline-rich regions, whereas the mutations S396E and S404E caused a movement of the C-terminus away from the repeat region. Moreover, since MTSL attached to the N-terminus did not reveal any intensity-differences for the C-terminus and vice versa, we showed that a distribution of paramagnetic probes over the whole sequences is necessary to reveal all structural differences in IDPs. The opening of the paperclip conformation is clearly driven by a weakening of electrostatic attraction. This conformational enlargement was confirmed by diffusion experiments. We found an increase of the hydrodynamic radius from 61,6 (± 0.2) for wt httau40 to 65,2 (± 0.3) for the E-mutant, in agreement with a previous study [259]. On the other

Figure 4.14 (following page): Residual secondary structure of the E-Mutant. A) C^α secondary chemical shifts are plotted for wt httau40 (blue bars) and the E-mutant (black line). B) The difference plot shows only very weak differences of the C^α secondary chemical shifts for residues away from the mutation sites. Very close to the mutation sites, uncorrelated differences can be observed, however, not indicating any structural change. C-D) A similar picture is found for the $^3J_{HNH^\alpha}$ coupling constants. The difference plot reveals that no peak shows a significant different $^3J_{HNH^\alpha}$ value (by crossing the red lines). The domain organization is shown above panel A). On top, cartoon-representation of the secondary structure elements found for wt httau40. E) 1H - ^{15}N chemical shift differences clearly show the sites of mutation.



hand, only uncorrelated local secondary structural changes in close proximity to the mutation-sites were found, suggesting no increase in the secondary structure. However, it was shown that exactly this mutant of tau aggregates faster than wt tau [258], suggesting that loosening of the intricate network of long-range residue-residue contacts must be an important step towards aggregation.

4.4 Molecular basis of the recognition of tau aggregates by conformation-specific antibodies

4.4.1 Introduction

Since more than 25 years it is known that tau plays a role in AD [260, 261], whereas it took another ten years to clarify and accept, that tau alone is a sufficient agent to cause neurodegenerative diseases [262, 263]. In 2011, it is beyond controversy that the two major targets in AD diagnosis and treatment are tau and $A\beta$ [264–266]. For improved AD diagnosis and treatment antibodies were developed that detect changes in the phosphorylation and the conformation of tau [221, 267, 268]. Conformation-specific antibodies have the advantage to detect conformational changes that are one of the earliest alterations of tau in AD [269, 270]. The monoclonal antibodies Alz50 and MC1 recognize conformational changes in the tau protein that appear before the assembly of PHFs and are then also found in PHF, but are not present in normal brain [221, 271]. The specificity of Alz50 and MC1 for pathological tau is due to a unique conformation of tau in the disease state and requires two discontinuous epitopes that are separated by about 300 residues and are located at the N-terminus (residues 1-18) and in the repeat region (residues 313-322) [221, 271, 272] (Figure 4.15A). The molecular mechanism of recognition, however, is unknown due to missing high-resolution information about the pathological conformations of tau. To obtain insight into the pathological conformation of tau and its mechanism of recognition by conformation-specific antibodies, we characterized the dynamic structure of PHFs formed by 441-residue tau using solution-state NMR spectroscopy enhanced by high-resolution magic angle spinning (HRMAS) [190].

4.4.2 Assignment of PHF-tau

We aggregated 441-residue tau (htau40), the longest isoform of tau that is present in the human central nervous system (Figure 1.4 and 4.15A), into insoluble filaments. According to electron microscopy, the tau-filaments (Figure 4.15 and 4.16) were morphologically similar to those isolated from AD neuronal tissue. Moreover, NMR diffusion experiments [165] demonstrate that the observed NMR signals arise from aggregated tau protein with a

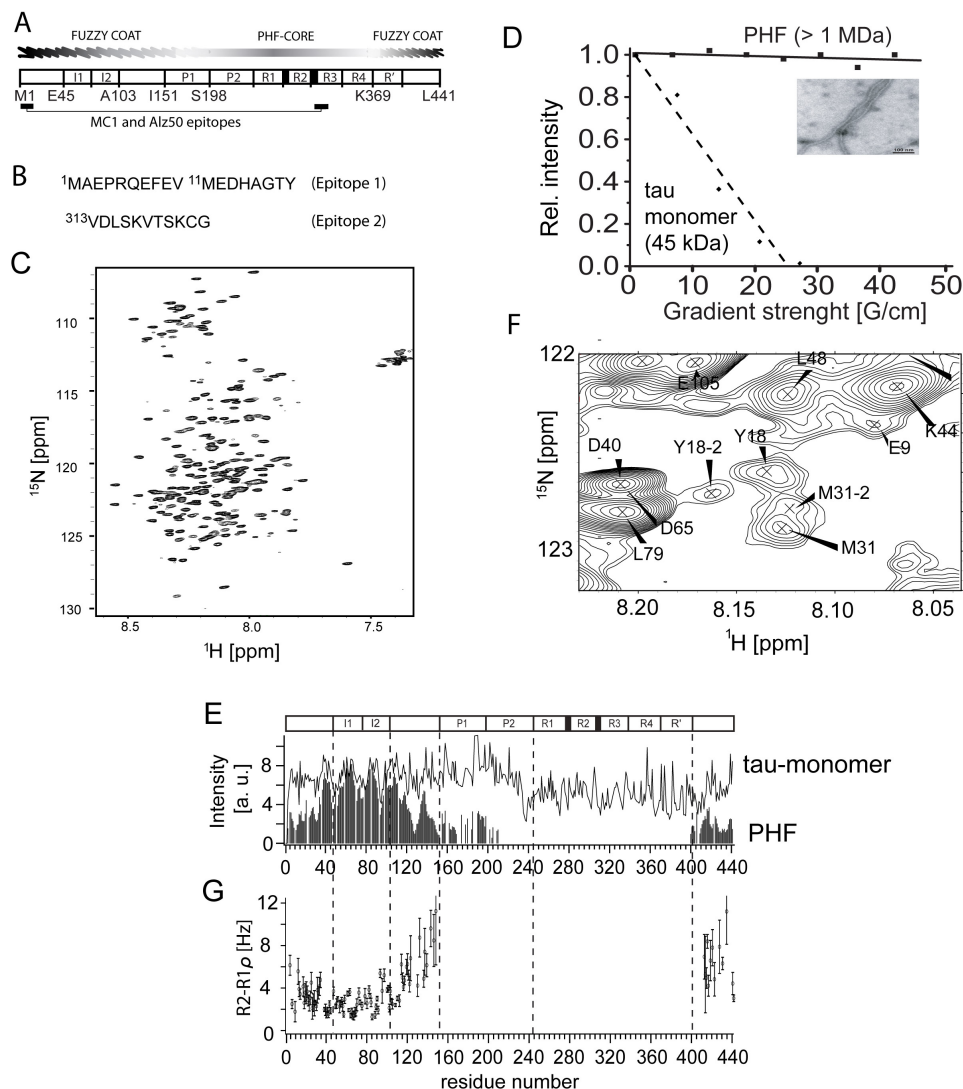


Figure 4.15: HRMAS of PHFs from httau40. A) The domain organization is shown, as well as the discontinuous epitopes for Alz50 and MC1. On top, the fuzzy coats and the PHF-core are indicated. B) Primary sequence of the two discontinuous epitopes. C) HSQC spectrum from PHF-tau, recorded with an HRMAS probehead. D) Diffusion experiments of monomeric tau (dashed line) and PHF-tau (solid line) with an electron micrograph from PHFs used. E) Selected region of the HSQC demonstrating multiple conformations in PHF-tau for Y18 and M31. F) Comparison of the absolute signal intensities of monomeric tau (line) and PHF-tau (bars). G) Difference of R2 and R1ρ, indicating conformational exchange.

molecular mass of more than one megadalton (Figure 4.15D). The enormous molecular weight of PHFs lead to severe signal broadening that makes it difficult to characterize them by solution-state NMR spectroscopy. We used solution NMR in combination with high-resolution magic angle spinning [190] to remove variations in magnetic susceptibility and detect residues which remain flexible in PHF-tau. Spinning htau40 filaments at the magic angle with a frequency of 6 kHz resulted in a drastic improvement in the NMR line-width consistent with previous studies [216], pointing to the presence of dynamics over a broad spectrum of time scales. About 260 signals were observed in a two-dimensional HRMAS heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N -labeled PHFs (Figure 4.15C). Sequence-specific resonance assignment of 244 of these signals identified most of the residues in the N-terminal domain up to T212 and at the C-terminus starting at V399 in agreement with previous studies [214,216]. No signals were detected for residues in between T212 and V399, suggesting that residues in the central domain are too immobile to be detected by liquid-state NMR spectroscopy. The absence of signals in the central domain proves that the concentration of residual or dynamically exchanging monomeric protein is below $5\ \mu\text{M}$ and does not affect the subsequent analysis. Analysis of the intensity of NMR signals of PHF tau revealed strong variations along its sequence (Figure 4.15F); for residues $^1\text{M-G}^{37}$ the intensity of the $^1\text{H}^N$ - ^{15}N cross-peaks was about two- to three-fold lower than for residues in the inserts that are subject to alternative splicing. After the inserts intensities strongly dropped for $^{121}\text{H-K}^{130}$, to shortly rise again and finally drop and remain at low values comparable to those observed at the N-terminus. In addition, all residues within the proline-rich region and the 40 C-terminal residues showed low signal intensities. Highly similar intensity profiles were observed in three separately aggregated htau40 PHF samples, emphasizing a high morphological similarity of the fuzzy coat of PHFs aggregated under identical conditions.

Comparison with monomeric htau40 revealed that the NMR signal intensities of many residues were strongly reduced in filamentous htau40. Most strikingly, $^{121}\text{H-K}^{130}$ and $^1\text{M-G}^{37}$, which are separated from the fibrillar core by 170 residues or more showed both changes in position and intensity of NMR signals (Figure 4.15F). The presence of chemical exchange in this region was further supported by ^{15}N spin relaxation measurements (Figure 4.15G). In agreement with chemical exchange, additional peaks were observed in close proximity to several of these residues (Figure 4.15E and 4.16J). The

additional signals could not be connected in triple-resonance experiments or using exchange spectroscopy due to low signal-to-noise and signal overlap, but chemical shift similarity and paramagnetic relaxation enhancement allowed their assignment to residues at the N- and C-terminus. In contrast, no peak doubling was observed at the N- and C-terminus in monomeric tau, highlighting the specificity of the multiple conformations in PHF-tau.

4.4.3 Transient long-range interactions of the fuzzy coat

We revealed the identity of the PHF-specific conformations through PRE measurements [170](Figure 4.16). The PRE effect scales as the inverse sixth power of the distance between the unpaired electron of the nitroxide and the NMR spin, providing a powerful probe of structure (see chapter 2). ^{15}N spin relaxation times indicate that the fuzzy coat of PHFs is highly dynamic on a broad scale suggesting that the correlation time of the electron-amide proton internuclear vector is comparable to that of small water soluble proteins. Initially, we measured PRE broadening for PHFs with a nitroxide attached to position 15 (Figure 4.16A-C). 40 residues at the N-terminus were broadened with PRE intensity ratios below 0.6. In addition, NMR revealed that $^{119}\text{A-D}^{133}$, the proline-rich region (residues $^{151}\text{I-T}^{212}$) and the C-terminal fuzzy coat ($^{399}\text{V-A}^{429}$) are in transient contact with the N-terminus of PHF-tau. The similarity of the PRE intensity ratio profile (Figure 4.16B) with the pattern of NMR signal intensities of PHF tau (Figure 4.15F) provides strong evidence that the PHF-specific conformations are caused by formation of transient long-range interactions. The global folding of the fuzzy coat of PHFs was further supported by paramagnetic effects observed for a nitroxide radical at position 125 (Figure 4.16D-F). Importantly, experiments using an equimolar mixture of nitroxide-tagged ^{14}N PHF-tau with ^{15}N -labeled cysteine-free PHF-tau proves that the observed paramagnetic effects are mostly of intramolecular nature (Figure 4.16K). To directly probe the interaction between the PHF core structure and the fuzzy coat, we attached the nitroxide spin label to the native cysteine C322 of PHF tau (Figure 4.16G-I). C322 is near the second epitope of the conformation-specific antibodies Alz50 and MC1 and is located in the protease-resistant PHF core [221, 271]. Attachment of the nitroxide label to C322 caused signal broadening in residue stretches close to Q124, a region that transiently populates helical conformations (Figure 4.17), A152 and $^{167}\text{N-T}^{212}$, as well as $^{409}\text{S-A}^{426}$ at the C-terminus (Figure

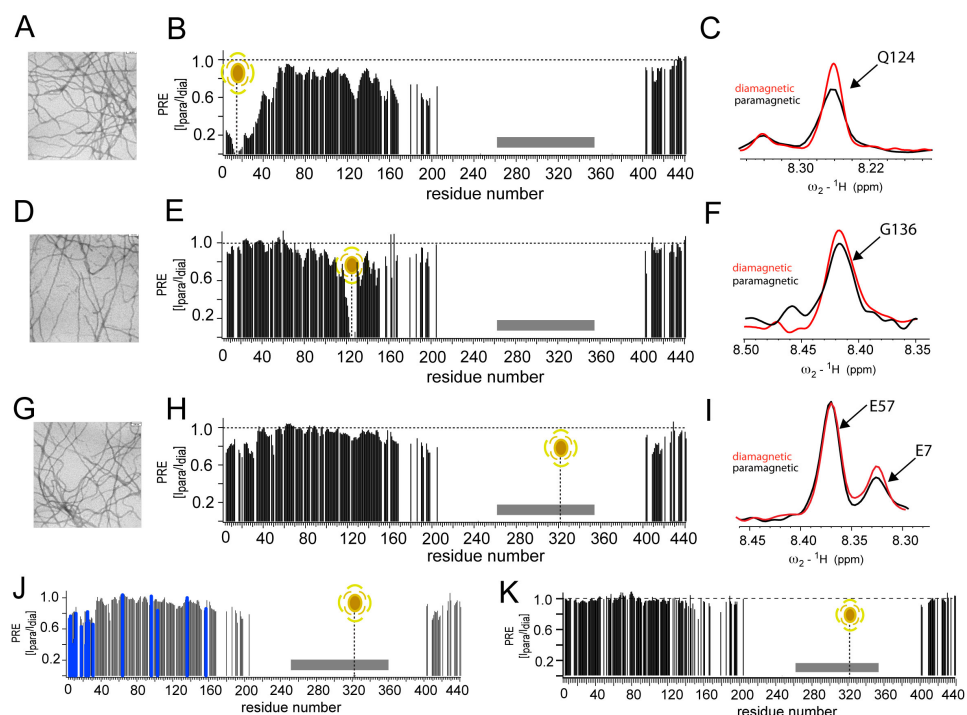


Figure 4.16: PRE experiments of PHF-tau. A, D, G) Electron micrographs of the different tau-mutants with MTSL attached to C15, C125, and C322, respectively. B) PRE intensity profile of PHF-tau with MTSL attached to residue 15. C) The 1D-trace of Q124 additionally confirms the transient contacts of the N-terminus and the region $^{119}\text{A-D}^{133}$. E) PRE intensity profile of PHF-tau with MTSL@C125. F) 1D-trace of G136. H) PRE intensity profile with MTSL@C322. I) 1D-trace of residues E7 and E57. J) Blue bars indicate the residues with a second peak set. K) PRE broadening profile of the equimolar mixture of nitroxide-tagged ^{14}N PHF-tau with ^{15}N -labeled cysteine-free PHF-tau.

4.16H and I). Thus, residues in the projection domain and at the C-terminus directly contact the PHF-core residue C322, consistent with partial protection of the C-terminus to proteolysis in PHF tau [273]. In addition, the ability to quench the nitroxide radical by a reducing agent indicates that the side chain of C322 is not fully buried in the fibrillar core but accessible to the solvent in agreement with the ability of Alz50 and MC1 to recognize the 313-322 epitope. The most pronounced paramagnetic effect, however, was observed for the first 30 residues at the N-terminus using either the major or the minor peak set observed for the N-terminal residues. Taken together NMR spectroscopy revealed a dense network of long-range interactions be-

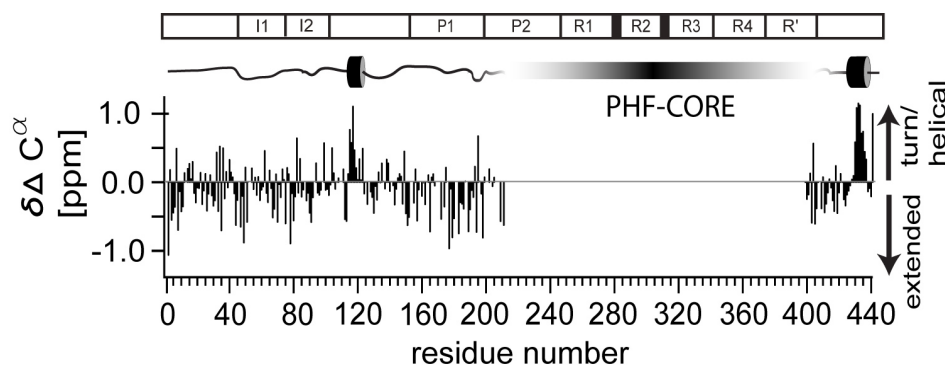


Figure 4.17: Secondary structure propensities in PHF-tau. C^α secondary structure propensities were investigated by using a 3D HNCA experiment. We found two α -helical propensities for $^{114}\text{LEDEAAGHVT}^{123}$ and $^{428}\text{LADEVASLA}^{437}$ that coincide with the two α -helical stretches found in monomeric htau40.

tween regions of PHF tau that are conserved in all six isoforms of tau, in particular between the two discontinuous epitopes (residues 1-18 and 313-322) that are required for recognition of pathological tau by the Alz50 and MC1 antibody.

4.4.4 Nature of the long-range interactions

To obtain insight into the mechanism of formation of long-range interactions in PHF tau, we performed NMR measurements of PHF tau at high ionic strength and of a mutant version of PHF tau in which F8 and V10 were replaced by serines. Mutation of F8 and V10, two hydrophobic residues in the N-terminal epitope of Alz50 and MC1, did not affect chemical exchange broadening in PHF tau (Figure 4.18A). In contrast, at high ionic strength the intensity profile of PHF tau was very similar to the profile observed for monomeric tau (Figure 4.18B). Only in regions neighboring the fibril core, that is residues 170-212 and 399-441, the intensity in the fuzzy coat remained low, most likely due to restricted motion inferred by the nearby fibrillar core. The strong impact of ionic strength demonstrates that electrostatic interactions are important for formation of the network of intramolecular long-range interactions in PHF tau. The monoclonal antibodies Alz50 and MC1 specifically recognize pathological tau in an aggregated state. Therefore we asked how the network of long-range interactions is changed when tau assembles

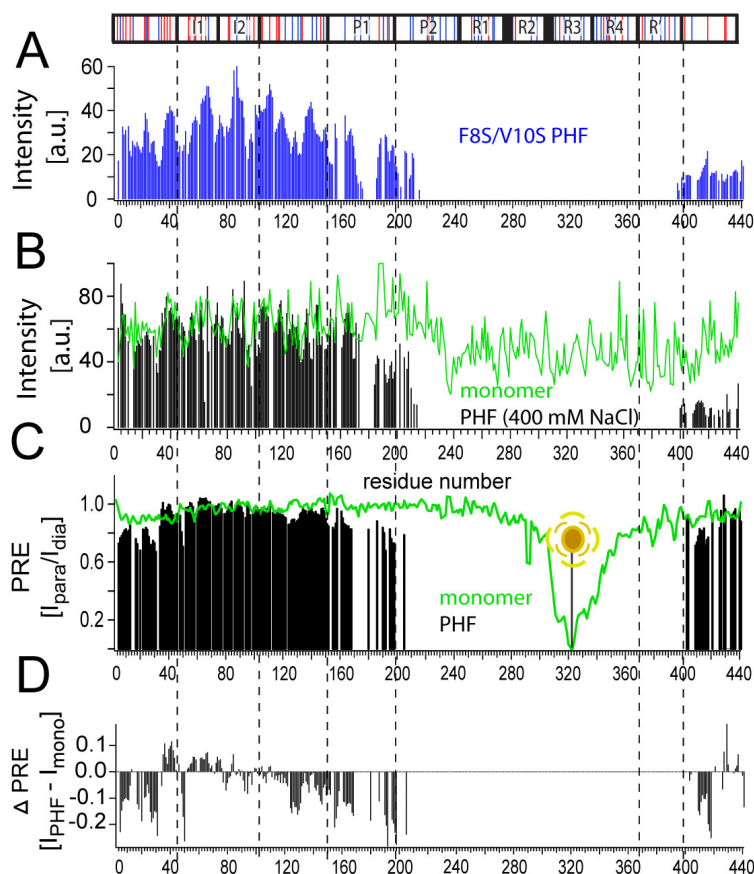


Figure 4.18: Nature of the long-range interactions. A) The mutation of two the hydrophobic F8 and V10 to serines resulted in the same intensity pattern found for non-mutated PHF-tau, excluding hydrophobic interactions responsible for the low signal intensities at the N-terminus. B) The addition of 400 mM NaCl resulted in an intensity pattern for PHF-tau similar to monomeric tau, revealing a tight network of electrostatic interactions. C) Comparison of the PRE intensity profiles from monomeric tau (green line) and PHF-tau (black bars) with MTSL@C322. D) The difference plot of the PRE profiles from C) shows enhanced broadening for PHF-tau for the residue stretches $^1M-G^{37}$, $^{119}A-D^{133}$, $^{151}I-T^{212}$ and $^{399}V-A^{429}$.

into PHFs. In monomeric htau40, the paramagnetic effects were mostly restricted to the immediate vicinity of the nitroxide label at Cys322 (Figure

4.18C). Only for the 30 N-terminal residues NMR signals were slightly broadened in agreement with a weak transient interaction between the N-terminus and repeat three in monomeric tau [49, 218]. In contrast, in PHF tau long-range contacts were observed with the proline-rich domain, the region next to M127 and the N- and C-terminus (Figure 4.18C and D). In addition, the paramagnetic effects induced at the N-terminus were stronger in PHF tau than in monomeric tau (Figure 4.19A) demonstrating that the interaction between the two discontinuous epitopes of the monoclonal Alz50 and MC1 antibodies is tightened in PHF tau.

4.4.5 Discussion

Identification of neurons that appear morphologically normal, but are undergoing pathological changes, is an important step in determining the earliest AD-specific modifications of tau. As the major constituent of PHFs in AD is hyperphosphorylated tau, phosphorylation-dependent tau-antibodies were developed [266]. Many phosphorylation-dependent antibodies, however, have epitopes that are also present in biopsy brain tissue from non-demented patients [274], and thus do not appear to represent post-translational changes necessary for tau aggregation. In contrast, antibodies that detect conformational changes in tau are specific for tau aggregation and can detect the earliest alterations of tau in AD [221, 269]. Two very useful conformation-dependent tau-antibodies are Alz50 and MC1, recognizing pathological tau conformations before the manifestation of PHFs and independent of phosphorylation by using two discontinuous epitopes [221, 271]. The molecular mechanism of recognition, however, is unknown due to the lack of structural information of pathological tau. Determining the molecular structure of tau filaments poses a significant challenge in view of their insolubility and enormous size. Previous solution-state NMR studies grappled with low resolutions and incomplete assignment [215, 216]. Here we used HRMAS solution-state NMR on full-length httau40, the longest tau isoform with 441 residues, and assigned 244 residues of PHF-tau with a molecular mass of more than one megadalton (Figure 4.15). We found that $^1\text{M-T}^{212}$ and $^{399}\text{V-L}^{441}$ remain highly dynamic on a broad time scale giving rise to the so-called fuzzy coat observed in electron micrographs of PHF-tau. Residues $^{213}\text{P-V}^{398}$ were not observed by solution NMR due to slow tumbling of the fibrillar core in agreement with previous studies [111, 216]. On the other hand, solid-state

NMR spectroscopy of a 99-residue fragment comprising three repeats of tau indicated that the rigid PHF-core ranges from residues 262 to 354 [214]. Together the data suggest the following structural organization of PHF-tau: (i) a N-terminal highly dynamic fuzzy coat ranging from M1 to roughly T212, (ii) a semi-rigid linker of about 50 residues from P213 to G261, (iii) a rigid fibril-core ranging from S262 to I354, (iv) a second semi-rigid linker of about 50 residues comprising S355 to V398 and (v) the mobile C-terminal fuzzy coat from V399 to L441. Paramagnetic NMR at single residue resolution revealed that aggregation of tau into paired helical filaments results in the formation of a dense network of intramolecular long-range interactions (Figure 4.16). Residue stretches $^{1}\text{M-G}^{37}$, $^{119}\text{A-D}^{133}$, $^{151}\text{I-T}^{212}$ and $^{399}\text{V-A}^{429}$, regions that are conserved in all six isoforms of the tau molecule, are in direct contact in PHF-tau. In contrast, residues in the inserts of tau that are subject to alternative splicing contribute only little to the network of long-range interactions (Figure 4.16). Importantly both the N-terminal and C-terminal parts of the fuzzy coat interact with the surface of the fibrillar core consistent with partial protection of the C-terminus to proteolysis in PHF-tau [273]. Further support for the intricate network of long-range interactions revealed by paramagnetic NMR (Figure 4.16) is provided by antibody studies: (i) the reactivity of Alz50 and MC1 was attenuated upon deletion of residues 46-241 in tau, which are between the C- and N-terminal portions of the Alz50 and MC1 epitopes [221]; (ii) residues 155-244 and 305-314 are required for recognition of PHF tau by the conformation-specific antibody Tau66 [275]. In addition, the observed long-range contacts are mostly intramolecular (Figure 4.16K) despite the high local concentration of tau in PHFs, in agreement with failed attempts at creating the Alz50 and MC1 epitopes intermolecularly by combining complementary NH₂- and COOH-terminal deletion mutants [221,271].

Electrostatic interactions are important for formation of transient long-range interactions in PHF-tau (Figure 4.18) in agreement with the distribution of charges in tau: a predominantly positively charged repeat region and a negatively charged N-terminus (Figure 4.15B). In monomeric tau the repeat region exchanges between many different conformations without a well-defined distribution of charges. In contrast, in the rigid cross- β -structure of PHF-tau hydrophobic residues are expected to form the core while it is energetically more favorable for charged residues to lie on the surface. Thus, aggregation of tau will be connected to formation of positive patches in the repeat region, resulting in enhanced interaction with the negatively charged

N-terminal tail. In addition, a connection between formation of long-range interactions and secondary structure is suggested by the partial overlap of ¹¹⁹A-D¹³³, residues that are strongly involved in transient long-range interactions in PHF tau (Figure 4.16), with the region that transiently populates α -helical structure (¹¹⁴LEDEAAGHVT¹²³) (Figure 4.17).

There is increasing evidence that soluble oligomeric tau species, rather than fibrillar tau, may be the critical toxic moiety underlying neurodegeneration [276]. Biochemical studies using recombinant tau have demonstrated a clear selectivity of Alz50 and MC1 for PHF tau: the interaction of Alz50 with PHF-tau is nearly two orders of magnitude greater in affinity than its interaction with recombinant monomeric tau (Figure 4.19B) [271]. However, MC1 and Alz50 also bind to a non-filamentous, soluble pool of abnormal tau that is able to self-assemble into PHFs in a concentration-dependent manner [270], suggesting that both PHFs and soluble oligomers of tau are recognized by Alz50 and MC1. Our study reveals that the two epitopes of the Alz50 and MC1 antibody are already weakly in contact in monomeric tau, but the interaction is tightened in PHF tau (Figure 4.19A and B). This provides an explanation for the recognition of oligomeric tau before the formation of PHFs: a collapse of the hydrophobic repeat region in tau oligomers will lead to preferential distribution of hydrophobic and charged residues in the interior and on the surface of the collapsed structure, respectively, resulting in an electrostatic strengthening of the interaction between the negatively charged N-terminal tail and the positively charged repeat region. Later in aggregation when PHFs have formed a similar charge distribution will be present and both early and late forms of pathological tau would be recognized. Binding of the Alz50 and MC1 antibody will depend on the sequence composition of the two epitopes [277], but the increased affinity of the two antibodies toward PHF tau is due to a tightening of the interaction during aggregation (Figure 4.19B). Antibody binding will stabilize the closed conformation. In summary, our study demonstrates that the core structure of a filament or an amyloid may have many capacities for interactions with otherwise unstructured segments within the protein (Figure 4.19C). It rationalizes the conformation-specific antibodies of tau and highlights the heterogeneity within aggregate structures.

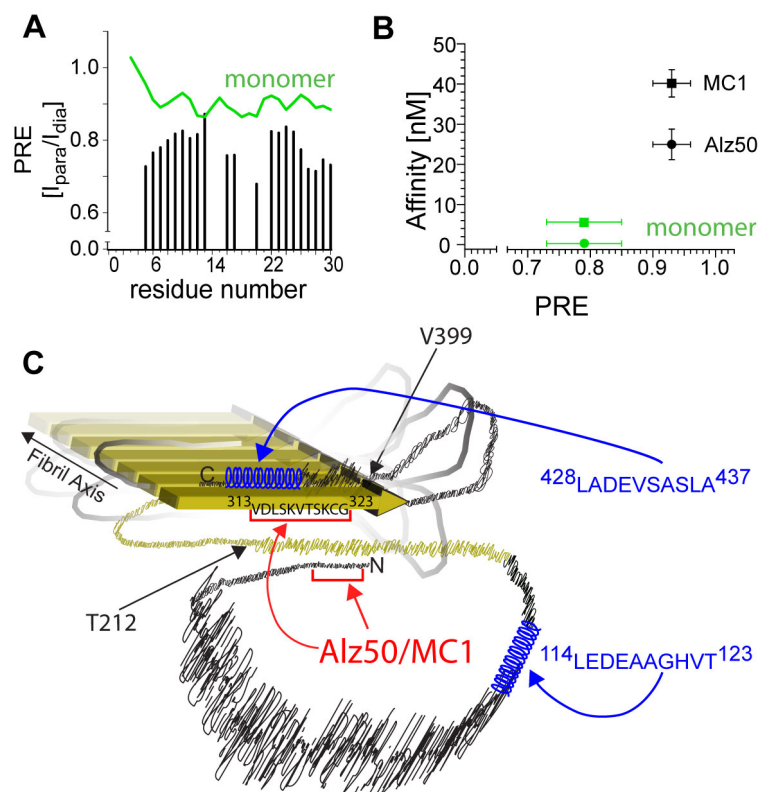


Figure 4.19: Transient contacts in PHF-tau, important for the Alz50 antibody. A) Comparison of the PRE intensity profile of monomeric tau (green line) and PHF tau (bars), when a nitroxide spin label was attached to the native C322. At the N-terminus, PRE intensity ratios are lower in PHF tau than in monomeric tau revealing a tightening of long-range contacts between the two epitopes of the Alz50 and MC1 antibody. B) Correlation between the affinity of the Alz50 and MC1 antibody [271] and the strength of interaction between their two discontinuous epitopes as quantified by the average PRE broadening of residues 1-18 when the nitroxide label was attached to the native C322. The interaction of the Alz50 antibody with PHF tau is nearly two orders of magnitude greater than its interaction with monomeric tau [271]. C) Model of the network of long-range interactions specific for PHF tau. Regions with transient helical structure are shown in blue and the fibrillar core identified by solid-state NMR [214] in yellow. The amplitude of dynamics in different parts of the fuzzy coat is indicated by the width of the protein chain. Spatial proximity between different parts of the chain indicates tertiary interactions that were revealed by paramagnetic NMR. The two discontinuous epitopes of the conformation-specific antibodies Alz50 and MC1, residues 1-18 and 313-322, are highlighted in red.

Bibliography

- [1] Mulder, G. J. (1839) *Journal fuer praktische Chemie* **16(1)**, 129–152.
- [2] Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) *Nature* **181(4610)**, 662–666.
- [3] Alberts, B. (2008) *Molecular Biology of the Cell* 5th ed, Garland Science, .
- [4] Fulton, A. B. and Isaacs, W. B. (1991) *Bioessays* **13(4)**, 157–161.
- [5] Xinsheng, L., Ke, F., and Wei, W. (2004) *Proteins* **54(3)**, 491–499.
- [6] Uversky, V. N. (2002) *Protein Sci* **11(4)**, 739.
- [7] Tompa, P. (2009) *Structure and function of intrinsically disordered proteins*, Chapman & Hall/CRC, .
- [8] Anfinsen, C. et al. (1973) *Science* **181(96)**, 223–230.
- [9] Levinthal, C. (1969) *Mossbauer Spectroscopy in Biological Systems* **3**, 22–24.
- [10] Levinthal, C. (1968) *Journal de Chimie Physique* **65(1)**, 44–45.
- [11] Dill, K. and Chan, H. (1997) *Nat Struct Biol* **4(1)**, 10–19.
- [12] Fischer, E. (1894) *Berichte der deutschen chemischen Gesellschaft* **27(3)**, 2985–2993.
- [13] Karush, F. (1950) *J Am Chem Soc* **72(6)**, 2705–2713.
- [14] Koshland, D. (1958) *Proc Natl Acad Sci U S A* **44(2)**, 98.

- [15] Csermely, P., Palotai, R., and Nussinov, R. (2010) *Trends Biochem Sci* **35**, 539–46.
- [16] Boehr, D., Nussinov, R., and Wright, P. (2009) *Nat Chem Biol* **5(11)**, 789–796.
- [17] James, L., Roversi, P., and Tawfik, D. (2003) *Science* **299(5611)**, 1362.
- [18] Lange, O., Lakomek, N., Fares, C., Schroder, G., Walter, K., Becker, S., Meiler, J., Grubmuller, H., Griesinger, C., and deGroot, B. (2008) *Science* **320(5882)**, 1471.
- [19] Ma, B., Kumar, S., Tsai, C., and Nussinov, R. (1999) *Protein Eng Des Sel* **12(9)**, 713.
- [20] Mittermaier, A. and Kay, L. (2006) *Science* **312(5771)**, 224.
- [21] Henzler-Wildman, K., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., et al. (2007) *Nature* **450(7171)**, 838–844.
- [22] Korzhnev, D. and Kay, L. (2008) *Acc Chem Res* **41(3)**, 442–451.
- [23] Mittag, T., Kay, L., and Forman-Kay, J. (2010) *J Mol Recognit* **23(2)**, 105–116.
- [24] Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., and Romero, P. (2001) *J Mol Graph Model* **19(1)**, 26–59.
- [25] Dyson, H. J. and Wright, P. E. (2005) *Nat Rev Mol Cell Biol* **6(3)**, 197–208.
- [26] Tompa, P. (2002) *Trends Biochem Sci* **27(10)**, 527–533.
- [27] Dunker, A., Oldfield, C., Meng, J., Romero, P., Yang, J., Chen, J., Vacic, V., Obradovic, Z., and Uversky, V. (2008) *BMC Genomics* **9(2)**, S1.
- [28] Minor, D. and Kim, P. (1994) *Nature* **367(6464)**, 660–663.
- [29] Jiang, Z., Zhang, L., Chen, J., Xia, A., and Zhao, D. (2002) *Polymer* **43(22)**, 6037–6047.

- [30] Garbuzynskiy, S., Lobanov, M., and Galzitskaya, O. (2004) *Protein Sci* **13(11)**, 2871–2877.
- [31] Pontius, B. (1993) *Trends Biochem Sci* **18(5)**, 181–186.
- [32] Brown, H. and Hoh, J. (1997) *Biochemistry* **36(49)**, 15035–15040.
- [33] Chen, J., Kanai, Y., Cowan, N., and Hirokawa, N. (1992) *Nature* **360(6405)**, 674–677.
- [34] Alber, F., Dokudovskaya, S., Veenhoff, L., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B., et al. (2007) *Nature* **450(7170)**, 695–701.
- [35] Strambio-De-Castillia, C., Niepel, M., and Rout, M. P. (2010) *Nat Rev Mol Cell Biol* **11(7)**, 490–501.
- [36] Lim, R., Huang, N., Koeser, J., Deng, J., Lau, K., Schwarz-Herion, K., Fahrenkrog, B., and Aepli, U. (2006) *Proc Natl Acad Sci U S A* **103(25)**, 9512.
- [37] Podlaha, O. and Zhang, J. (2003) *Proc Natl Acad Sci U S A* **100(21)**, 12241.
- [38] Pometun, M., Chekmenev, E., and Wittebort, R. (2004) *J Biol Chem* **279(9)**, 7982.
- [39] Uversky, V., Oldfield, C., and Dunker, A. (2008) *Annu Rev Biophys* **37**, 215–46.
- [40] Dobson, C. (2003) *Nature* **426(6968)**, 884–890.
- [41] Pechmann, S., Levy, E., Tartaglia, G., and Vendruscolo, M. (2009) *Proc Natl Acad Sci U S A* **106(25)**, 10159.
- [42] Mukrasch, M., Biernat, J., vonBergen, M., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2005) *J Biol Chem* **280(26)**, 24978.
- [43] Dehmelt, L. and Halpain, S. (2005) *Genome Biol* **6(1)**, 204.
- [44] Goedert, M., Spillantini, M., Jakes, R., Rutherford, D., and Crowther, R. (1989) *Neuron* **3(4)**, 519–526.

- [45] Lee, G., Cowan, N., and Kirschner, M. (1988) *Science* **239(4837)**, 285–285.
- [46] Cleveland, D. W., Hwo, S. Y., and Kirschner, M. W. (1977) *J Mol Biol* **116(2)**, 227–47.
- [47] Schweers, O., Schönbrunn-Hanebeck, E., Marx, A., and Mandelkow, E. (1994) *J Biol Chem* **269(39)**, 24290–7.
- [48] Lee, V. M., Goedert, M., and Trojanowski, J. Q. (2001) *Annu Rev Neurosci* **24**, 1121–59.
- [49] Jeganathan, S., Bergen, M., Brutzl, H., Steinhoff, H.-J., and Mandelkow, E. (2006) *Biochemistry* **45(7)**, 2283–93.
- [50] Dotti, C. G. and Poo, M.-m. (2003) *Nat Cell Biol* **5(7)**, 591–4.
- [51] Sotelo-Silveira, J. R., Calliari, A., Kun, A., Koenig, E., and Sotelo, J. R. (2006) *Traffic* **7(5)**, 508–15.
- [52] Thies, E. and Mandelkow, E.-M. (2007) *J Neurosci* **27(11)**, 2896–907.
- [53] Hirokawa, N., Funakoshi, T., Sato-Harada, R., and Kanai, Y. (1996) *J Cell Biol* **132(4)**, 667–79.
- [54] Aronov, S., Aranda, G., Behar, L., and Ginzburg, I. (2002) *J Cell Sci* **115(19)**, 3817–27.
- [55] Aronov, S., Aranda, G., Behar, L., and Ginzburg, I. (2001) *J Neurosci* **21(17)**, 6577–87.
- [56] Utton, M., Connell, J., Asuni, A., vanSlegtenhorst, M., Hutton, M., deSilva, R., Lees, A., Miller, C., and Anderton, B. (2002) *J Neurosci* **22(15)**, 6394.
- [57] Weingarten, M., Lockwood, A., Hwo, S., and Kirschner, M. (1975) *Proc Natl Acad Sci U S A* **72(5)**, 1858.
- [58] Amos, L. A. (2004) *Org Biomol Chem* **2(15)**, 2153–60.
- [59] Desai, A. and Mitchison, T. J. (1997) *Annu Rev Cell Dev Biol* **13**, 83–117.

- [60] Preuss, U., Biernat, J., Mandelkow, E., and Mandelkow, E. (1997) *J Cell Sci* **110**(6), 789.
- [61] Mukrasch, M., vonBergen, M., Biernat, J., Fischer, D., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2007) *J Biol Chem* **282**(16), 12230.
- [62] Goode, B. L., Denis, P. E., Panda, D., Radeke, M. J., Miller, H. P., Wilson, L., and Feinstein, S. C. (1997) *Mol Biol Cell* **8**(2), 353–65.
- [63] Eliezer, D., Barré, P., Kobaslija, M., Chan, D., Li, X., and Heend, L. (2005) *Biochemistry* **44**(3), 1026–1036.
- [64] Fischer, D., Mukrasch, M. D., vonBergen, M., Klos-Witkowska, A., Biernat, J., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2007) *Biochemistry* **46**(10), 2574–82.
- [65] Sillen, A., Barbier, P., Landrieu, I., Lefebvre, S., Wieruszeski, J.-M., Leroy, A., Peyrot, V., and Lippens, G. (2007) *Biochemistry* **46**(11), 3055–64.
- [66] Al-Bassam, J., Ozer, R. S., Safer, D., Halpain, S., and Milligan, R. A. (2002) *J Cell Biol* **157**(7), 1187–96.
- [67] Chau, M. F., Radeke, M. J., deInés, C., Barasoain, I., Kohlstaedt, L. A., and Feinstein, S. C. (1998) *Biochemistry* **37**(51), 17692–703.
- [68] Kar, S., Fan, J., Smith, M. J., Goedert, M., and Amos, L. A. (2003) *EMBO J* **22**(1), 70–7.
- [69] Santarella, R. A., Skiniotis, G., Goldie, K. N., Tittmann, P., Gross, H., Mandelkow, E.-M., Mandelkow, E., and Hoenger, A. (2004) *J Mol Biol* **339**(3), 539–53.
- [70] Hirokawa, N., Shiomura, Y., and Okabe, S. (1988) *J Cell Biol* **107**(4), 1449–59.
- [71] Makrides, V., Massie, M. R., Feinstein, S. C., and Lew, J. (2004) *Proc Natl Acad Sci U S A* **101**(17), 6746–51.
- [72] Gong, C.-X., Liu, F., and Iqbal, K. (2005) *J Neural Transm* **112**, 813–838.

- [73] Mawal-Dewan, M., Henley, J., Van deVoorde, A., Trojanowski, J. Q., and Lee, V. M. (1994) *J Biol Chem* **269**(49), 30981–7.
- [74] Stoothoff, W. H. and Johnson, G. V. W. (2005) *Biochim Biophys Acta* **1739**(2-3), 280–97.
- [75] Jucker, M. (2006) *Alzheimer: 100 Years and Beyond*, Springer Verlag Berlin Heidelberg, .
- [76] Gong, C.-X. and Iqbal, K. (2008) *Curr Med Chem* **15**(23), 2321–8.
- [77] Zheng-Fischhöfer, Q., Biernat, J., Mandelkow, E. M., Illenberger, S., Godemann, R., and Mandelkow, E. (1998) *Eur J Biochem* **252**(3), 542–52.
- [78] Johnson, G. V. W. and Stoothoff, W. H. (2004) *J Cell Sci* **117**(Pt 24), 5721–9.
- [79] Biernat, J., Wu, Y.-Z., Timm, T., Zheng-Fischhöfer, Q., Mandelkow, E., Meijer, L., and Mandelkow, E.-M. (2002) *Mol Biol Cell* **13**(11), 4013–28.
- [80] Litersky, J. M., Johnson, G. V., Jakes, R., Goedert, M., Lee, M., and Seubert, P. (1996) *Biochem J* **316**, 655–60.
- [81] Ahljianian, M. K., Barrezueta, N. X., Williams, R. D., Jakowski, A., Kowsz, K. P., McCarthy, S., Coskran, T., Carlo, A., Seymour, P. A., Burkhardt, J. E., Nelson, R. B., and McNeish, J. D. (2000) *Proc Natl Acad Sci U S A* **97**(6), 2910–5.
- [82] Brownlees, J., Irving, N. G., Brion, J. P., Gibb, B. J., Wagner, U., Woodgett, J., and Miller, C. C. (1997) *Neuroreport* **8**(15), 3251–5.
- [83] Harwood, A. J. (2001) *Cell* **105**(7), 821–4.
- [84] Cho, J.-H. and Johnson, G. V. W. (2003) *J Biol Chem* **278**(1), 187–93.
- [85] Godemann, R., Biernat, J., Mandelkow, E., and Mandelkow, E. M. (1999) *FEBS Lett* **454**(1-2), 157–64.
- [86] Liu, S. J., Zhang, J. Y., Li, H. L., Fang, Z. Y., Wang, Q., Deng, H. M., Gong, C. X., Grundke-Iqbal, I., Iqbal, K., and Wang, J. Z. (2004) *J Biol Chem* **279**(48), 50078–88.

- [87] Liu, F., Liang, Z., Shi, J., Yin, D., El-Akkad, E., Grundke-Iqbal, I., Iqbal, K., and Gong, C.-X. (2006) *FEBS Lett* **580(26)**, 6269–74.
- [88] Bramblett, G. T., Goedert, M., Jakes, R., Merrick, S. E., Trojanowski, J. Q., and Lee, V. M. (1993) *Neuron* **10(6)**, 1089–99.
- [89] Goedert, M., Jakes, R., Crowther, R. A., Six, J., Lübke, U., Vandermeeren, M., Cras, P., Trojanowski, J. Q., and Lee, V. M. (1993) *Proc Natl Acad Sci U S A* **90(11)**, 5066–70.
- [90] Schneider, A., Biernat, J., vonBergen, M., Mandelkow, E., and Mandelkow, E. M. (1999) *Biochemistry* **38(12)**, 3549–58.
- [91] Biernat, J., Gustke, N., Drewes, G., Mandelkow, E. M., and Mandelkow, E. (1993) *Neuron* **11(1)**, 153–63.
- [92] Landrieu, I., Lacosse, L., Leroy, A., Wieruszeski, J.-M., Trivelli, X., Sillen, A., Sibille, N., Schwalbe, H., Saxena, K., Langer, T., and Lip-pens, G. (2006) *J Am Chem Soc* **128(11)**, 3575–83.
- [93] DiTella, M. C., Feiguin, F., Carri, N., Kosik, K. S., and Cáceres, A. (1996) *J Cell Sci* **109**, 467–77.
- [94] Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., Sato-Yoshitake, R., Takei, Y., Noda, T., and Hirokawa, N. (1994) *Nature* **369(6480)**, 488–91.
- [95] Takei, Y., Teng, J., Harada, A., and Hirokawa, N. (2000) *J Cell Biol* **150(5)**, 989–1000.
- [96] Andorfer, C. A. and Davies, P. (2000) *Dev Neurosci* **22(4)**, 303–9.
- [97] Peterson, D. W., Ando, D. M., Taketa, D. A., Zhou, H., Dahlquist, F. W., and Lew, J. (2010) *Proc Natl Acad Sci U S A* **107(7)**, 2884–9.
- [98] Iqbal, K., Liu, F., Gong, C.-X., Alonso, A. D. C., and Grundke-Iqbal, I. (2009) *Acta Neuropathol* **118(1)**, 53–69.
- [99] Su, B., Wang, X., Drew, K. L., Perry, G., Smith, M. A., and Zhu, X. (2008) *J Neurochem* **105(6)**, 2098–108.

- [100] Arendt, T., Stieler, J., Strijkstra, A. M., Hut, R. A., Rüdiger, J., Van derZee, E. A., Harkany, T., Holzer, M., and Härtig, W. (2003) *J Neurosci* **23(18)**, 6972–81.
- [101] Spires-Jones, T. L., Stoothoff, W. H., deCalignon, A., Jones, P. B., and Hyman, B. T. (2009) *Trends Neurosci* **32(3)**, 150–9.
- [102] vonBergen, M., Barghorn, S., Biernat, J., Mandelkow, E.-M., and Mandelkow, E. (2005) *Biochim Biophys Acta* **1739(2-3)**, 158–66.
- [103] vonBergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E. M., and Mandelkow, E. (2000) *Proc Natl Acad Sci U S A* **97(10)**, 5129–34.
- [104] Goux, W. J., Kopplin, L., Nguyen, A. D., Leak, K., Rutkofsky, M., Shanmuganandam, V. D., Sharma, D., Inouye, H., and Kirschner, D. A. (2004) *J Biol Chem* **279(26)**, 26868–75.
- [105] Inouye, H., Sharma, D., Goux, W. J., and Kirschner, D. A. (2006) *Biophys J* **90(5)**, 1774–89.
- [106] vonBergen, M., Barghorn, S., Li, L., Marx, A., Biernat, J., Mandelkow, E. M., and Mandelkow, E. (2001) *J Biol Chem* **276(51)**, 48165–74.
- [107] Berriman, J., Serpell, L. C., Oberg, K. A., Fink, A. L., Goedert, M., and Crowther, R. A. (2003) *Proc Natl Acad Sci U S A* **100(15)**, 9034–8.
- [108] Goux, W. J. (2002) *Biochemistry* **41(46)**, 13798–806.
- [109] Sadqi, M., Hernández, F., Pan, U., Pérez, M., Schaeberle, M. D., Avila, J., and Muñoz, V. (2002) *Biochemistry* **41(22)**, 7150–5.
- [110] Iliev, A. I., Ganesan, S., Bunt, G., and Wouters, F. S. (2006) *J Biol Chem* **281(48)**, 37195–204.
- [111] Wischik, C. M., Novak, M., Edwards, P. C., Klug, A., Tichelaar, W., and Crowther, R. A. (1988) *Proc Natl Acad Sci U S A* **85(13)**, 4884–8.
- [112] Margittai, M. and Langen, R. (2004) *Proc Natl Acad Sci U S A* **101(28)**, 10278–83.

- [113] Abraha, A., Ghoshal, N., Gamblin, T. C., Cryns, V., Berry, R. W., Kuret, J., and Binder, L. I. (2000) *J Cell Sci* **113**, 3737–45.
- [114] Horowitz, P. M., LaPointe, N., Guillozet-Bongaarts, A. L., Berry, R. W., and Binder, L. I. (2006) *Biochemistry* **45(42)**, 12859–66.
- [115] Lapointe, N. E., Horowitz, P. M., Guillozet-Bongaarts, A. L., Silva, A., Andreadis, A., and Binder, L. I. (2009) *Biochemistry* **48(51)**, 12290–7.
- [116] LaPointe, N. E., Morfini, G., Pigino, G., Gaisina, I. N., Kozikowski, A. P., Binder, L. I., and Brady, S. T. (2009) *J Neurosci Res* **87(2)**, 440–51.
- [117] Mandelkow, E., vonBergen, M., Biernat, J., and Mandelkow, E.-M. (2007) *Brain Pathol* **17(1)**, 83–90.
- [118] Geschwind, D. H. (2003) *Neuron* **40(3)**, 457–60.
- [119] Konzack, S., Thies, E., Marx, A., Mandelkow, E.-M., and Mandelkow, E. (2007) *J Neurosci* **27(37)**, 9916–27.
- [120] Goedert, M., Jakes, R., Spillantini, M. G., Hasegawa, M., Smith, M. J., and Crowther, R. A. (1996) *Nature* **383(6600)**, 550–3.
- [121] Kampers, T., Friedhoff, P., Biernat, J., Mandelkow, E. M., and Mandelkow, E. (1996) *FEBS Lett* **399(3)**, 344–9.
- [122] Wilson, D. M. and Binder, L. I. (1997) *Am J Pathol* **150(6)**, 2181–95.
- [123] Alonso, A. d. C., Mederlyova, A., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2004) *J Biol Chem* **279(33)**, 34873–81.
- [124] Gamblin, T. C., Chen, F., Zambrano, A., Abraha, A., Lagalwar, S., Guillozet, A. L., Lu, M., Fu, Y., Garcia-Sierra, F., LaPointe, N., Miller, R., Berry, R. W., Binder, L. I., and Cryns, V. L. (2003) *Proc Natl Acad Sci U S A* **100(17)**, 10032–7.
- [125] Rapoport, M., Dawson, H., Binder, L., Vitek, M., and Ferreira, A. (2002) *Proc Natl Acad Sci U S A* **99(9)**, 6364.
- [126] Roberson, E., Scearce-Levie, K., Palop, J., Yan, F., Cheng, I., Wu, T., Gerstein, H., Yu, G., and Mucke, L. (2007) *Science* **316(5825)**, 750.

- [127] Braak, H. and Braak, E. (1988) *Neuropathol Appl Neurobiol* **14(1)**, 39–44.
- [128] Shahani, N., Subramaniam, S., Wolf, T., Tackenberg, C., and Brandt, R. (2006) *J Neurosci* **26(22)**, 6103–14.
- [129] Witte, H., Neukirchen, D., and Bradke, F. (2008) *J Cell Biol* **180(3)**, 619–32.
- [130] Götz, J., Probst, A., Spillantini, M., Schäffer, T., Jakes, R., Bürki, K., and Goedert, M. (1995) *EMBO J* **14(7)**, 1304.
- [131] Zempel, H., Thies, E., Mandelkow, E., and Mandelkow, E.-M. (2010) *J Neurosci* **30(36)**, 11938–50.
- [132] Oddo, S., Caccamo, A., Tran, L., Lambert, M., Glabe, C., Klein, W., and LaFerla, F. (2006) *J Biol Chem* **281(3)**, 1599.
- [133] Guo, J., Arai, T., Miklossy, J., and McGeer, P. (2006) *Proc Natl Acad Sci U S A* **103(6)**, 1953.
- [134] Ittner, L. M., Ke, Y. D., Delerue, F., Bi, M., Gladbach, A., vanEersel, J., Wölfing, H., Chieng, B. C., Christie, M. J., Napier, I. A., Eckert, A., Staufenbiel, M., Hardeman, E., and Götz, J. (2010) *Cell* **142(3)**, 387–97.
- [135] Ittner, L. and Götz, J. (2011) *Nat Rev Neurosci* **12(2)**, 6772.
- [136] Jeganathan, S., vonBergem, M., Mandelkow, E.-M., and Mandelkow, E. (2008) *Biochemistry* **47(40)**, 10526–39.
- [137] Syme, C., Blanch, E., Holt, C., Jakes, R., Goedert, M., Hecht, L., and Barron, L. (2002) *Eur J Biochem* **269(1)**, 148–156.
- [138] Munowitz, M. (1988) *Coherence and NMR*, Wiley New York, .
- [139] Haken, H. and Wolf, H. (1987) *Atom-und Quantenphysik*, Springer Berlin, 8th edition.
- [140] Cohen-Tannoudji, C. and Diu, B. (2007) *Quantenmechanik I*, Walter de Gruyter, 3rd edition.

- [141] Cavanagh, J. (2007) Protein NMR spectroscopy: principles and practice, Academic Press New York, 2nd edition.
- [142] Sorensen, O., Eich, G., Levitt, M., Bodenhausen, G., and Ernst, R. (1983) *Progr in NMR spec* **16(2)**, 163–192.
- [143] Kessler, H., Gehrke, M., and Griesinger, C. (1988) *Angew Chem Int Ed Engl* **27**, 490–536.
- [144] Keeler, J. (2005) Understanding NMR Spectroscopy, Wiley New York, .
- [145] Levitt, M. (2003) Spin dynamics: basics of nuclear magnetic resonance, Wiley New York, .
- [146] Bibow, S. Solution state nmr investigations of the amyloid beta-peptide Master's thesis Humboldt University, Berlin, Germany (2007).
- [147] Contreras, R. and Peralta, J. (2000) *Progr in NMR spec* **37(4)**, 321–425.
- [148] Vuister, G., Delaglio, F., and Bax, A. (1993) *J Biomol NMR* **3(1)**, 67–80.
- [149] Karplus, M. (1959) *J Chem Phys* **30**, 11–15.
- [150] Karplus, M. (1963) *J Am Chem Soc* **85(18)**, 2870–2871.
- [151] Pardi, A., Billeter, M., and Wuthrich, K. (1984) *J Mol Biol* **180**, 741–751.
- [152] Habeck, M., Rieping, W., and Nilges, M. (2005) *J Magn Reson* **177(1)**, 160–165.
- [153] Permi, P., Kilpelainen, I., Annala, A., and Heikkinen, S. (2000) *J Biomol NMR* **16(1)**, 29–37.
- [154] Mandal, P. and Majumdar, A. (2004) *Conc Magn Reson A* **20(1)**, 1–23.
- [155] Ponstingl, H. and Otting, G. (1998) *J Biomol NMR* **12(2)**, 319–324.

- [156] deDios, A., Pearson, J., and Oldfield, E. (1993) *J Am Chem Soc* **115**(21), 9768–9773.
- [157] Saito, H. (1986) *Magn Reson Chem* **24**(10), 835–852.
- [158] Spera, S. and Bax, A. (1991) *J Am Chem Soc* **113**, 5490–5492.
- [159] deDios, A. C., Pearson, J. G., and Oldfield, E. (1993) *Science* **260**(5113), 1491–6.
- [160] Mulder, F. A. A. and Filatov, M. (2010) *Chem Soc Rev* **39**(2), 578–90.
- [161] Schwarzingher, S., Kroon, G., Foss, T., Chung, J., Wright, P., and Dyson, H. (2001) *J Am Chem Soc* **123**(13), 2970–2978.
- [162] Wishart, D. S. and Sykes, B. D. (1994) *Methods Enzymol* **239**, 363–92.
- [163] Wishart, D., Bigam, C., Holm, A., Hodges, R., and Sykes, B. (1995) *J Biomol NMR* **5**(1), 67–81.
- [164] Johnson Jr, C. (1999) *Progr in NMR Spec* **34**(3), 203–256.
- [165] Stejskal, E. and Tanner, J. (1965) *J Chem Phys* **42**(1), 288.
- [166] Bertini, I., Luchinat, C., Parigi, G., and Pierattelli, R. (2005) *Chem-BioChem* **6**(9), 1536–1549.
- [167] Pintacuda, G., John, M., Su, X., and Otting, G. (2007) *Acc Chem Res* **40**(3), 206–212.
- [168] Battiste, J. and Wagner, G. (2000) *Biochemistry* **39**(18), 5355–5365.
- [169] Gillespie, J. and Shortle, D. (1997) *J Mol Biol* **268**(1), 158–169.
- [170] Gillespie, J. and Shortle, D. (1997) *J Mol Biol* **268**(1), 170–184.
- [171] Clore, G. and Iwahara, J. (2009) *Chem Rev* **109**(9), 4108–4139.
- [172] Solomon, I. (1955) *Phys Rev Lett* **99**(2), 559–565.
- [173] Daune, M. and Quant, S. (2000) *Molekulare biophysik*, Springer, .
- [174] Duer, M. (2004) *Introduction to solid-state NMR spectroscopy*, Wiley-Blackwell, .

- [175] Bax, A. (2003) In Neurath Award Lecture : .
- [176] Gayathri, C. and Bothner-By, P. (1982) *Chem Phys Letters* **87(2)**, 192–196.
- [177] Bax, A., Kontaxis, G., and Tjandra, N. (2001) *Methods Enzymol* **339**, 127–174.
- [178] Ottiger, M., Delaglio, F., and Bax, A. (1998) *J Magn Reson* **131(2)**, 373–378.
- [179] Hiller, S., Fiorito, F., Wuthrich, K., and Wider, G. (2005) *Proc Natl Acad Sci U S A* **102(31)**, 10876.
- [180] Hiller, S., Wasmer, C., Wider, G., and Wuthrich, K. (2007) *J Am Chem Soc* **129(35)**, 10823–10828.
- [181] Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) *Proc Natl Acad Sci U S A* **94**, 12–366.
- [182] Andrew, E., Bradbury, A., and Eades, R. (1958) *Nature* **182(1659)**, 12–13.
- [183] Lowe, I. (1959) *Phys Rev Lett* **2(7)**, 285–287.
- [184] Skora, L., Becker, S., and Zweckstetter, M. (2010) *ChemBioChem* **13(11)**, 1829–32.
- [185] Lindon, J., Beckonert, O., Holmes, E., and Nicholson, J. (2009) *Progr in NMR Spec* **55(1)**, 79–100.
- [186] Weybright, P., Millis, K., Campbell, N., Cory, D., and Singer, S. (1998) *Magn Reson Med* **39(3)**, 337–345.
- [187] Lippens, G., Bourdonneau, M., Dhalluin, C., Warrass, R., Richert, T., Seetharaman, C., Boutillon, C., and Piotto, M. (1999) *Curr Org Chem* **3**, 147–169.
- [188] Laws, D., Bitter, H., and Jerschow, A. (2002) *Angew Chem Int Ed Engl* **41(17)**, 3096–3129.
- [189] Kuchel, P., Chapman, B., Bubb, W., Hansen, P., Durrant, C., and Hertzberg, M. (2003) *Conc Magn Reson A* **18(1)**, 56–71.

- [190] Garroway, A. (1982) *J Magn Reson* **49(1)**, 168–171.
- [191] Barghorn, S., Biernat, J., and Mandelkow, E. (2005) *Methods Mol Biol* **299**, 35–51.
- [192] Mori, S., Abeygunawardana, C., Johnson, M., and Vanzijl, P. (1995) *J Magn Reson B* **108(1)**, 94–98.
- [193] Schwieters, C., Kuszewski, J., Tjandra, N., and Marius Clore, G. (2003) *J Magn Reson* **160(1)**, 65–73.
- [194] Iwahara, J., Schwieters, C., and Clore, G. (2004) *J Am Chem Soc* **126(18)**, 5879–5896.
- [195] Smallcombe, S., Patt, S., and Keifer, P. (1995) *J Magn Reson A* **117(2)**, 295–303.
- [196] Plaxco, K., Morton, C., Grimshaw, S., Jones, J., Pitkeathly, M., Campbell, I., and Dobson, C. (1997) *J Biomol NMR* **10(3)**, 221–230.
- [197] Mills, R. (1973) *J Phys Chem* **77(5)**, 685–688.
- [198] Chou, J., Baber, J., and Bax, A. (2004) *J Biomol NMR* **29(3)**, 299–308.
- [199] Wilkins, D., Grimshaw, S., Receveur, V., Dobson, C., Jones, J., and Smith, L. (1999) *Biochemistry* **38(50)**, 16424–16431.
- [200] Wu, D., Chen, A., and Johnson, C. (1995) *J Magn Reson A* **115(2)**, 260–264.
- [201] Jones, J., Wilkins, D., Smith, L., and Dobson, C. (1997) *J Biomol NMR* **10(2)**, 199–203.
- [202] Cantor, C. and Schimmel, P. (1980) *Biophysical Chemistry Part II: Techniques for the study of biological structure and function*, W. H. Freeman and Company, 1st edition.
- [203] Zheng, G., Stait-Gardner, T., Anil Kumar, P., Torres, A., and Price, W. (2008) *J Magn Reson* **191(1)**, 159–163.
- [204] Eberstein, W., Georgalis, Y., and Saenger, W. (1994) *J Cryst Growth* **143(1-2)**, 71–78.

- [205] Fadda, G. and Lairez, D. (2000) *EPL* **52**, 712.
- [206] Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J Biomol NMR* **6(3)**, 277–293.
- [207] Panchal, S., Bhavesh, N., and Hosur, R. (2001) *J Biomol NMR* **20(2)**, 135–147.
- [208] Zweckstetter, M. and Bax, A. (2001) *J Am Chem Soc* **123(38)**, 9490–9491.
- [209] Kay, L., Keifer, P., and Saarinen, T. (1992) *J Am Chem Soc* **114(26)**, 10663–10665.
- [210] Cavanagh, J. and Rance, M. (1993) *Ann Rep NMR Spect* **27**, 1–1.
- [211] Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J Magn Reson* **102**, 241–245.
- [212] Mukrasch, M., Markwick, P., Biernat, J., vonBergen, M., Bernadó, P., Griesinger, C., Mandelkow, E., Zweckstetter, M., and Blackledge, M. (2007) *J Am Chem Soc* **129(16)**, 5235–5243.
- [213] Barre, P. and Eliezer, D. (2006) *J Mol Biol* **362(2)**, 312–326.
- [214] Andronesi, O., Bergen, M., Biernat, J., Seidel, K., Griesinger, C., Mandelkow, E., and Baldus, M. (2008) *J Am Chem Soc* **130(18)**, 5922–5928.
- [215] Sillen, A., Leroy, A., Wieruszeski, J., Loyens, A., Beauvillain, J., Buée, L., Landrieu, I., and Lippens, G. (2005) *ChemBioChem* **6(10)**, 1849–1856.
- [216] Sillen, A., Wieruszeski, J., Leroy, A., Younes, A., Landrieu, I., and Lippens, G. (2005) *J Am Chem Soc* **127(29)**, 10138–10139.
- [217] Sibille, N., Sillen, A., Leroy, A., Wieruszeski, J., Mulloy, B., Landrieu, I., and Lippens, G. (2006) *Biochemistry* **45(41)**, 12560–12572.
- [218] Mukrasch, M. D., Bibow, S., Korukottu, J., Jeganathan, S., Biernat, J., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2009) *PLoS Biol* **7(2)**, 399–414.

- [219] Hirokawa, N., Shiomura, Y., and Okabe, S. (1988) *J Cell Biol* **107(4)**, 1449.
- [220] Wille, H., Drewes, G., Biernat, J., Mandelkow, E., and Mandelkow, E. (1992) *J Cell Biol* **118(3)**, 573.
- [221] Jicha, G., Bowser, R., Kazam, I., and Davies, P. (1997) *J Neurosci Res* **48(2)**, 128–132.
- [222] Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C., Poulsen, F., and Vendruscolo, M. (2004) *J Am Chem Soc* **126(10)**, 3291–3299.
- [223] Mylonas, E., Hascher, A., Bernado, P., Blackledge, M., Mandelkow, E., and Svergun, D. (2008) *Biochemistry* **47(39)**, 10345–10353.
- [224] Pirovano, W. and Heringa, J. (2010) *Methods Mol Biol* **609**, 327–348.
- [225] Chou, P. and Fasman, G. (1974) *Biochemistry* **13(2)**, 222–245.
- [226] Garnier, J., Osguthorpe, D., and Robson, B. (1978) *J Mol Biol* **120(1)**, 97–120.
- [227] Garnier, J., Gibrat, J., and Robson, B. (1996) *Methods Enzymol* **266**, 540.
- [228] Chothia, C. and Lesk, A. (1986) *The EMBO journal* **5(4)**, 823.
- [229] Illergard, K., Ardell, D., and Elofsson, A. (2009) *Proteins: Structure, Function, and Bioinformatics* **77(3)**, 499–508.
- [230] Pollastri, G. and Mclysaght, A. (2005) *Bioinformatics* **21(8)**, 1719.
- [231] Sickmeier, M., Hamilton, J., LeGall, T., Vacic, V., Cortese, M., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V., et al. (2006) *Nucl Acids Res* **35(suppl 1)**, D786.
- [232] Ward, J., McGuffin, L., Bryson, K., Buxton, B., and Jones, D. (2004) *Bioinformatics* **20(13)**, 2138.
- [233] Ward, J., Sodhi, J., McGuffin, L., Buxton, B., and Jones, D. (2004) *J Mol Biol* **337(3)**, 635–645.

- [234] Vucetic, S., Brown, C., Dunker, A., and Obradovic, Z. (2003) *Proteins-Structure Function and Genetics* **52(4)**, 573–584.
- [235] Peng, K., Radivojac, P., Vucetic, S., Dunker, A., and Obradovic, Z. (2006) *BMC bioinformatics* **7(1)**, 208.
- [236] Romero, P., Obradovic, Z., Li, X., Garner, E., Brown, C., and Dunker, A. (2001) *Proteins* **42(1)**, 38–48.
- [237] Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) *Bioinformatics* **21(16)**, 3433–3434.
- [238] Thomas, P. and Dill, K. (1996) *Proc Natl Acad Sci U S A* **93(21)**, 11628.
- [239] Sawaya, M., Sambashivan, S., Nelson, R., Ivanova, M., Sievers, S., Apostol, M., Thompson, M., Balbirnie, M., Wiltzius, J., McFarlane, H., et al. (2007) *Nature* **447(7143)**, 453–457.
- [240] Deng, X., Eickholt, J., and Cheng, J. (2009) *BMC Bioinformatics* **10(1)**, 436.
- [241] Wang, Z., Eickholt, J., and Cheng, J. (2010) *Bioinformatics* **26(7)**, 882.
- [242] Noivirt-Brik, O., Prilusky, J., and Sussman, J. (2009) *Proteins* **77(S9)**, 210–216.
- [243] Cheng, J., Wang, Z., Tegge, A., and Eickholt, J. (2009) *Proteins* **77(9)**, 181–184.
- [244] Altschul, S., Madden, T., Schaeffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. (1997) *Nucl Acids Res* **25(17)**, 3389.
- [245] Brandt, R., Léger, J., and Lee, G. (1995) *J Cell Biol* **131(5)**, 1327.
- [246] Munoz, V. and Serrano, L. (1994) *Nat Struct Biol* **1(6)**, 399–409.
- [247] Munoz, V. and Serrano, L. (1995) *J Mol Biol* **245(3)**, 297–308.
- [248] Munoz, V. and Serrano, L. (1995) *J Mol Biol* **245(3)**, 275–296.
- [249] Munoz, V. and Serrano, L. (1997) *Biopolymers* **41(5)**, 495–509.

- [250] Rost, B., Yachdav, G., and Liu, J. (2004) *Nucl Acids Res* **32(suppl 2)**, W321.
- [251] Meiler, J., Mueller, M., Zeidler, A., and Schmaeschke, F. Jufo: secondary structure prediction for proteins.
- [252] Mészáros, B., Simon, I., and Dosztányi, Z. (2009) *PLoS Comp Biol* **5(5)**, 1–18.
- [253] Dosztányi, Z., Mészáros, B., and Simon, I. (2009) *Bioinformatics* **25(20)**, 2745–2746.
- [254] Necula, M. and Kuret, J. (2005) *FEBS Lett* **579(6)**, 1453–1457.
- [255] Sun, Q. and Gamblin, T. (2009) *Biochemistry* **48(25)**, 6002–6011.
- [256] Wishart, D. and Sykes, B. (1994) *J Biomol NMR* **4(2)**, 171–180.
- [257] Schwarzingler, S., Kroon, G., Foss, T., Wright, P., and Dyson, H. (2000) *J Biomol NMR* **18(1)**, 43–48.
- [258] Jeganathan, S., Hascher, A., Chinnathambi, S., Biernat, J., Mandelkow, E., and Mandelkow, E. (2008) *J Biol Chem* **283(46)**, 32066.
- [259] Eidenmüller, J., Fath, T., Maas, T., Pool, M., and Brandt, J. (2001) *Biochem J* **357(1)**, 759–767.
- [260] Grundke-Iqbal, I., Iqbal, K., Tung, Y., Quinlan, M., Wisniewski, H., and Binder, L. (1986) *Proc Natl Acad Sci U S A* **83(13)**, 4913.
- [261] Grundke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y., Zaidi, M., and Wisniewski, H. (1986) *J Biol Chem* **261(13)**, 6084.
- [262] Foster, N., Wilhelmsen, K., Sima, A., Jones, M., D’Amato, C., and Gilman, S. (1997) *Ann Neurol* **41(6)**, 706–715.
- [263] Spillantini, M., Murrell, J., Goedert, M., Farlow, M., Klug, A., and Ghetti, B. (1998) *Proc Natl Acad Sci U S A* **95(13)**, 7737.
- [264] Blennow, K., deLeon, M., and Zetterberg, H. (2006) *Lancet* **368(9533)**, 387–403.

- [265] Shaw, L., Korecka, M., Clark, C., Lee, V., and Trojanowski, J. (2007) *Nat Rev Drug Discovery* **6(4)**, 295–303.
- [266] Hampel, H., Blennow, K., Shaw, L., Hoessler, Y., Zetterberg, H., and Trojanowski, J. (2010) *Exp Gerontol* **45(1)**, 30–40.
- [267] Sigurdsson, E. (2009) *Curr Alzheimer Res* **6(5)**, 446–450.
- [268] Mercken, M., Vandermeeren, M., Luebke, U., Six, J., Boons, J., Voorde, A., Martin, J., and Gheuens, J. (1992) *Acta neuropathologica* **84(3)**, 265–272.
- [269] Mandelkow, E. and Mandelkow, E. (1998) *Trends Cell Biol* **8(11)**, 425–427.
- [270] Weaver, C., Espinoza, M., Kress, Y., and Davies, P. (2000) *Neurobiol Aging* **21(5)**, 719–727.
- [271] Carmel, G., Mager, E., Binder, L., and Kuret, J. (1996) *J Biol Chem* **271**, 32789–32795.
- [272] Jicha, G., Berenfeld, B., and Davies, P. (1999) *J Neurosci Res* **55(6)**, 713–723.
- [273] vonBergen, M., Barghorn, S., Mueller, S., Pickhardt, M., Biernat, J., Mandelkow, E., Davies, P., Aebi, U., and Mandelkow, E. (2006) *Biochemistry* **45(20)**, 6446–6457.
- [274] Matsuo, E., Shin, R., Billingsley, M., Van deVoorde, A., O’Connor, M., Trojanowski, J., and Lee, V. (1994) *Neuron* **13(4)**, 989–1002.
- [275] Ghoshal, N., García-Sierra, F., Fu, Y., Beckett, L., Mufson, E., Kuret, J., Berry, R., and Binder, L. (2001) *J Neurochem* **77(5)**, 1372–1385.
- [276] De Calignon, A., Fox, L., Pitstick, R., Carlson, G., Bacskai, B., Spires-Jones, T., and Hyman, B. (2010) *Nature* **464(7292)**, 1201–1204.
- [277] Ksiezak-Reding, H., Leibowitz, R., Bowser, R., and Davies, P. (1995) *Brain research* **697(1-2)**, 63–75.

Appendices

Appendix A

Preliminary results of phosphorylated tau

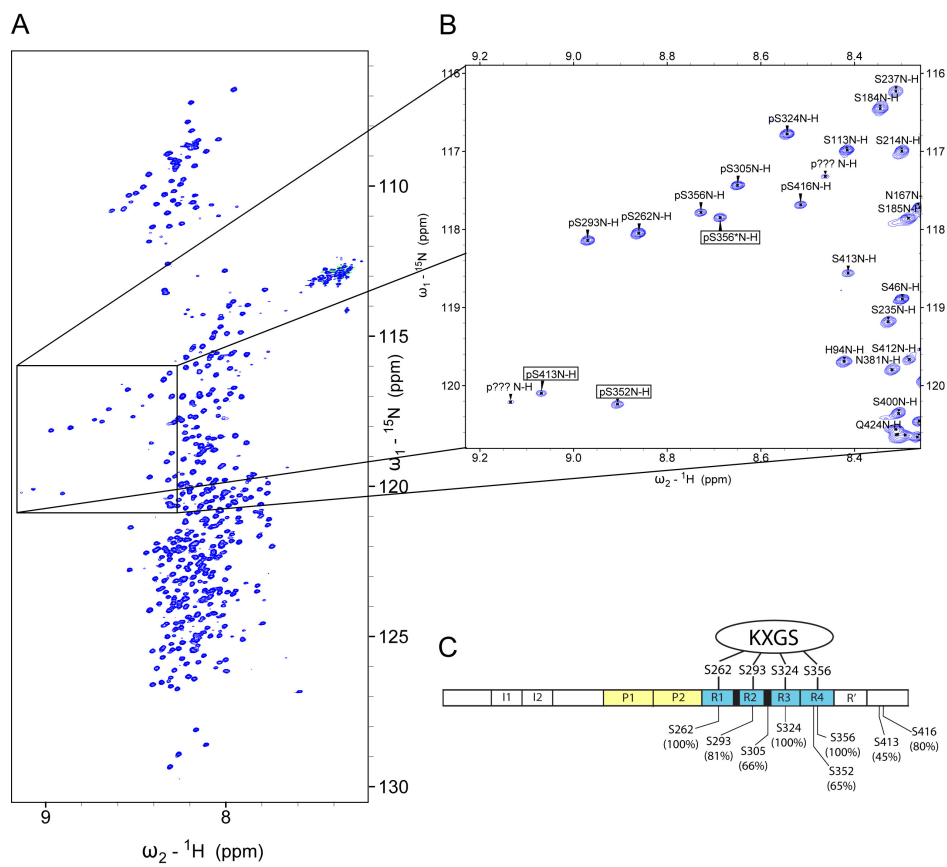


Figure A.1: Tau phosphorylated with MARK

50 μ M tau (pH 6.8, 50 mM phosphate buffer) was incubated with MARK. The HSQC spectrum showed several new peaks that appear in a region previously free of resonances (Figure A.1A and B). The phosphorylated backbone resonances of htau40 were assigned using 3D HACANNH and HNN experiments at 5 °C [207,208]. We found that the serines in the KXGS motifs (S262, S293, S324 and S356), were phosphorylated to 100%, except S293, which was phosphorylated to 81%. In addition, S305 (66%), S352 (65%), S413 (45%) and S416 (80 %) were phosphorylated (Figure A.1B and C, boxed peaks were assigned by Dr. Martin Schwalbe). The incomplete phosphorylation of S352 resulted in the appearance of two peaks for S356. Similarly, incomplete phosphorylation of S413 and S416 might cause the additional, currently unassigned peaks close to pS413 and pS416. However, the very low signal intensities prevented an unambiguous assignment with the 3D HACANNH and HNN experiments.

Appendix B

Pulseprograms

```
;PGSTE WATERGATE 2D
;set up by sebi (March 2010)
;optimized April 2010
;2D sequence for diffusion measurement with
;water suppression using watergate W5 pulse sequence with gradients
;
;G. Zheng, T. Stait-Gardner, P.G. Anil Kumar, A.M. Torres and W.S. Price
;J. Magn. Reson. 191 (2008), pp. 159-163
;
;
;
;
;for error message "cannot find difframp" type "dosy" but don't start
;should then work!!!
;
;
;-----
;I choose the nomenclature according to Bruker (d20, p30 etc)
;d1: 7 to 10 seconds (avoid sample heating)
;p30: little delta
;gpz5: around 20%
;gpz6: incremented from 25-95%
;gpz7: 6.66%
;d20: specifies big Delta (around 100ms)
;d19: delay for binomial water suppression
```

```

; d19 = distance of next null (200us approx. 5000 Hz)
;ns=4*n
;use 'Difftrap' as gradient pulses
;use AU-program 'dosy' to calculate gradient ramp-file Difframp
;2 times delta2 total time is 2 times 2ms (W5 time is subtracted)
;delta2 time is now like in paper
;

```

```

# include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

```

```

define list<gradient> diff=<Difframp>

```

```

define delay delta1
define delay delta2
define delay delta3
define delay bigD

```

```

"p19=p1"
"delta1= 0.2m"
"delta2= 2m-0.93m" ;W5 is 1,86ms long
"delta3=delta1-10u"

```

```

"bigD=d20-3*p30-2*delta2-3*delta1"

```

```

1 ze
2 d1
10u p1:f1
50u UNBLKGRAD
p1 ph1

```

delta1
p30:gp5*-1 ;start of diffusion (big Delta)
delta2
;-----start 1st W5-----
p19*0.087 ph11
d19
p19*0.206 ph11
d19
p19*0.413 ph11
d19
p19*0.778 ph11
d19
p19*1.491 ph11
d19
p19*1.491 ph12
d19
p19*0.778 ph12
d19
p19*0.413 ph12
d19
p19*0.206 ph12
d19
p19*0.087 ph12
;-----end 1st W5-----
delta2
p30:gp6*-1*diff
delta1

p1 ph3
delta1
p30:gp7
bigD

p1 ph1
delta1
p30:gp6*diff ;End of diffusion time bigDelta

```

delta2
;-----start 2nd W5-----
p19*0.087 ph13
d19
p19*0.206 ph13
d19
p19*0.413 ph13
d19
p19*0.778 ph13
d19
p19*1.491 ph13
d19
p19*1.491 ph14
d19
p19*0.778 ph14
d19
p19*0.413 ph14
d19
p19*0.206 ph14
d19
p19*0.087 ph14
;-----end 2nd W5-----
delta2
p30:gp5
delta3
10u BLKGRAD
go=2 ph31
30m mc #0 to 2 F1QF(igrad diff)
exit

```

```

ph1=0 2
ph3=0 1 2 3
ph4=0 1 2 3

```

```

ph11= 0 0 1 1
ph12= 2 2 3 3

```

ph13= 0

ph14= 2

ph31=0 2 0 2

end

```
;Nz-EXSY.sebi
;15N-HSQC type experiment
;should show cross-peaks for slowly exchanging
;residues
;
;based on Farrow, Zhang, Kay JBNMR 1994 727-734
;
;version 1.0 (19. Oct. 2010) written by sebi
;removed water flip black because its bad
;this sequence contains 3-9-19 watergate
;because it works better than a tailored watergate
;
;-----power levels-----
;p11 : power for 1H
;p110 : power for sp10
;p12 : power for 13C
;p13 : power for 15N
;p113 : power for 15N-decoupling
;
;-----pulses-----
;p1 : 90 degree hard pulse 1H
;p2 : 180 degree hard pulse 1H
;p5 : 90 degree hard pulse 15N
;p6 : 180 degree hard pulse 15N
;p10 : shaped pulse 90 degree(125-270 Hz wide on H2O)@p110
;p11 : shaped pulse for WTG
;p17 : shaped pulse in mixing time (2m)
;pcpd3 : 90 deg cpd-pulse15N(waltz16,160u
;
;-----delays-----
;d1 : recycle delay
;d2 : 2.7m (=1/4 Jnh)
;d3 : 2.7m (=1/4 Jnh)
;d19 : 3-9-19 delay to next null
;d22 : mixing time
;in0 : 1/(2 SW) (Hz)
```

```
;-----Gradients-----
;p21 : 1m (Gradient before first INEPT)
;p22 : 500u (Gradient for first INEPT)
;p23 : 1m (cleaning gradient after INEPT)
;p24 : 1m cleaning gradient after 15N-evolution)
;p25 : 500u (Gradient after mixing time)
;p26 : 500u (Gradient for second INEPT)
;p27 : 125u (Gradient for tailored WTG in second INEPT)
;gpz1 : 10% (5 Gauss)
;gpz2 : 8% (4 Gauss)
;gpz3 : 30% (15 Gauss)
;gpz4 : -40% (-20 Gauss)
;gpz5 : -20% (-10 Gauss)
;gpz6 : 16% (8 Gauss)
;gpz7 : 50% (25 Gauss)

;use States-TPPI

# include <Avance.incl>

define delay INEPT_1
define delay INEPT_2

#define GRADIENT1 10u p21:gp1 200u
#define GRADIENT2 10u p22:gp2 180u
#define GRADIENT3 10u p23:gp3 200u
#define GRADIENT4 10u p24:gp4 200u
#define GRADIENT5 10u p25:gp5 200u
#define GRADIENT6 10u p26:gp6 200u
#define GRADIENT7 10u p27:gp7 200u

" in0=inf1/2"
" d0=in0/2-p5*2/3.14159-p1"
" p2=2*p1"
" p6=2*p5"
```

```

”d3=2.7m”
”d16=d3-p26”
;”d22=250m”
”d18=d19/2-p5”
”d17=2.5m”

”INEPT_1=d2-(p22+210u)”
”INEPT_2=d2-(p27+210u+p1*2.3846+d19*2.5)”
”l1=d22/(2*d17+p2)”

```

```

1 10u ze
2 1m do:f3
d1 pl1:f1
20u pl3:f3
20u LOCKH.ON
(p5 ph20):f3 ;ensures that only proton magnetization..
GRADIANT1 ;...contributes to signal
;—————first INEPT—————
(p1 ph20):f1
GRADIANT2
INEPT_1
(center(p2 ph20):f1 (p6 ph20):f3)
GRADIANT2
INEPT_1 pl1:f1
(p1 ph21):f1 ;HzNz and H2Oy
GRADIANT2
;(p10:sp10 ph22):f1 ;flips water to -z
10u pl3:f3 pl1:f1
GRADIANT3 ;cleaning gradient
(p5 ph1):f3
;—————15N evolution with simultaneous refocussing—————
d0
d3
(p2 ph20):f1 ;pulse flips water to -z
d0
(p6 ph2):f3

```

d3
(p5 ph21):f3 ;puts pure N magn. onto +z-axis
GRADIENT4
;—————mixing time—————
3 d17
(p2 ph20):f1
d17 pl3:f3 pl1:f1
lo to 3 times l1
;—————establish antiphase magn—————
GRADIENT5
(p5 ph20):f3
GRADIENT6
d16
(center(p2 ph20):f1 (p6 ph20):f3)
GRADIENT6
d16
(p5 ph21):f3
;—————second INEPT—————
;(p10:sp10 ph3):f1 ;water flip back
10u
(p1 ph4):f1
GRADIENT7
INEPT_2
(p1*0.2308 ph20 d19 p1*0.6923 ph20 d19 p1*1.4615 ph20):f1
(d18 p6 ph1 d18):f3
(p1*1.4615 ph22 d19 p1*0.6923 ph22 d19 p1*0.2308 ph22):f1
GRADIENT7
INEPT_2 pl13:f3 LOCKH_OFF
;—————acquisition
go=2 ph31 cpd3:f3
1m do:f3 mc #0 to 2 F1PH(ip1,id0)

;—phase cycle—

ph20=0

ph21=1

ph22=2

ph23=3

ph1=0

ph2=0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3

ph3=0 0 2 2

ph4=0 0 2 2

ph31=0 0 2 2 2 2 0 0

end

Appendix C

Curriculum Vitae

Name	Stefan Bibow
Address	Eisenacher Str. 15
City	Göttingen
Postal Code	37077
Country	Germany
E-Mail	stefan.bibow3@arcor.de
City of birth	Berlin
Date of birth	01/16/1981
Nationality	German
2000	University entrance certificate ("Abitur")
2001-2007	Study of Biophysics, Humboldt University of Berlin
05/2007	Diploma in Biophysics Diploma thesis: Solution state NMR investigations of the amyloid β -peptide, Advisor Prof. Dr. Bernd Reif
since 07/2007	PhD student at the Max Planck Institute for Biophysical Chemistry in the group of Prof. Dr. M. Zweckstetter (NMR based Structural Biology) in Göttingen

September 16, 2011