ABSTRACT

Title of Document: CONFORMATIONAL AND ASSEMBLY
                            STUDIES OF TWO TAU REPEATS:
                            2R AND 3R

                            Christina Panayiotou Koshiari,
                            Master of Science, 2006

Directed By: Associate Professor, Dr. Victor Muñoz,
                            Department of Chemistry and Biochemistry

Our aim was to investigate the conformational properties of two tau fragments, 2R

and 3R, in response to increasing amounts of a hydrogen-bond stabilizing solvent

trifluoroethanol. 2R and 3R contain two or three of the microtubule binding domain

repeats respectively, which are known to form the core of Pair Helical Filaments

(PHFs) found in an Alzheimer's Disease brain. Using Circular Dichroism and Fourier

Transform Infrared spectroscopy we find that upon increasing the trifluoroethanol

concentration, a conformational transition occurs from mostly random coil with

residual α-helix and β-turn structure into a full α-helix. The high α-helical propensity

of 2R and 3R is consistant with tau's biological role as a natively unfolded protein

that binds to the α-helical C-terminus of tubulin, thereby stabilizing microtubules.

The increase in α-helix content correlates with the formation of fibrillar aggregates *in*

vitro. Atomic Force Microscopy shows that these aggregates show morphological traits with *ex vivo* PHFs.

CONFORMATIONAL AND ASSEMBLY STUDIES
OF TWO TAU REPEATS: 2R AND 3R

By

Christina Panayiotou Koshiari

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2006

Advisory Committee:
Dr. Victor Muñoz, Chair
Dr. George Lorimer
Dr. Dorothy Beckett

DEDICATION

To my parents, Panayiotis and Androulla Koshiari,

and all my friends

-Christina

Στους γονείς μου, Παναγιώτη και Αντρούλλα Κοσιάρη,

και σε όλους μου τους φίλους

-Χριστίνα

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor, Dr. Victor Muñoz, for giving me the opportunity to work in his lab and for guidance and support throughout my graduate career. A special thanks also to my committee members, Dr. Dorothy Beckett and Dr. George Lorimer for all their helpful comments in completing this thesis. Also, I would like to thank Dr. Mourad Sadqi for initializing the tau project and Dr. Rani Kunjithapatham for the 2R and 3R preliminary purification experiments. I would like to also thank my parents, Panayiotis and Androulla Koshiari, for their unconditional support before and during my studies in the United States. Also, I would like to thank the rest of my family: my grandparents, Evyenia and Cristodoulos Chrisanthou and Maria and Nicolas Koshiari; my sisters, Maria and Evyenia, my two brothers-in-law, Elias Markadji and Justin Brown; and my nieces and nephews, Andrea, Christos, Panos, Apostolos, Ariana, and Daniel. My undergraduate advisor, Dr. Jan Teller, for introducing me to laboratory research. My friends, Nicky Mala, Jenny Sigurdardottir, Mona Bryn, Panayiotis Theodorou, Maria Charalambous, and Tanay Desai for always having time to listen. My neighbors in Fordham Court, College Park MD or Melrose Court as we call it. Kathleen, Erica, and Becky thank your for our conversations. The CVS store employees at College Park Plaza. CVS has served as my security blanket during my stay in College Park. On my way home, a bag of candy was absolutely essential for my survival! Also, I would like to thank my fellow labmates, it has been an exciting experience working with each one of you and I wish you all the best in life and research.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION, RESEARCH BACKGROUND, AND RESEARCH OBJECTIVES

1.1 Introduction

1.1.1 Protein Folding and Conformation

Proteins are polymers composed of 20 amino acids and fold to their unique and functional three-dimensional structure depending on their linear amino acid sequence and aqueous environment [1] to carry out a wide range of biological functions such as transport, storage, structural support, or regulation [2]. The key to understanding how different proteins function in these different ways lies in an understanding of protein conformation.

Protein folding is mainly driven by the free energy contributions of intramolecular ionic interactions forming salt bridges between side chains bearing opposite electrical charge, van der Waals induced dipole - induced dipole attractions between closely packed molecules, hydrophobic interactions to exclude the hydrophobic side chains from the polar solvent, and hydrogen bonding between side

α-helix   β-pleated sheet
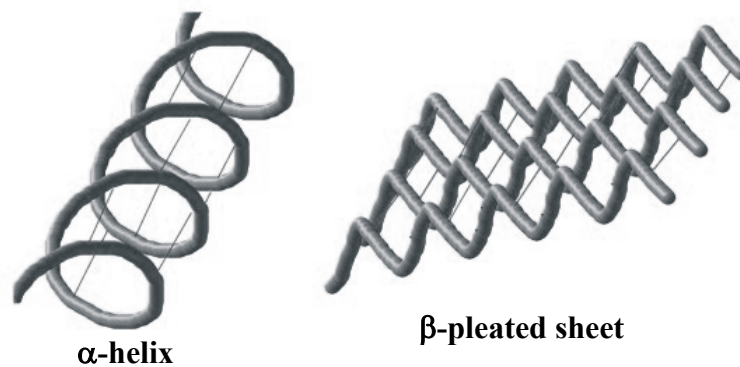
Figure 1.1 Protein secondary structure of α-helix (left image) and β-pleated sheet (right image). The peptide backbone can form intramolecular hydrogen bonds, indicated by gray lines.

chain and backbone N–H and C=O groups [3]. In 1951, Pauling and Corey proposed two periodic folding structures known as the α-helix and β-pleated sheet (Figure 1.1) and the conformation of each depends on the protein's amino acid composition. The α-helix structure is stabilized by the backbone hydrogen bonding between the C=O of amino acid i and N–H group of amino acid i+4; whereas, β-pleated sheets are stabilized by the backbone hydrogen bonding of C=O and N–H between adjacent β-strands [2, 4]. In addition, turns play a role in reversing the direction of the polypeptide chain, which is critical to attain a globular structure. Various turn structures have been identified depending of the number of amino acids involved in the turn. Most common is the β-turn involving four amino acid residues stabilized by hydrogen bonding between the carbonyl oxygen of the first amino acid to the amino group hydrogen of the fourth amino acid [2, 4].

The conformational elements of polypeptide chains (α-helices, β-sheets, and turns) play an important role in globular proteins from both structural and functional points of view. The former because the active native fold of the protein is energetically driven by the formation of those structural elements and the latter because structure elements exposed on the surface of proteins represent molecular recognition sides for the function of the protein [2, 4]. In addition to the natively folded proteins, a new group of proteins has emerged known as "natively unfolded proteins" [5]. As their name implies, natively unfolded proteins lack ordered structure under physiological conditions and some of them can fold upon binding to their cellular target [5, 6]. These proteins are characterized by their low overall hydrophobic amino acid content and high net charge.

1.1.2 Protein Misfolding and Fibrillization

In the search for their natively folded conformation, proteins can be trapped in misfolded states. Proteins that fail to achieve a folded functional state, may result in the formation of deposits consisting of assembled monomeric protein, a process known as protein fibrillization.

The "classical textbook picture" of protein fibrillization is as follows: Protein fibrillization is widely considered to be an assembly of misfolded or partially folded polypeptides. The main driving force behind protein fibrillization is believed to be the hydrophobic clustering of the solvent exposed hydrophobic protein regions. Other energetic factors include intermolecular electrostatic attractions and hydrogen bonding [7]. In order for fibrillization to be thermodynamically favorable, hydrophobic regions need to be exposed to the solvent which requires the folded structure to be destabilized. The exposure of hydrophobic regions will result in decrease of entropy in the system by ordering water molecules around the non hydrogen bonding hydrophobic regions; thus the presence of another hydrophobic region in another polypeptide will result in assembly driven by overall increase in entropy [4]. Changes in protein stability is generally attained by mutations, either sporadic or familial, protein truncations, post-translational modifications, such as glycosylation and phosphorylation, or changes in temperature, pH, ionic strength, and organic solvents [8]. Over 20 conformational diseases have been identified so far including Alzheimer's disease (AD), Parkinson's disease, and Huntington's disease, that are a result of protein misfolding. Recently, it has been proposed that perhaps all proteins can aggregate *in vitro* into fibrils indistinguishable from observed *ex vivo* and

associated to disease [9].

Fibrillar aggregates, studied *ex vivo* or *in vitro*, have been found to consist of fibers containing misfolded protein with β-sheet conformation, and are known as amyloid fibrils. X-ray diffraction studies demonstrated a repeating primary molecular structure in amyloid fibrils [10], which is known as the *cross-β* motif. The fibril structure consists of β-sheets parallel to the fiber axis separated by hydrogen bonds at a 4.7 Å spacing (meridional reflection) and adjacent β-strands orthogonal to the fibril axis separated by 10 Å spacing (equatorial reflection) [10, 11]. The formation of aggregates with similar molecular conformation by numerous proteins and peptides, unrelated in sequence and function, suggested that a generic mechanism governing the aggregation process may exist [8].

An important question is whether all physiological fibrillar aggregates have an amyloid molecular structure. In this regard, a study by Sadqi et al in 2002 suggested that the structure of *ex vivo* Pair Helical Filaments (PHFs) are composed of 85% α-
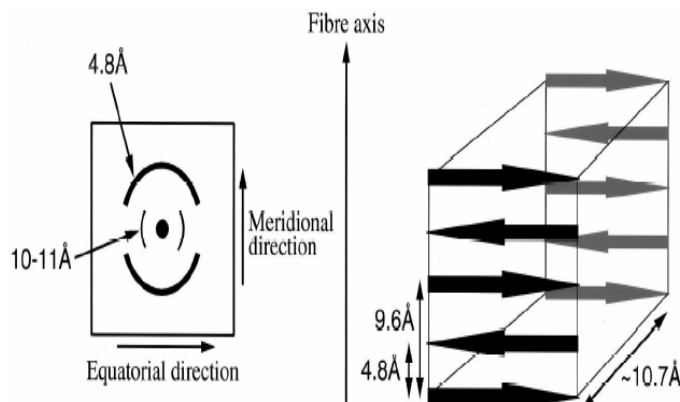
Figure 1.2 The β-amyloid molecular structure. The amyloid x-ray diffraction is characterized by the meridional reflection at 4.8 Å and the equatorial reflection at 10-11 Å (left image). The molecular conformation of amyloid fibrils forming intermolecular hydrogen bonds between parallel β-sheets and orthogonal β-strands is shown (right image). Figure adapted from [9].

helices [12]. PHFs are bundles of fibrillar aggregates composed of tau protein and are deposited extracellularly in AD effected brains. Examples of α-helical polymers with functional roles are well known, but this was the first time α-helices were associated with pathological fibrillar assembly; however, this idea is still controversial. As described above, the major energetic contribution factor of β-sheet polymers is the formation of backbone hydrogen bonding independent of the side chain content. The formation of α-helical polymers requires a more specific design so that favorable side chain-side chain hydrophobic, hydrophilic, and electrostatic intermolecular interactions are capable of compensating the decrease in entropy.

Understanding α-helical assembly and its role in aberrant protein aggregation is still in its first steps. Here, our fundamental goal is to investigate the fundamental properties of conformation and conformation-assembly interplay leading to aberrant fibrillization in two tau fragments from the microtubule binding domain (MBD): 2R and 3R. Previous studies have indicated that tau belongs to a unique group of "natively unfolded proteins" that upon binding to its target gains structure to stabilize microtubules. The MBD region has been identified as the region responsible for microtubule stabilization and also forms the core of PHFs. Therefore, understanding the conformational properties of those two fragments will aid in an understanding of the biological role of tau and of aberrant protein aggregation.

1.2   Research Background

1.2.1   Tau Physiological Role and Fibrillization

Tau was first identified as a neuronal microtubule associated protein due to its co-purification with microtubule preparations [13]. It was later suggested that it plays a role in the stability of microtubules [14]. Tau stabilizes microtubules by binding to the α-helical C-terminus of tubulin [13, 15]. The structural details of tau in its microtubule bound and free form are still largely unknown. Tau has a high hydrophilic amino acid content and it has resisted all efforts of crystallization, whereas the full-length tau molecular weight impedes structural analysis by nuclear magnetic resonance (NMR) methods. Most of the available structural data come from spectroscopy, hydrodynamic measurements, and NMR on smaller fragments. All of these experiments indicate that the polypeptide chain is a natively unfolded protein [16-19].

Tau has attracted considerable interest in the last few decades since its discovery as the major component of the intracellular fibrillar aggregates found in AD brain named Pair Helical Filaments (PHFs). PHFs accumulate in brain regions associated with memory retention, and the amount of filaments correlate with the severity of dementia [20]. It is currently unclear as to how tau makes the transition from an important component in the microtubule cytoskeleton to an insoluble dysfunctional polymer. In the past few decades, much effort has been applied to understanding the ultrastructure of PHFs, the molecular composition of PHFs from *ex vivo* and *in vitro* grown filaments, the structural properties of tau, and the mechanism of assembly.

6

1.2.2 Tau Gene and Important Regions

In the central nervous system, multiple tau isoforms are expressed from a single 15 exon gene located on chromosome 17. By alternative mRNA splicing of exons 2, 5, and 10, six major tau isoforms are expressed ranging from 441 amino acids in the longest isoform to 353 amino acids in the shortest fetal isoform [21]. In SDS-PAGE, these are seen as a group of bands between 55-62 KDa [18]. In the peripheral nervous system, a higher molecular weight tau, around 100 KDa, is expressed [22]. Tau isoforms can be divided into three regions that play distinct functions in stabilizing microtubules. The N-terminal projection domain and the C-terminal domain are known to play a role in the spacing between axonal microtubules, and the microtubule binding domain (MBD) stabilizes microtubules by binding in the exterior and interior of microtubules to the C-terminus of tubulin. The MBD domain is composed of three or four imperfectly repeated sequences containing 31 or 32 residues, depending of whether exon 10 is expressed,. The highest microtubule binding affinity is found when all four repeats are present [21]. The MBD has been also identified as the region forming the core of PHFs. Greenberg and Davies, using PHFs extracted from AD brain homogenates and antibodies for different regions of tau found that the four or three repeat regions, depending on the isoform, contained the packed core of the PHFs whereas the N- and C- terminal parts
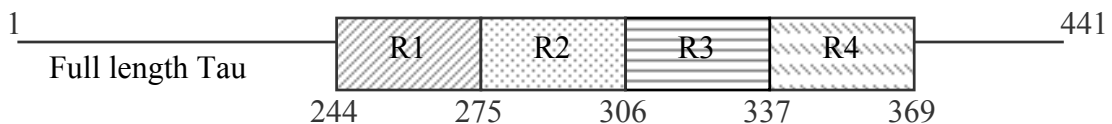
Figure 1.3 The longest tau isoform consists of 441 amino acids and contains four microtubule binding repeats regions (R1, R2, R3, R4).

7

form a "fuzzy coat" around the PHF core [23]. Later, *in vitro* tau assembly studies have further supported the important contribution of MBD to tau assembly [24-26]. The role of the N- and C- termini in tau assembly is still under investigation, but it has been suggested to play an assembly inhibitory role [27], or an enhancing role [28].

1.2.3 Ultrastructure of PHFs

Tau polymerizes into PHFs were all six tau isoforms are incorporated in the filament. The MBD region forms the core and the N- and C- termini the "fuzzy coat" around the core of the PHFs. In 1963, Kidd and colleagues studied *ex vivo* PHFs which showed a characteristic "paired helical" structure under the electron microscope, resembling two filaments twisted around each other in a left handed helical configuration with a crossover periodicity of ≈ 80 nm [29] and diameters of ≈ 10 and ≈ 20 nm at the narrowest and widest regions respectively (Figure 1.4) [23, 30]. In addition to PHFs, ex vivo polymers of tau can also attain a straight filament (SF) morphology with a diameter of ≈ 15 nm [26].

150 nm
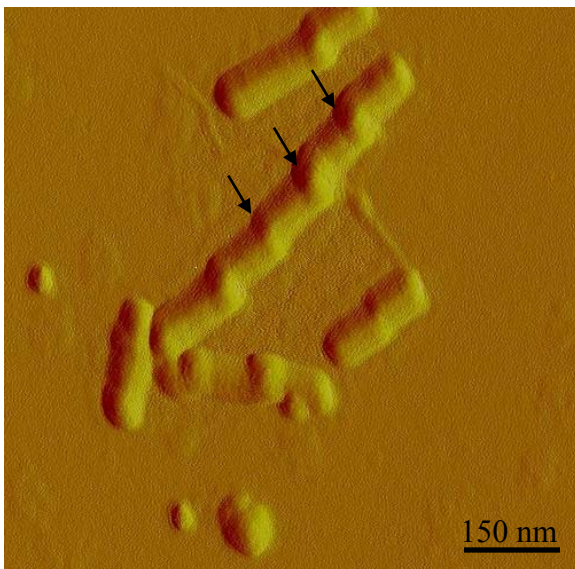
Figure 1.4 Atomic Force Microscopy (AFM) image of *ex vivo* Pair Helical Filaments (PHFs). The characteristic PHFs periodicity at 80 nm is indicated by arrows.

1.2.4 Molecular Structure of PHFs

Although the ultrastructure of PHFs is well established, the molecular structure of PHFs is still debated. One view supports the β-sheet amyloid like fibrils [31-33], the other view supports the α-helical structure model [18, 34], and even random coil fibrils have been suggested [17]. Why is the molecular structure characterization of PHFs so challenging? Conflicting results can be rationalized by the variables of each study: various purification procedures employed in different labs and co-purification of other cellular proteins such as ferritin, which is 75% helical, and β-amyloid which is known to form amyloid fibrils. In general, cellular purification of intact fibrillar particles is rather challenging due to the harsh purification procedures that are typically required. Also, classification of fibrils can be easily obtained using x-ray diffraction where regular building blocks in the molecular structure can be identified. In the case of PHFs, however, this has not been achieved yet due to the inherent difficulties of aligning the highly heterogeneous and irregular PHFs [32, 35].

However, the molecular structure studies in conjunction with the ultrastructure information available, suggest that PHFs cannot be easily categorized as amyloid fibrils, as it has been proposed. PHFs are thicker (widths of 20 nm and 10 nm at the widest and narrowest points) than amyloid fibrils (widths of 10 nm) and morphologically more complex. The structural β-sheet propensity of tau is minimal, and polymerization of tau into PHFs does not follow the classical principles of amyloid fibrillization such as dependency on protein concentration [36]. Interestingly, the fact that PHFs appear in a number of neurodegenerative diseases, termed

taupathies, in addition to AD, and combined with genetic linkage of tau to a specific subclass of these disorders, suggests that tau polymerization may be distinct from amyloid formation [36]. To understand what triggers the accumulation of abberant aggregates *in vivo*, is it therefore necessary to understand the assembly properties of tau *in vitro.*

1.2.5   Inducing *Bona Fide* PHFs in Vitro

In the past decade, there have been many efforts to develop assays to study the conformational and assembly properties of tau *in vitro*. Due to tau's high hydrophilic amino acid content, and perhaps the requirement of post-translational modifications such as phosphorylation or dephosphorylation, glycosylation, and oxidation, fibrillization *in vitro* does not proceed spontaneously even at supraphysiological protein concentrations [21].

During purification of PHFs from diseased AD brains, it was observed that other non-protein components naturally present in the cytosol were co-purified with PHFs. This observation suggested that components could facilitate tau polymerization *in vivo* that lead to their use in *in vitro* studies to enhance tau assembly. Compounds used for *in vitro* studies include: polyanions (e.g., heparin, polyglutamate, and RNA) and fatty acids (e.g., arachidonic acid). Other solvents used include trifluoroethanol (TFE), which is known to stabilize hydrogen bonds of α-helices and β-sheets.

The exact mechanism of how these compounds act is not known, nor is their physiological relevance. However it is suggested that the clustering of negative
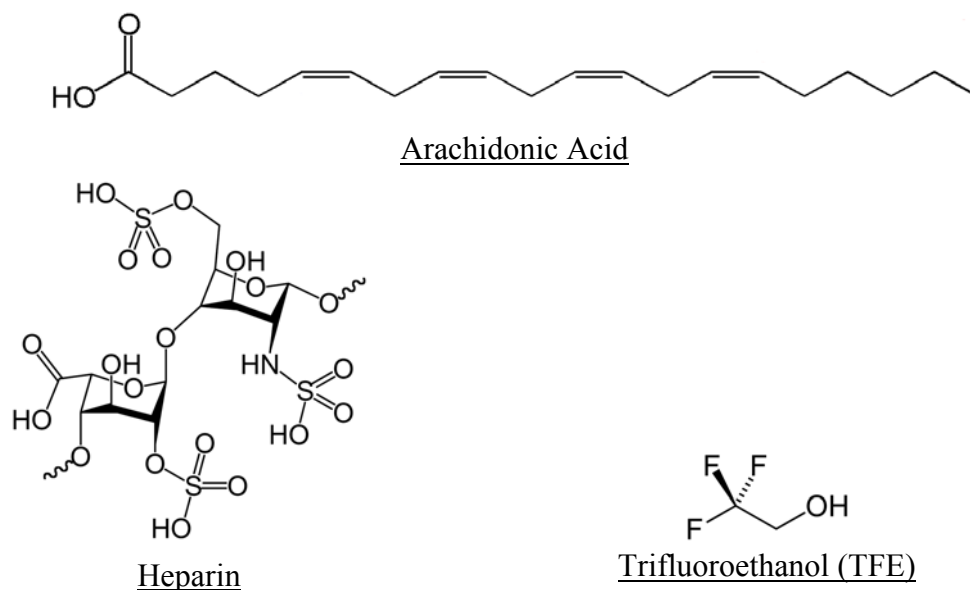
Figure 1.5 Structure of molecules used to induce tau assembly in vitro: arachidonic acid, heparin, and trifluoroethanol (TFE). The exact mechanism of how these inducers are acting to promote tau assembly is not known, but it is believed that the clustering of negatives charges neutralize the highly positively charged tau and promote polymerization.

charges to neutralize the highly positively charged tau plays a role in stabilizing the

assembly process.

1.2.6   Tau Conformational Transition to Assembly

The transition of the soluble unfolded tau into insoluble fibrils using different

inducers, has been shown to be accompanied by a change in conformation and the

adoption of several secondary structures. The structures suggested are α-helix, β-

sheet, and turns. Also, it has been suggested that the capacity of tau to assemble may

depend on a specific conformation and a specific region of the polypeptide chain [24,

37, 38].

It has been suggested, mainly using heparin as the inducing agent, that

induction of a β-sheet conformation leads to tau assembly into fibrils resembling amyloid characteristics. This idea was supported by spectroscopic methods signaling a β-sheet conformation, by x-ray diffraction indicating the typical *cross-β* amyloid like pattern, and by Electron Microscopy indicating the presence of fibrils with morphology similar to fibrils resemble that of *ex vivo* PHFs [32, 39]. The minimal region for tau assembly was localized in the second and third repeats hexapeptides, ^{275}VQIINK^{280} and ^{306}VQIVYK^{311} [25]. Also, it has been suggested that inducing a β-sheet conformation in these regions could result in the formation of a nucleus to promote assembly and to form the core of the PHFs [39]. The use of truncated tau, that contains the MBD region, and formation of tau dimers through cysteine crosslinking, are thought to enhance assembly [39]. The identification of β-sheet character in this hexapeptide has been extrapolated to draw conclusions for the 441 amino acid protein; however, it is been shown that the MBD region may show β-amyloid character nonspecifically [34, 40]. Also, the use of small hydrophobic peptides to study fibrillization has been used in many studies and is known to form non-specific amyloid fibrils[40].

On the other hand it has been found, using TFE as the inducing agent, that induction of an α-helical conformation leads to tau assembly into fibrils resembling coiled-coil polymers. This idea was supported by spectroscopic studies indicating an α-helical structure with coiled-coil structural characteristics, and AFM studies, using full-length mutated tau, indicating fibrils resembling that of *ex vivo* PHFs [19, 34, 41]. The α-helical character of the MBD was also supported by using arachidonic acid. Also, phosphorylation of serine and threonine in tau, which is thought to

enhance tau assembly into PHFs, is suggested to play a role in stabilizing α-helix formation by acting as helix caps [42]. The α-helical propensity was found to be localized in the 13-14 amino acid stretch known as the interrepeat, which has been shown to influence tau binding and stabilize microtubules [19]. Studies utilizing the third repeat, have proposed that the interrepeat has an amphipathic character where lysines 273 and 280 are located on the same face of the helix to interact with the negatively charged surface of microtubules [19, 41]. Also, it has been suggested that a β-sheet conformation is stabilized in the hexapeptide region; proline-lysine in the third repeat may play a role in forming strand-turn-helix assembly, and turn structures in the PGGG regions in the end of the four repeats. The application of the inducing agents to study polymerization has also indicated that inclusion of the MBD results in a greater efficiency to form filaments [21, 24].

The results of these studies suggest that the various inducing molecules do not act via a common mechanism to promote tau assembly. Different inducers stabilize different secondary structures and regions in the polypeptide. Interestingly, the MBD region alone can assemble into *bona fide* PHFs, suggesting the importance of this region to adopt a secondary structure. Also, since the molecular structure of PHFs is not yet resolved, and β-sheets and α-helices have been proposed as the PHFs' building blocks, targeting *in vitro* PHFs structure has been elusive. However, even though the molecular conformation is not yet resolved, the biological role of tau can give clues about its conformation, based on its function of binding tubulin. Structural data have indicated an α-helical conformation in the region known to be the tau

binding site, which suggests that tau could also adopt an α-helical conformation. Also, structure prediction programs indicate that tau MBDs have significantly high intrinsic propensity to form α-helices. As previously mentioned in section 1.2.3, PHFs have a different ultrastructure than the typical amyloid fibrils, suggesting that PHFs and the increasing number of taupathies may have a molecular structure composed of α-helices as opposed to the classical amyloid fibrils composed of β-sheets. *In vitro* experiments, suggesting the formation of amyloid fibrils, were carried out at high protein concentration in the presence of heparin, whereas experiments resulting in α-helical structure by the presence of TFE or arachidonic acid, were carried out at low protein concentration. The formation of tau amyloid fibrils could be non-specific and simply induced by charge compensation induced by heparin. In fact, heparin was used to study the amyloidogenic properties of α-synuclein [43], and was also shown by NMR to stabilize tau in a β-sheet conformation in 50% of the MBD including the interepeat region that is presumed to form an α-helix [44]. Also, it has been shown by NMR experiments that heparin is indeed incorporated into the polymerized tau, but it has yet to be shown that heparin interacts with tau *in vivo*. On the other hand, heparin in not localized intracellularly as PHFs are thus their biological interaction *in vivo* is unclear. The advantage of TFE is that it is known to induce formation of both amyloid fibrils [8] and α-helical fibrils [34], thus it is not as biased toward amyloid formation as heparin seems to be. However, neither TFE nor heparin are physiologically relevant; TFE is a solvent commonly used in conformation studies and heparin is located extracellularly. Even though heparin was shown to be present in PHF purifications it might just be due to non-specific

electrostatic binding during the purification of *ex vivo* PHFs. The α-helical coiled-coil

character is interesting since it is the first time it has been connected to pathological

fibrillization; however, this hypothesis is still in its early stages and it remains to be

identified what triggers the α-helical polymerization, and the role of the amphipathic

helix proposed, and its biological significance.

Previous work has led to a basic understanding of the *in vitro* conditions

needed to induce a conformational change in the unfolded full-length tau and the

conformation significance of the MBD to stabilize microtubules and promote

assembly. However, what is fascinating about this system is that the α-helical and β-

sheet contributions may both play a role in the function of tau and assembly of PHFs.

However, more studies are needed to elucidate the conformation of tau and

conformation-assembly interplay in PHF assembly.

1.3   Research Objectives

Previous experiments have proposed that assembly of tau can be enhanced using the MBD region and structure determination can be improved by eliminating the N- and C- termini, which are thought to be not incorporated in the PHFs core. Also, it has been proposed that upon utilizing TFE, the conformation of α-helix and β-sheet can be studied and their influence in tau assembly can be assessed.

The primary goal in this project was to purify and investigate the conformational and assembly properties of two tau fragments, 2R and 3R, which contain two and three of the microtubule repeat regions (Figure 1.6), respectively. As discussed above, the microtubule binding repeats bind to microtubules and seem to form the core of the PHFs. The conformation of tau was studied upon stepwise increments of TFE from 0 to 94%. Our experimental approach was based on the combined use of Circular Dichroism (CD) and Fourier Transform Infrared (FTIR) in Attenuated Total Reflection (ATR) and in Transmission mode to investigate the conformational changes of the molecule, and the use of Atomic Force Microscopy (AFM) to investigate if our conditions induced *bona fide* PHFs and characterize their ultrastructure.

This study will be useful not only in elucidating the conformation-assembly interplay in tau, but more generally might provide information on the conformation of properties of the new group of proteins that are natively unfolded.

1.4   Research Outline

1.4.1   Purification of 2R and 3R

Our first aim was to purify the 2R and 3R fragments. The purification of natively unfolded tau can be challenging due to its low mean hydrophobicity and a relatively high positive net charge, which facilitates non-specific binding to other cellular components during purification, such as nucleic acids. Additionally, due to its unfolded structure, it is easily cleaved proteolytically during the purification procedure by the release of proteases. Here, we used a combination of the expression procedure used for the full-length tau, and developed a modified purification procedure. In our purification procedure we omitted the used of protease inhibitors, because they would have been present in the final purified stock solution, and could potentially alter the structural properties of our fragments. Also, we applied reverse-phase chromatography in our final purification step, which provided an efficient method to purify 2R and 3R based on their molecular size and hydrophobicity.

1.4.2   Conformational Properties of 2R and 3R

Our second aim was to study the conformational properties of 2R and 3R. Here, our objective was to use protein concentration in the low micromolar range, and study the conformational properties of 2R and 3R in the presence of TFE as the inducing agent. The experiment was carried out in low protein concentrations since it has been previously suggested that high protein concentrations might promote non-specific stabilization of β-sheets leading to formation of β-amyloid fibrils [45]. TFE was chosen because it is known to act as hydrogen bond stabilizer thus promoting

both β-sheet and α-helices [8, 34].

The outcome of the spectroscopic experiments should aid in answering fundamental questions in the role of conformational adoption to stabilize microtubules and to fibril formation. It was previously shown for full length and various tau repeats, that TFE progressively enhances the α-helical content of the molecule [41]. Also, it was suggested that adoption of β-sheet and turn structures may be important in stabilizing a favorable conformation for tau polymerization [37, 46]. We would like to investigate the secondary structures content of 2R and 3R as a function of TFE concentration and the role of secondary structure in stabilizing assembly. Since the two fragments differ only in the 4th repeat, it also allows us to determine if the induction of α-helical conformation in tau has a linear dependence on the TFE concentration or if there is a cooperative effect. Conformation will be studied using several techniques, that monitor different properties of the amide bond and require different sample preparation (see Methods and Materials). We are also interested in investigating if the determination of secondary structure is consistent among those different spectroscopic techniques.

1.4.3   Assembly Characterization of 2R and 3R

Our third aim was to study the 3R assembly in various TFE concentrations by AFM. Previous studies have used mostly electron microscopy and hanging drop techniques. EM involves fixing the sample that could alter the ultrastructure of the fibril, and hanging drop involves sample concentration that is believed to favor β-amyloid fibrillization. AFM will allow us to study assembly in atomic resolution with

minimal sample preparation.

Previous research that studied tau assembly with TFE, suggested that assembly can proceed with the use of mutated full length tau [34]. Other studies suggested that assembly cannot proceed utilizing only the MBD regions unless heparin is used concurrently [37]. Here, we would like to characterize the ultrastructure of 3R aggregates. In our study only the MBD region is included, with exclusion of the N- and C-termini the "fuzzy coat" is eliminated, thus the fibril height distribution, width, and periodicity might be altered. Also, in combination with the spectroscopic studies, we would like to systematically characterize assembly and determine what is the interplay between conformation and its propensity to aggregate and what are the optimal conditions leading to assembly.

(a) Tau Fragments
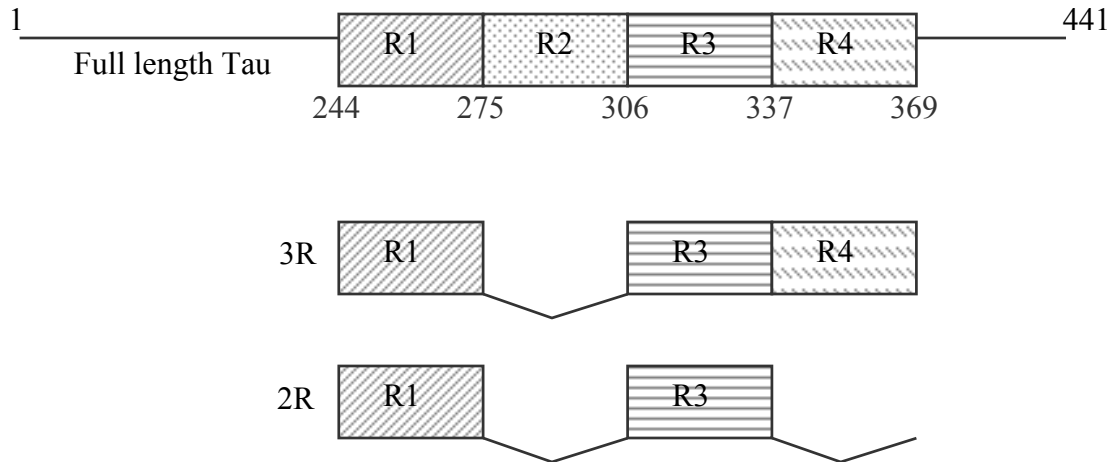
(b) Repeat Sequence

```
                      10          20          30
R1   NH₂-QTAPVPMPDLKNVKSKIGSTENLKHQPGGGKV-COOH
R2   NH₂-VQIINKKLDLSNVQSKCGSKDNIKHVPGGGS-COOH
R3   NH₂-VQIVYKPVDLSKVTSKCGSLGNIHHKPGGGQ-COOH
R4   NH₂-VEVKSEKLDFKDRVQSKIGSLDNITHVPGGGN-COOH
```

Figure 1.6. Schematic of the 2R and 3R tau fragments. 2R contains the 1st and 3rd repeat where as 3R contains the 1st, 3rd, and 4th repeat (a), amino acid numbering is based on the longest tau isoform containing 441 residues. The amino acid sequence of each repeat is indicated in (b).

2. METHODS AND MATERIALS

2.1 2R and 3R Expression

Clones for 2R and 3R fragments were kindly provided by Dr. Jesús Avila (Universidad Autónoma de Madrid, Madrid, Spain). Clones were inserted in the IPTG inducible plasmid pRK172 and transformed in BL21 (DE3) strain of *Escherichia coli*. Transformation was completed in our lab by Dr. Rani Kunjithapatham. The sequence of 2R and 3R is shown in Figure 2.1.

Proteins were expressed and as described previously [34] with a few modifications in the procedure. The overall procedure is herein described. Overnight cultures from single colonies were diluted 1/200 and served for large-scale cell amplification. Cell growth occurred in LB medium at 37 °C in a shaking incubator at 200 rpm until the cultures reached an optimal density of 0.8 at 600 nm. Expression was then induced with 1 mM isopropyl-β-D-thiogalactosidase (IPTG) and the culture was incubated for an additional 3 hours. The cells were then harvested by centrifuging for 30 minutes at 3,750 rpm. Protein extraction was done by

```
               MPDLKNVKSKIGSTENLKHQPGGGKV      Repeat-1
  2R (64aa)    QIVYKPVDLSKVTSKCGSLGNIHHKPGGGQV  Repeat-3
               EVKSEKL                          Repeat-4


               MPDLKNVKSKIGSTENLKHQPGGGKV      Repeat-1
  3R (100aa)   QIVYKPVDLSKVTSKCGSLGNIHHKPGGGQV  Repeat-3
               EVKSEKLDFKDRVQSKIGSLDNITHVPGGGNK  Repeat-4
               KIETHKLTFRE                      N-Terminus
```

Figure 2.1 Sequence of the purified tau fragments: 2R and 3R. The length of 2R and 3R is 63 and 100 amino acids respectively. 2R contains the first, third repeat regions, and 6 residues from the fourth repeat; 3R contains the first, third, fourth repeat regions, and 10 residues from the N-terminus of the full-length tau protein.

21

resuspending the pellet in 50 mM sodium phosphate buffer (pH 7.0), containing 2

mM phenylmethanesulfonyl fluoride, 5 μg/mL leupeptin, 5 μg/mL aprotinin, 45

μg/mL protease inhibitor cocktail, and 7 μM β-mercaptoethanol. Samples were

sonicated 7 times with 1 minute duration at power setting 10 in an ultrasonic cell

disruptor. Bacterial debris were pelleted by spinning for 30 minutes at 15,000 rpm,

and the supernatant was boiled for 30 minutes at 100 °C and further centrifuged at

14,000 rpm for 30 min. The soluble protein was precipitated by the addition of

ammonium sulfate (25% m/v) incubated for 1 hour at 4°C with constant stirring, then

the pH was raised to 12.0 with the addition of 12 M sodium hydroxide, and the

protein further precipitated by the addition of ammonium sulftate (50% m/v). The

increment in pH was done due to the high positively charged content of the tau

protein to avoid co-purification with DNA. The protein was pelleted out by

centrifugation at 25,000 rpm at 4 °C for 2 hours and then resuspended in double

distilled water. The expression and extraction of 2R and 3R was characterized by

SDS-PAGE (Figure 2.2).
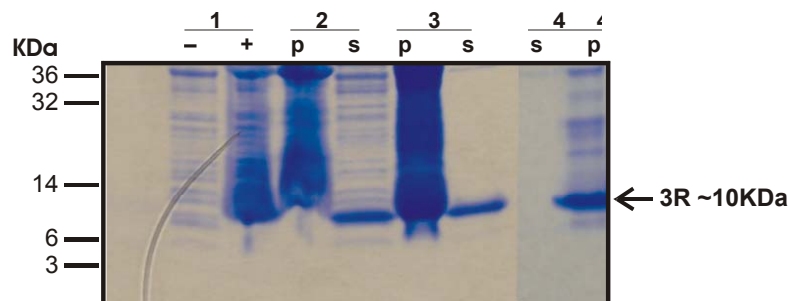
Figure 2.2. SDS-PAGE of 3R expression and extraction. Four steps of the expression
and extraction procedure are shown: 1. before (- lane) and after (+ lane) induction
with IPTG, 2. the pellet (p) and supernatant (s) after the seven sonication cycles, 3.
after boiling at 100 °C for 30 minutes, and 4. after precipitation with 50% ammonium
sulfate (ASP). The SDS-PAGE shows the 3R major band at ~10KDα, a lower
molecular band at ~8KDα, and a range of high molecular bands at 21-36KDα. Gel
was visualized with comassie staining.

2.2   2R and 3R Purification:

The pellet after 50% ammonium sulfate precipitation was dissolved in water (1 mL water per 2 liters cell culture) and centrifuged for 10 minutes at 13,000 rpm to remove any precipitates. Supernatant was applied to the HiPrep 26/10 desalting column (Amersham Biosciences, Piscataway, NY) and eluted with HPLC water at 5 ml/min. Protein was monitored by absorbance at 205 nm. Both 2R and 3R, were eluted in the exclusion volume at ~7 min (Figure 2.3). Desalting is an excellent alternative to dialysis when handling samples sensitive to degradation, like tau. HiPrep 26/10 column was prepacked with Sephadex G-25 fine which is a bead-formed gel prepared by crosslinking dexran and epichlorohydrin. With an exclusion limit of 5 KDa, proteins and peptides larger than 5 KDa will be eluted in the exclusion volume and be separated from molecules (like salts) with a molecular weight less than 1 KDa. Our protein fractions were free from DNA contamination as confirmed with UV absorbance at 260 nm. Protein fractions were pooled together and lyophilized in order to concentrate the sample for further purification.

Figure 2.3. Desalting purification of 3R. (a) Chromatogram of 3R purification. 3R was eluted with a flow rate of 5 mL/min at ~7.5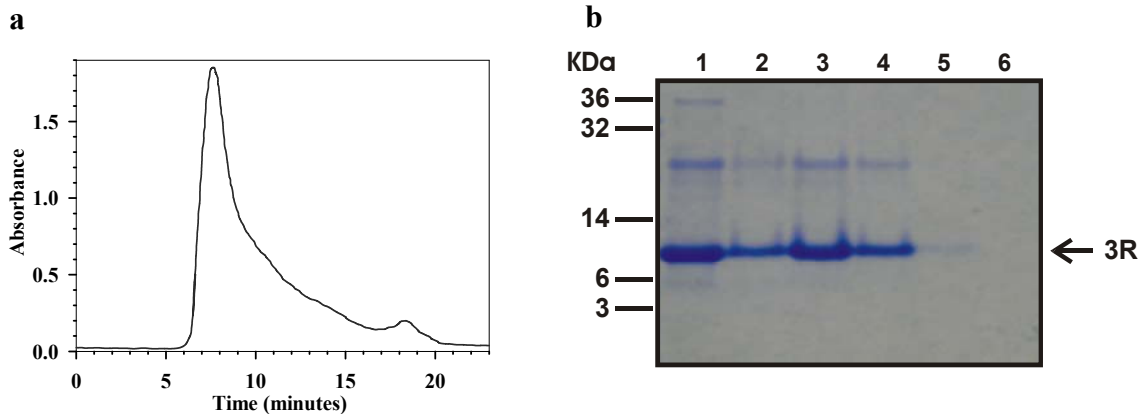 min. (b) SDS-PAGE comassie stained gel of the collected column fractions. Lane 1, the loaded sample. Lane 2, 5-7 minutes fraction. Lane 3, 7-10 minutes fraction. Lane 4, 10-17 minutes fraction. Lane 5, 17-20 minutes fraction. Lane 6, >20 minutes fraction.

The lyophilized powder was dissolved in water (1 ml water per 2 liters cell culture) and centrifuged for 10 minutes at 13,000 rpm in order to eliminate any precipitates. Supernatant was applied to the C18 reverve phase (RP) column (Vydac, Hesperia, CA) and eluted using Solvent A and Solvent B. Solvent A contained 5% Acetonitrile and 0.1% Trifluoroacetate; Solvent B contained 95% Acetonitrile and 0.1% Trifluoroacetate. Both fragments were eluted at ~42 minutes (Figure 2.4.) at 5 ml/min flow rate. The C18 column was equilibrated for 10 minutes with 100% Solvent A, followed by a 50 minute continued gradient from 0% Solvent A to 100% Solvent B, followed by 10 minutes constant 100% Solvent B. Reverse-phase chromatography provides a quick protein separation method based on two factors: molecular size and hydrophobicity. Fractions containing the protein were then combined and lyophilized.

The purity of the lyophilized powder was confirmed with MALDI-TOF mass spectrometry (Figure 2.5). The actual molecular weight of 2R is 6674.9 (~2%
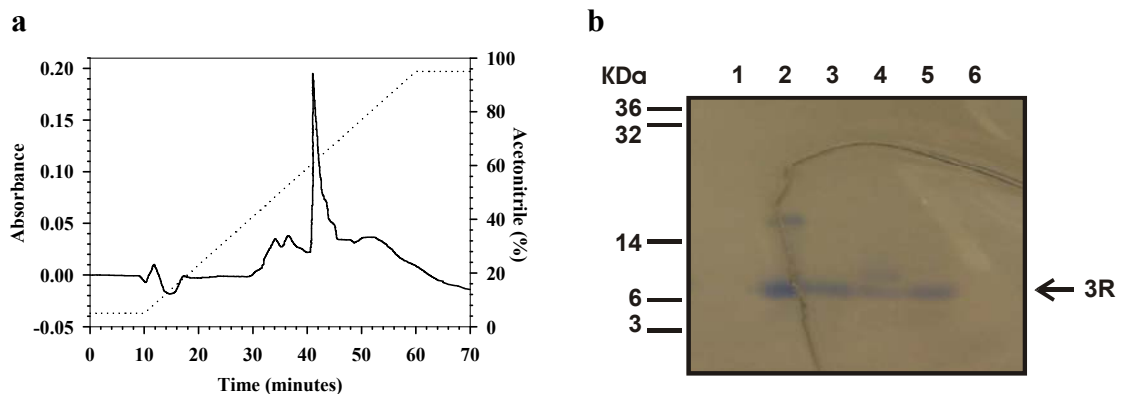
Figure 2.4 RP-HPLC purification of 3R. (a) Chromatogram of 3R purification. 3R was eluted with a flow rate of 5 ml/min at ~ 42 minutes (solid line) with a continued gradient of Solvent A, 5% Acetonitrile and 0.1% TFE; and Solvent B, 95% Acetonitrile and 0.1% TFE (dotted line). (b) SDS-PAGE comassie stained gel of the collected column fractions. Lane 1, 35-41 minutes fraction. Lane 2, 41-43 minutes fraction. Lane 3, 44 minutes fraction. Lane 4, 45 minutes fraction. Lane 5, 45-50 minutes fraction. Lane 6, >50 minutes fraction.

difference from theoretical molecular weight of 6806.9) and of 3R is 10795.7 (~0.5%

difference from theoretical molecular weight of 10855.4). The 3R isoform, apart from

the monomeric protein also contained a small fraction of dimeric protein, whereas the

2R isoform contain fragments of the 2R isoform.
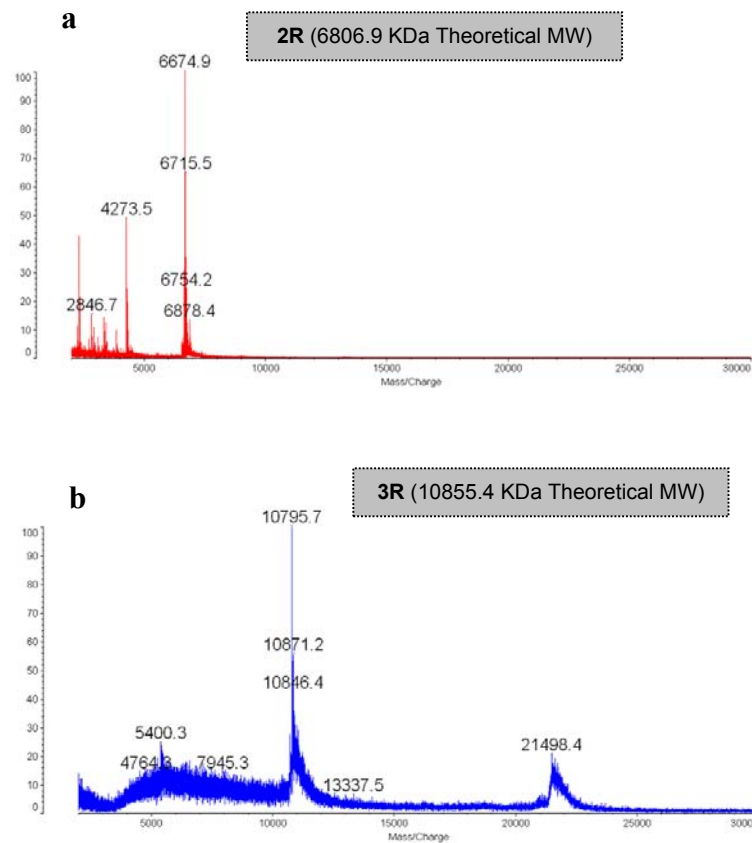
Figure 2.5 MALDI-TOF mass spectrometry spectra of 2R (a) and 3R (b). Analysis was performed using a LINEAR method calibrated with three protein standards ranging from 6-22 KDα. The theoretical molecular weight of 2R is 6806.9 and 3R is 10855.4 as calculated with ProtParam.

2.3   Determination of Protein Concentration

As tau is free of tyrosine and tryptophan residues, three methods were used for determination of concentration: bicinchoninic acid assay (BCA) Colorimetric Assay (Pierce, Rockford, IL), Bradford Assay (Pierce, Rockford, IL), and UV Spectroscopy.

Before concentration determination, the protein powder was carefully weighed on an analytical balance and dissolved in 20 mM sodium phosphate buffer (pH 7.0). First, protein concentration was determined using BCA Colorimetric assay calibrated to a standard curve generated with bovine serum albumin. The BCA assay relies on two reactions. Firstly, the Cu^{+2} ions, BCA, and the peptide bonds or some of the amino acid residues in the protein form a complex, which reduces Cu^{+2} to Cu^{+1}, where the amount of Cu^{+2} reduction is proportional to the amount of protein present in the solution. Then, the Cu^{+1} ion and BCA form a complex, which has a purple-blue color and strong absorption at 562 nm. With this method, protein concentration was determined to be 90% of the original weighed powder. Considering the unfolded nature of tau, more amino acids should be expected to form a binding complex than in a regular folded protein of the same size. Thus this method was likely to overestimate protein concentration. In the second procedure, protein concentration was determined with the Bradford Assay. The Bradford assay relies on the absorbance shift in the dye coomassie, to an absorption maximum at 595 nm, when bound to arginine and hydrophobic amino acid residues present in the protein.The increase of absorbance at 595 nm is proportional to the amount of bound dye, and thus to protein concentration. With this method, concentration was determined to be 10% of the original weighed protein powder. Considering that 2R and 3R contain a minimal amount of

hydrophobic residues this method was likely to greatly underestimate protein concentration. Since both of these conventional methods were not adequate for protein concentration determination of our fragments, we alternatively measured protein concentration by absorbance at 193 nm, which is based on backbone absorbance. With this method, and using a molar extinction coefficient of 8000 cm^{-1} M^{-1} we determined that 65% of the initial weighted material was protein. This result is consistent with those of many other freeze-dried proteins.

2.4 Circular Dichroism (CD) spectroscopy

2.4.1 Theory

CD spectroscopy measures the difference in the absorption of left- and right-handed circularly polarized light and can provide information about the continuous structure of chiral macromolecules.

Light is an electromagnetic wave composed of two wave planes. One plane is electric and the other is magnetic, which oscillate perpendicular to each other and to the direction of light propagation. For *unpolarized light*, the electric vector oscillates in any possible direction perpendicular to the direction of propagation (Figure 2.6a).
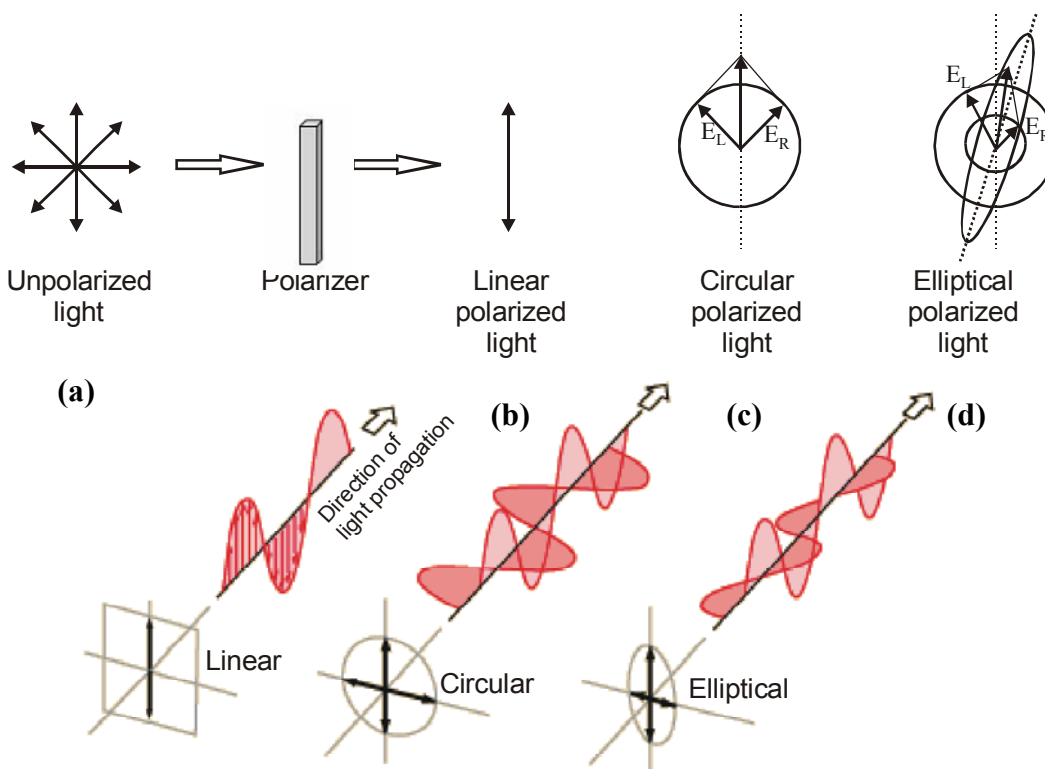
Figure 2.6 Circular Dichroism (CD) and light polarization. Unpolarized light (a) can be converted to linear-polarized light (b) with the use of a polarizer. Linearly polarized light can be represented as the sum of equal magnitudes of left- and right- circularly-polarized (c) light 90° out of phase. In optically active molecules, the of left- and right- circularly polarized light will not have equal magnitudes producing elliptical-polarized light (d). Part of the figure was adapted from [http://hyperphysics.phyastr. gsu.edu/hbase/hframe.html]

Light can be polarized into a single plane using a polarizer. For linearly-*polarized light*, the electric vector oscillates in a single plane perpendicular to direction of propagation (Figure 2.6b) and repeats sinusoidally as a function of time with frequency v, and in distance with wavelength λ . The electric vector of a linearly polarized light can be represented as the sum of equal magnitudes of left- and right-*circularly-polarized light* 90° out of phase (Figure 2.6c). When linear-polarized light passes through an optically active solution, the magnitudes of left- and right-circularly polarized light will not have equal magnitudes producing *elliptical-polarized light* (Figure 2.6d) [47]. Optically active molecules are chiral, molecules whose mirror image is non-superimposable. The simplest example of a chiral molecule is a carbon center that is tetrahedrally bonded to four different atoms or groups. [47]. For a three-dimensional CD wave illustration you may visit the site:

http://www.enzims.hu/~szia/cddemo/edemo0.htm

CD involves the absorption difference (ΔA) of left- and right- circularly polarized light at a given wavelength and is defined by:

$$\Delta A = A_L - A_R$$
<div align="right">Equation 2.1</div>

It can also be expressed, by applying Beer's Law as:

$$\Delta A = (\varepsilon_L - \varepsilon_R) \cdot C \cdot l$$
<div align="right">Equation 2.2</div>

where ε is the extinction coefficient (mol·cm/L), C is the molar concentration (mol/L), and l is the cuvette pathlength in centimeters (cm).

Even though commercial instruments measure the difference in absorption of left- and right- circularly polarized light, for historical reasons the measurement in molar ellipticity is reported. Molar ellipticity (θ_{obs}) is related to the tangent of

ellipticity and is related to the difference in absorbance by:

$$\theta_{obs} = 32.98\Delta A$$

Circular Dichroism of Proteins

Based on the energy of the electronic transitions that dominate a given range, CD spectra of proteins are divided into three regions (Figure 2.7). The far-UV region, between 190 and 250 nm where the peptide backbone contributions dominate; the near-UV region, between 250-300 nm where the aromatic side-chains contribute; and the near-UV visible region, between 300 to 700 nm where extrinsic chromophores contribute [48].

The far-UV region has been extensively used to study protein conformation using the amide bond as the chromophore. The amide bond consists of the nitrogen, carbon, and oxygen atoms (Figure 2.8a), and is a chiral center for all amino acids, except glycine, and thus optically active. The local conformation of a protein produce distinct CD spectra that allow identification of α-helix, β-sheet, random coil, and β-turns. Three electronic transitions are recognized, but note that they may show a shift due to the solvent and environment [48]:

- the n-π* transition centered around 220 nm, in which the non-bonding electrons of the carbonyl oxygen are promoted to the antibonding π orbital.
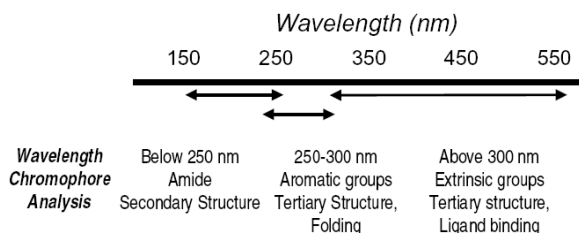
Figure 2.7 CD spectra regions and contributing chromophores.

30

- the π-π* transition centered around 190 nm, in which the π-electrons are delocalized oven the entire peptide fragment.

- below 180 nm, another π-π* transition is present around 160 nm but CD absorbance is difficult to measure in this range.

The intensity and energy of these transitions depends on the angles the peptide bond assumes, namely the φ and ψ angles, which define the secondary structure of the protein (Figure 2.8b). The α-helix spectrum is characterized by the negative band at 222 nm due to the n-π* transition, and a negative and positive couplet at 208 and 190 due to the parallel and perpendicular π-π* transitions, respectively. The splitting of one transition to two is called exciton coupling. The β-sheet spectrum is also
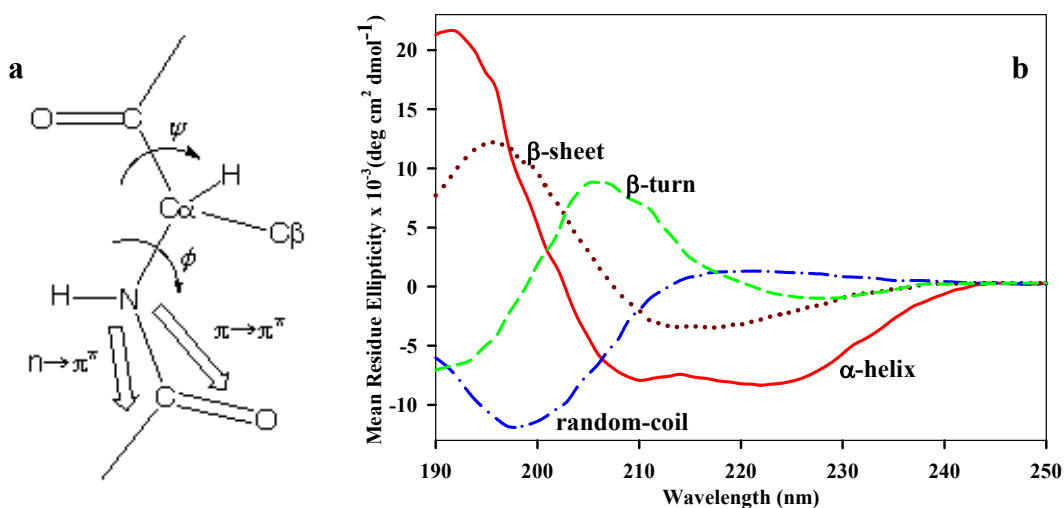
Figure 2.8 The far-UV CD region, 190 to 250 nm, the amide bond is used as the chromophore to study protein secondary structure. The amide bond consists of the nitrogen, carbon, and oxygen atoms, and two electronic transitions have been recognized the n-π* and π-π* (a). Proteins assume different conformations resulting in different φ and ψ angles and produce distinct CD spectra (b): α-helix (red solid line), β-sheet (brown dotted line), random-coil (blue dashed dotted line), and β-turn (green dashed line).

characterized by exciton coupling of the π-π* transition into a negative band near 215 nm and a positive band around 198 nm. The random coil spectrum is characterized by a positive band at 230 nm due to the n-π* transition and a negative band at 195 nm due to π-π* transition. The β-turn spectrum is characterized depending on the type of turn. Type I turns have CD spectra similar to that of α-helix, and Type II turns have CD spectra similar to that of β-sheet but with a 10-15 nm red shift [48]. A summary of the characteristic transition states of secondary structures and the intensity of "pure" secondary structures is summarized in Table 2.2.

Table 2.1   Secondary structures as studied by CD

	Transition	Band (nm)	Intensity $(deg \cdot cm^2 \cdot dmol^{-1})$
α-helix	π-π*	190-195 nm	60,000 to 80,000
	π-π*	208	$-36,000 \pm 3,000$
	n-π*	222	$-36,000 \pm 3,000$
β-sheet	π-π*	195 – 200	30,000 to 50,000
	n-π*	215 - 220	-10,000 to -20,000
random coil	π-π*	200	-20,000
	n-π*	220	small

2.4.2   Materials and Sample Preparation

The purified 2R and 3R tau fragments were used at a concentration of 7 and 11 μM respectively in 20 mM phosphate buffer solution (pH 7.0). Protein concentration was measured as described in section 2.3. Chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

2.4.3 Instrumentation

All CD spectra were collected using a Jasco J-180 spectropolarimeter equipped with a Peltier Thermoset. Spectra were recorded in continuous scanning mode in the range of 190 to 250 nm at 298 K using a 1 mm quartz cuvette. The scanning speed was 10 nm/min with a 16 seconds response time and a 2 nm bandwidth. TFE titration was carried out starting from three initial TFE concentrations (0%, 65%, and 85%). Titration of TFE was carried out with a ~3% stepwise increment of TFE percentage for a total of 31 steps. Data collection was repeated two times for 2R and three times for 3R and data were averaged. The spectra were converted into ASCII format using the Jasco software.

2.4.4 Method and Rationale for CD Spectra Analysis

All data were analyzed with in-house programs written in Matlab code (The MathWorks Inc, Natick, MA) and graphed using SigmaPlot (Systat Software, Point Richmond, CA). Spectral units were obtained in units of millidegrees θ_{obs} and subtracted from the buffer background spectra. Results are then expressed in units of mean residue Ellipticity θ in units of deg·cm^2·dmol^{-1} and were determined using the following equation: $\theta = \theta_{obs}/10 \cdot l \cdot c \cdot n$ where l is the cuvette pathlength in cm, c is the molar concentration, and n is the number of peptide bonds. Finally, θ was corrected for the dilution factor during TFE titration. To obtain the secondary structure populations, global singular value decomposition (SVD) was applied. SVD is a widely used technique to decompose a data matrix into several component matrices, exposing many of the useful and interesting properties of the original

matrix. In our global SVD calculation, a 61x62 θ matrix was used, which has a dimensionality of 61 rows for each wavelength (190 to 250 nm) and 62 columns for each 31 TFE concentrations for 2R and 3R. The θ matrix is decomposed into three matrices U, S, and V so that $\theta = U \times S \times V$. Matrix U is an orthogonal matrix that has columns that resemble the data as a function of wavelength, S is a diagonal matrix that contains the singular values (useful components) arranged in descending order, and V is an orthogonal matrix which contains the amplitude coefficients. The percentage of α-helix, β-sheet, and random-coil structure content in the average spectra of our fragments were calculated using the higher rank components obtained by SVD.

2.5 Fourier Transform Infrared (FTIR) Spectroscopy

2.5.1 Theory

Infrared (IR) spectroscopy is used to identify molecular components and structures based on the vibrations of atoms or molecules. IR radiation spans the 0.78 to 1000 μm (13000 to 10 cm^{-1}) electromagnetic spectrum, and based upon the wavenumber it can be categorized into three regions: far-infrared (4 ~ 400 cm^{-1}), mid-infrared (400 ~ 4,000 cm^{-1}), and near-infrared (4,000 ~ 13,000 cm^{-1}) [48].

A molecule, with n number of atoms, is in continuous vibration with respect to each other for a total of 3n degrees of freedom. Three degrees of freedom are translational, three or two degrees of freedom are rotational for linear and non-linear molecules respectively, and 3n-6 or 3n-5 degrees of freedom are vibrational for linear and non-linear molecules respectively. There are two major types of molecular vibrations, *stretching* involving changes in the inter-atomic distance along the bond axis, and *bending* motions involving changes in the angle between two bonds (Figure 2.9). Additionally interactions between vibrational modes (coupling) can occur if the vibrating bonds are joined by a single atom [48].
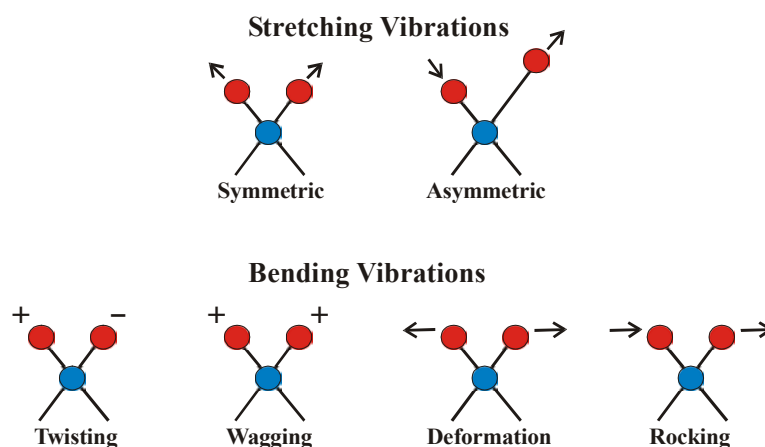

Figure 2.9   The stretching and bending vibrational modes in molecules.

To approximate the vibration frequency of two atoms held together by a bond Hooke's law can be applied. In this approximation, the two atoms and the connecting bonds are treated as a simple harmonic oscillator composed of two masses held together by a spring. The reduced mass (μ) of two atoms can be expressed as:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}$$

Equation 2.4

where m_1 and m_2 are the masses for the two atoms. To obtain the vibration frequency Hooke's law can then by applied:

$$v = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

Equation 2.5

where v is the frequency of vibration and k is the force constant.

The IR spectrum is obtained by passing an IR light through the sample and molecules will absorb radiation only when the incoming IR beam frequency coincides with the molecules' vibration frequency. For vibration to result in absorption of IR light, it must change the dipole moment of the molecule; the larger the change, the more intense the absorption band. The frequency, intensity, and probability of absorption is influenced by both intramolecular factors of the bond strength, mass of atoms, and polarity (dipole moment) of the vibrating bonds, as well as intermolecular factors, of neighboring atoms that alter the electron density in the bonds and affect the vibrational spectrum [49].

Fourier transform infrared (FTIR) spectrophotometers are mostly used nowadays, where all frequencies are examined simultaneously using an interferometer which has replaced the old fashioned dispersive IR instruments, where a monochromator was used. FTIR can be used in different modes: transmission,

specular reflectance, diffuse reflectance, and attenuated total reflectance [50]. The transmission mode is based on the absorption of IR light at specific wavelengths as it passes through the sample (Figure 2.10). The specular reflectance mode is based on IR reflection off of a smooth and mirror-like sample to record its spectrum. The diffuse reflection mode is based on powders and rough surfaces and can be recorded by illuminating these surfaces and collecting the radiation that is scattered at a wide range of angles with the aid of ellipsoidal collection mirrors. Attenuated total reflection (ATR) is based on IR light passing through an infrared transmitting crystal with a high refractive index, usually composed of zinc selenide, or germanium, allowing for the radiation to reflect within the ATR element several times. When a sample is brought into contact with the reflecting surface, it attenuates the beam by absorbing radiation perpendicularly to the surface of the crystal.
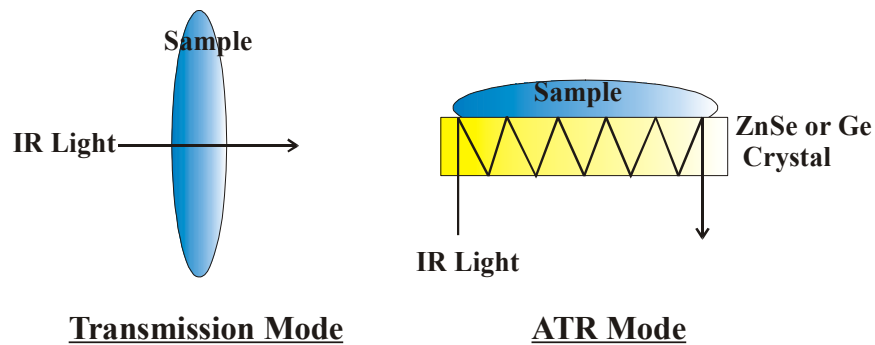
Figure 2.10   Schematics of transmission mode (left image) and attenuated total reflection (ATR) mode (right image). In transmission mode, IR light directly passes through the sample, where as in ATR mode IR light is reflected through a crystal.

Protein FTIR

FTIR spectra of proteins exhibit absorption bands associated with the amide bond (Figure 2.11) and can be used to characterize the secondary structure of proteins. Three amide regions are used to study conformations of proteins. Amide I region, centered around 1650 cm^{-1} with contribution from C=O stretching and minor contributions from C–N stretching and N–H bending; Amide II region centered around 1570 cm^{-1} from contributions of N–H bending and C–N stretching; and Amide III region centered around 1300 cm^{-1} from contributions of C–N stretching, N–H bending, C=O stretching, and O=C–N bending [49, 50]. The amide I region is mostly used due to its strong signal. Proteins have a continuous chain of amino acids connected through amide bonds. The difference in frequency in which amide bonds vibrate can be used to identify different secondary structures in which the amide bond is present and are summarized in Table 2.2. The difference of vibration frequency is due to different hydrogen bonding among amino acids in the polypeptide chain.

Table 2.2. Characteristic amide I band protein frequencies [44]

Frequency (cm^{-1})	Assignment
1621-1640	
1671-1679	β-structure
1641-1647	Random coil
1651-1657	α-helix
1658-1671	
1681-1696	Turns and bends
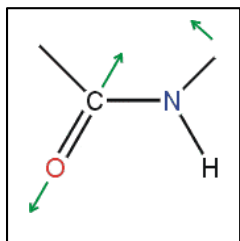
Figure 2.11 FTIR spectra of proteins exhibit absorption bands associated with the amide bond

2.5.2 Materials and Sample Preparation

For the ATR spectra, the purified 3R was used at a concentration of 9 μM in 20 mM phosphate buffer solution (pH 7.0). Protein concentration was determined as described in section 2.3.

For the transmission spectra, stock protein solution was prepared by dissolving the lyophilized protein powder in deuterium water and then was divided into eppendorf tubes for lyophilization. Lyophilized powder was dissolved in 20 mM sodium phosphate buffer (pH 6.6), prepared with deuterium water, for a 200 μM final protein concentration.

2.5.3 Instrumentation

All FTIR spectra were collected using an Excalibur FTS 3000 spectrometer from MX-BIORAD.

ATR spectra were recorded at room temperature using 1024 interferograms at 4 cm^{-1} resolution. The experiment was carried out at 9μM 3R starting from eleven TFE concentrations (0, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95% TFE). Spectra were collected using nitrogen purge and corrected for background absorbance. Spectra were obtained by applying 8 μL of the sample into the surface of the clean zinc selenide ATR crystal and solvent was evaporated using a low purge of nitrogen in the chamber to leave a thin-film residue. Application of sample was repeated for a total of four times. If smaller quantity of protein was used, spectra were too noisy for any structural determination.

Transmission spectra were recorded at room temperature using 124 interferograms at 2 cm^{-1} resolution. The experiment was carried out at a concentration of 200 μM 3R starting from eleven TFE concentrations (0, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95% TFE). Spectra were collected using nitrogen purge and corrected for background absorbance. Spectra were obtained by applying 12 μL of the sample between two calcium fluoride glasses separated by a 0.05 μm thick Teflon spacer. After five minutes in the chamber to remove any moisture, spectra were collected. If lower protein concentration was used, spectra were too noisy for any structural contained determination.

2.5.4 Method and Rationale for FTIR Spectra Analysis

All data were analyzed with in-house programs written in Matlab code (The MathWorks Inc, Natick, MA) and graphed using SigmaPlot (Systat Software, Point Richmond, CA). Amide I was used for structural determination. The amide I band was broad, due to band overlapping, and bands were resolved by calculating the second derivative of the spectrum or by deconvoluting the amide I band into a series of Gaussian bands as a function of wavenumber $P(x)$ in the form:

$$P(x) = \frac{1}{\sigma\sqrt{2\cdot\pi}} e^{\frac{-(x-\mu)^2}{2\cdot\sigma^2}}$$ where σ is the standard deviation, μ is the mean.

2.6   Atomic Force Microscopy (AFM)

2.6.1   Theory

The Atomic Force Microscope (AFM) also known as Scanning Force Microscope (SFM) or Scanning Probe Microscope (SPM) is designed to measure the topography of a nonconductive sample. The advantage of AFM compared to light or electron microscopes is that images can be generated in atomic resolution by "feeling" the sample rather than "looking" and in natural conditions without the usage of coatings used in EM which tend to alter the morphology of the specimen [51]. Since its initial development as an imaging technique AFM has undergone several augmentations and has been used to study a wide range of processes among them: protein folding-unfolding transitions and enzyme-substrate function. The information presented here pertains to the simplest use of the AFM - obtaining an image of a sample - but the general principles apply to most applications.

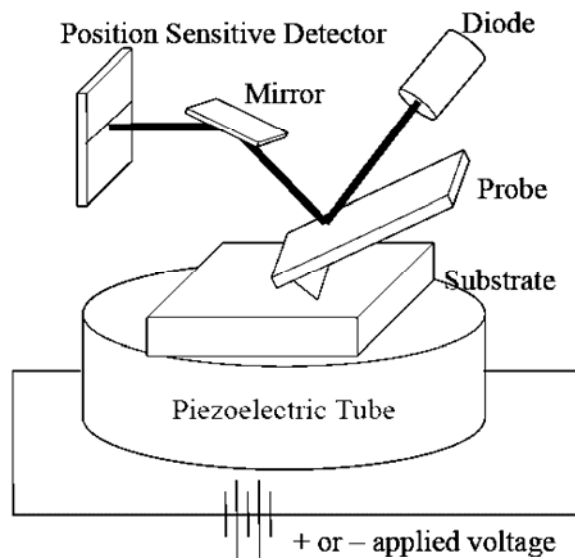The use of AFM to image specimens in atomic resolution is obtained by

Figure 2.12   Schematic representation of the atomic force microscope (AFM).

scanning the specimen with a probe and detecting the vertical displacement of the cantilever (Figure 2.12). A laser beam is focused at the end of the top surface of a probe and reflected to a position sensitive photodetector where a change in angle of reflection can be monitored and recorded. Depending on the version of the AFM, either the specimen is rastered under the probe or visa versa, using a piezoelectric scanner. As the tip moves in response to the topology of the specimen its change in position is recorded by the photodetector. The feedback from a sensor maintains the probe at a constant force, i.e. constant distance from the sample surface. The piezoelectric ceramic moves the specimen in the Cartesian coordinates (x, y, or z) precisely at angstrom resolution by altering its AC voltage potential. Rastering the scanner across the sample, produces a three-dimensional topographic map of the sample surface based on deflections of the cantilever, and thus the height of the topographic surface can be measured at a nanometer scale [51].

Above, a basic description was given of how AFM works where the acquisition of an image depends on the separation between the tip and the specimen. The intermolecular forces driving the process are:

Van der waals forces due to distribution of electric charge between the sample and the tip. The force between the tip and the specimen can be characterized by a force-distance relationship known as a pair-potential energy $E^{pair}(r)$ function and can be modeled with a Lennard-Jones function:

$$E^{pair}(r) = 4 \cdot \varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \qquad \text{Equation 2.6}$$
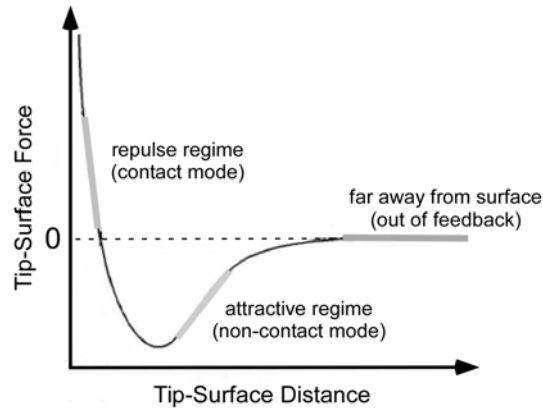
Figure 2.13   Schematic diagram of Lennard-Jones potential. At small tip-surface distance a repulsive force is experienced and at large tip-surface distance an attractive force is experienced. Based on the force interaction, the AFM can be operated in different modes: Contact mode in the repulsive regime, non-contact mode in the attractive regime, and tapping mode in both regimes. Figure was adapted from: [http://www.science.siu.edu/chemistry/zang/image/AFM-force.JPG].

where ε and σ are constants and depend on the material. Atoms repel or attract each other depending on their separation. The 12 power term accounts for the repulsive term during small probe-sample separation and the 6 power term accounts for the attractive force during large probe-sample separation (Figure 2.13).

Electrostatic forces present in ionic bonds, with two charges, q_1 and q_2. The force between them can be represented with Coulomb force law:

$$F = \frac{1}{4 \cdot \pi \cdot \varepsilon_o} \cdot \frac{q_1 \cdot q_2}{r^2}$$
                                                                       Equation 2.7

where ε_o depends on the medium and r is the distance of separation between the two charges. Specifically, at short distance two opposite charges are attracted to each other but eventually, the outer shell electrons around each ion interact and the force becomes repulsive.

Capillary and adhesive forces due to the condensation of water vapor on the specimen and tip when imaging in air. The tip will be pulled down to the sample by a strong liquid meniscus giving rise to capillary forces which "glue" the tip to the specimen. This can be corrected by adjusting the feedback control.

Double layer force when imaging in liquid where the mica can attract oppositely charged ions from the solution. This can be corrected by using a high ionic strength buffer.

The AFM can be operated in different modes depending on the forces of interaction between the probe and the specimen (Figure 2.13). Generally, the AFM works in three modes [51]:

- Contact mode in which the tip is in contact with the specimen as it is raster-scanned experiencing a repulsive force. Here, the value of imaging force is pre-set and when the tip encounters a sample feature its force charges. The feedback control applies a voltage to the piezo to raise or lower the sample relative to the cantilever to restore the desired value of force. The voltage that the feedback amplifier applies to the piezo is a measure of the height of features on the sample surface.

- Non-contact mode in which the tip is in close distance to the sample, but not touching it, experiencing an attractive force. Here, the value of frequency at which the cantilever is oscillated is pre-set, and when the tip encounters a sample feature its oscillating frequency changes. The feedback control applies a voltage to the piezo to raise or lower the sample relative to the cantilever to

restore the desired value of oscillating frequency. The voltage that the feedback amplifier applies to the piezo is a measure of the height of features on the sample surface.

- Tapping mode in which the cantilever is oscillated above the sample experiencing both repulsive and attractive forces. Here, the cantilever is oscillated at a pre-set amplitude above the specimen and when the tip interacts with a surface feature, its amplitude is changed. The feedback control applies a voltage to the the piezo to raise or lower the sample relative to the cantilever to restore the desired amplitude of oscillation. The voltage that the feedback amplifier applies to the piezo is a measure of the height of features on the sample surface. The tapping mode is the most commonly used mode for examining biological specimens because is not detrimental to the sample.

2.6.2 Materials and Sample Preparation

The end point samples after the CD experiment were imaged at 63, 87, and 94 % TFE and at ~ 3 μM 3R concentration. Also, freshly prepared samples at 0, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95% TFE and at 9 μM 3R concentration were also prepared. Samples were prepared by deposition of 5 μL of sample on freshly cleaved muscovite mica and incubated for 30 seconds at room temperature. The mica substrate was tilted and rinsed twice with 1 mL purified water, then dried with filtered compressed tetrafluoroethane (Dust Blaster, Kensington).

3.3.2 Instrumentation

AFM was operated in TappingMode^{TM} using a Nanoscope IIIa controller system (Digital Instruments, Santa Barbara CA) with a J-series vertical engage scanner. AFM probes were single-crystal silicon microcantilevers with 300 kHz resonant frequency and 40 newton/meter spring constant (model MPP-11200, Veeco Probes, Santa Barbara, CA). Images were obtained using a resolution of 512 lines, 1 Hz scanning speed, the drive amplitude setpoint was kept to a minimum, and image size varied.

2.6.4 Method and Rationale for AFM Image Analysis

Images were flattened using the AFM software and their characteristic height and periodicity was obtained with in house program written in Matlab code (The MathWorks Inc, Natick, MA) and graphed using SigmaPlot (Systat Software, Point Richmond, CA).

3.   RESULTS

3.1   Conformational Analysis of 2R and 3R Using Far-UV Circular Dichroism (CD) Spectroscopy

To investigate the conformational properties of 2R and 3R, the far-UV CD spectra of 2R and 3R were recorded with the increment of TFE concentration from 0 to 94% by titration. The experiment was carried out starting at a concentration of 7 μM for 3R and 11 μM for 2R. Prior to the CD experiments, we confirmed that our selected starting concentrations were free of fibrillar or amorphous aggregates by AFM (data not shown). The native conformation of 2R and 3R in sodium phosphate buffer (pH 7.0) is mostly random coil, which is characterized by the negative maximum at 198 nm, with a small content of α-helix conformation (Figure 3.1). This is consistent with a previously published analysis of full-length tau and the individual repeat regions [34, 37]. The α-helical content increases steadily with the stepwise
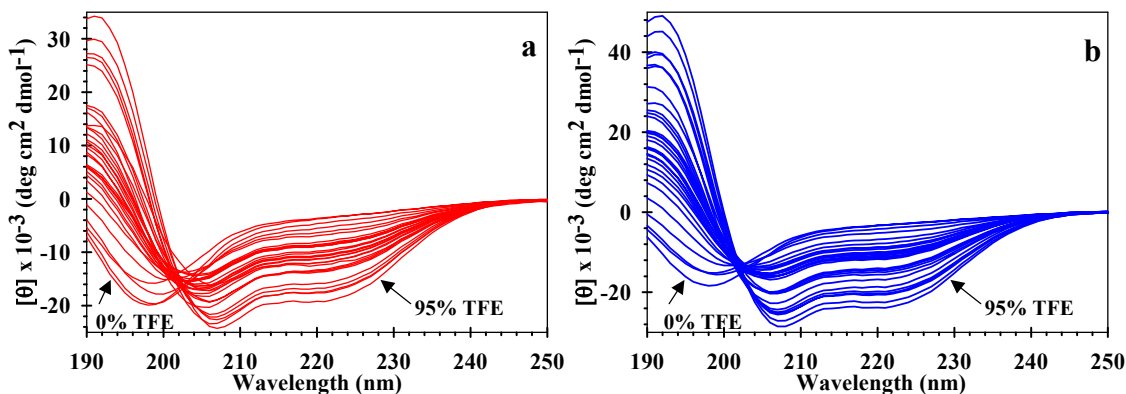

Figure 3.1. Mean molar Ellipticity values (θ) obtained by far-UV CD spectra of 2R (a) and 3R (b) as a function of TFE. The lowest and the highest TFE concentrations are indicated by arrows.

increase in TFE concentration, characterized by the double negative maxima at 208 and 222 nm and a positive maximum at 192 nm.

The average spectra for 2R and 3R as a function of TFE concentration do not exhibit an isodichroic point, signifying that the conformational transition from random coil to an α-helix is a heterogeneous process involving more than one phase. Figure 3.2 records the experimentally observed values at 222 nm as a function of TFE concentration. It is seen that the mean residue ellipticity values change in the direction toward an α-helix through two transitions. The first derivative of the spectra indicates that the first transition occurs between 0-45% TFE, and the second transition between 45-94% TFE for both 2R and 3R (Figure 3.2. insert) suggesting that either the α-helical structure is stabilized through two transitions, or another conformation appears at intermediate TFE concentration. In order to obtain the secondary structure contents from the far-UV CD spectra, singular value decomposition (SVD) was applied.

SVD produces three main components (Figure 3.3a). The first component, which corresponds to the average of all the data, is characteristic of an α-helix. The second component is characteristic of a random-coil spectrum. The noisy third component exhibits a broad positive maximum at 200 nm and a broad negative maximum at 220-230 nm and can be loosely assigned either to a β-turn or β-sheet conformation, or even to the characteristic length dependence of the α-helix spectrum. The CD spectra of a β-sheet is characterized by a positive maximum near 198 nm and a negative maximum at 218 nm, whereas a β-turn spectrum is similar to the β-sheet spectrum but with a 5-10 nm red shift [48], thus it is difficult to distinguish them in a noisy spectrum. The fourth component shows an anticorrelated

spectrum to the third component suggesting that possibly β-turn isomerization could contribute to the spectra. After the fourth component, all other components were just noise. The populations for α-helix, β-sheet, and random coil were calculated from the SVD components (Figure 3.3b). The native conformation of 2R and 3R is mostly random coil with a 20% α-helical character. At the highest TFE concentration (94% TFE), 2R is ~52% and 3R is ~82% helical respectively. No indication of a β-sheet conformation was observed in the entire TFE concentrations range. The α-helical content can be alternatively calculated by the signal ratio at wavelengths 193 and 222 nm. We obtained values of -1.65, -2.00 corresponding to an α-helical content of 64% for 2R and 78% for 3R, which are consistent with values obtained by SVD.

The conversion of 3R and 2R from random coil to α-helix is clearly characterized by two transitions. The first transition spanning 0-45% TFE, 2R converts from 20 to 27% α-helix and 3R from 20 to 32% α-helix. In the second transition spanning 45-94% TFE, 2R converts from 27 to 52% α-helix and 3R from 32 to 82% α-helix. First, these results indicate that the main α-helical propensity in tau is located in MBD, as it was previously reported that full length tau contains ~ 60% at the highest TFE concentration [34, 37]. Thus elimination of the N- and C-terminal domains, which are also known to restrain *in vitro* aggregation of full length tau, simplifies the spectroscopic properties of the protein and allows *in vitro* aggregation study. In order to further examine the tau conformations, we further examined the conformational properties by FTIR and the conformational-assembly properties by AFM.
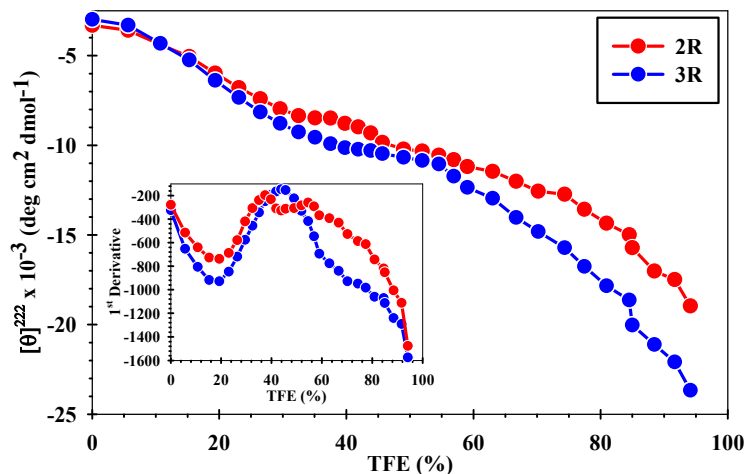
Figure 3.2 Mean molar Ellipticity values (θ) at 222 nm for 3R (blue line) and 2R (red line). The insert shows the 1st derivative data at 222 nm for 3R (blue line) and 2R (red line). Data indicate two α-helix stabilization phases: Phase I: 0-45% and Phase II: 45-95 %TFE.

Figure 3.3 (a) Singular Value Decomposition (SVD) components of the far-UV CD spectra. 1st component is characteristic to α-helix (red line), 2nd component to random coil (green line), and the 3rd component to β-turn or β-sheet (cyan line). (b) Secondary structure populations of 2R (triangles) and 3R (circles) for the three components.

3.2 Conformational Analysis of 3R Using Fourier Transform Infrared Spectroscopy (FTIR) in Attenuated Total Reflectance (ATR) mode

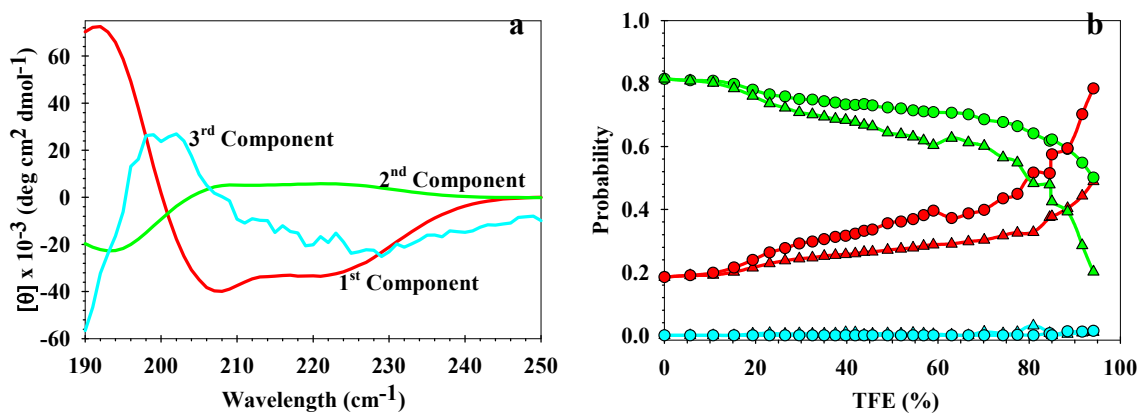The ATR experiment of 3R was carried out at eleven TFE concentrations (0, 5, 15, 25, 35, 45, 55, 65, 75, 85, and 95%) and at 9 μM concentration. The amide I spectra are presented in Figure 3.4a The amide I region ranges from 1600 to 1700 cm^{-1} reporting stretching vibrations from the backbone C=O, bending of the N–H bond, and stretching of the C–N bond. It is the most widely used region in characterizing protein secondary structure in addition to the amide II and III regions.
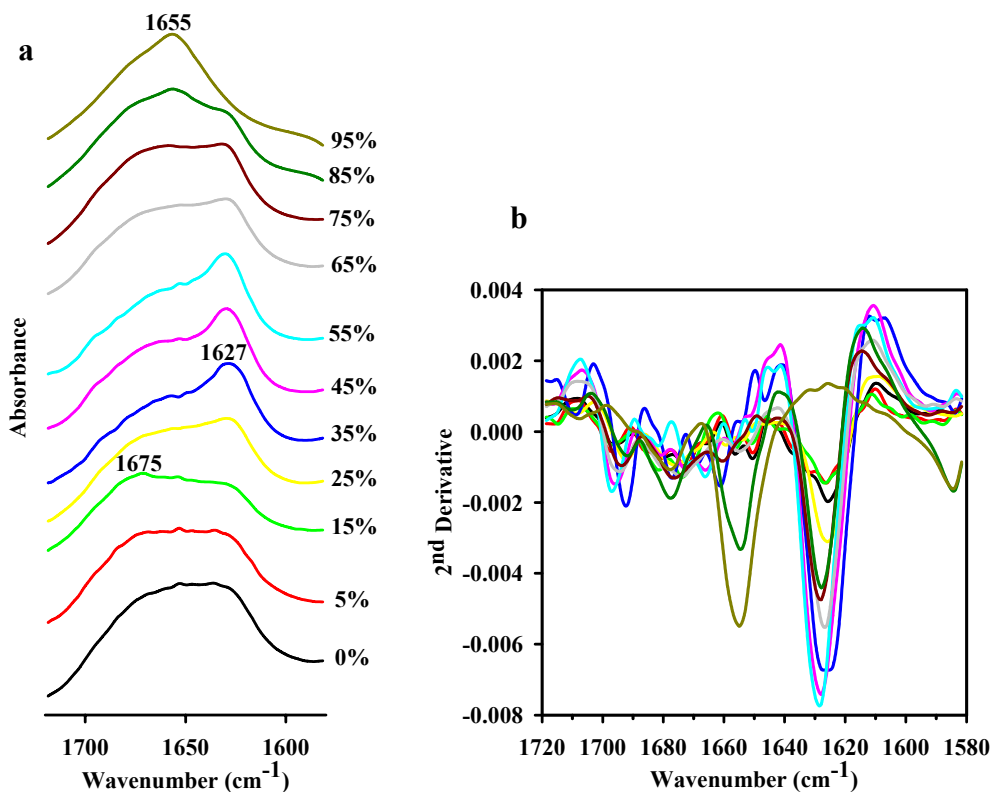

Figure 3.4 The Amide I spectra of 3R at eleven TFE concentrations (a) and the second derivative of the spectra (b). Line colors in (a) correspond to line colors in (b).

At a first glance, the spectra appear broad and featureless due to overlapping of bands, but their differences become more apparent by calculating the second derivative of the spectra (Figure 3.4b) or deconvoluting the spectra to several Gaussian bands (Figure 3.5).

The native conformation of 3R in sodium phosphate buffer (pH 7.0) up to 25% TFE is composed of a combination of three bands at 1627 cm^{-1}, 1655 cm^{-1}, and 1676 cm^{-1} corresponding to β-sheet, α-helix, and β-turn respectively. Upon further increment of TFE concentration up to 55%, the bands at 1627 cm^{-1} and 1675 cm^{-1} are more prominent. An increase of TFE concentration beyond 55%, results in the 1627 cm^{-1} and 1675 cm^{-1} bands to descend and the 1655 cm^{-1} band to ascend. At the final TFE concentration of 95%, the band at 1655 cm^{-1} dominates the spectrum.

The results obtained by ATR are inconsistent with our previous results obtained by CD spectroscopy. Here, the native conformation of 3R is a combination of α-helix, β-sheet, and β-turn as opposed to the mostly random coil observed by CD. Also, there seems to be β-sheet conformation at all TFE concentrations by ATR FTIR. Interestingly, the stabilization of the β-sheet population is mainly observed in the 35 to 55% TFE range. This optimal TFE concentration coincides with previously identified amyloid stimulating conditions and has been previously reported for amyloid formation in tau under the presence of heparin [37]. In addition to the increase of β-sheet signal, we observed increases in the β-turn signal, which is consistent with the third and fourth SVD components obtained in the CD experiment. SVD analysis of the ATR spectra (Figure 3.5a), suggests three main components. The first component is the average spectral signal, the second component corresponds to

β-sheet with the characteristic wavenumber maximum at 1627 cm^{-1}, and the third

component to α-helix with the characteristic maximum wavenumber at 1655 cm^{-1}.

All other components were noise. The amplitudes of the SVD components (Figure

3.5b) suggest that the α-helix and β-sheet signals increase until 55% TFE. At higher

TFE concentrations, the β-sheet signal decreases while the α-helix continues

increasing. The increment in β-sheet conformation could be explained by the

experimental procedure in the ATR experiment. In ATR FTIR, soluble protein is

slowly deposited on the ATR crystal by dissolution until a protein film is formed.

This process involves a progressive increment of protein concentration upon

dehydration, which could lead to non-specific protein-backbone interactions assisted

by the interaction with the crystal surface of the ATR. To further investigate the

conformation of 3R and how specific were the signals obtained here, we employed

FTIR in transmission mode.
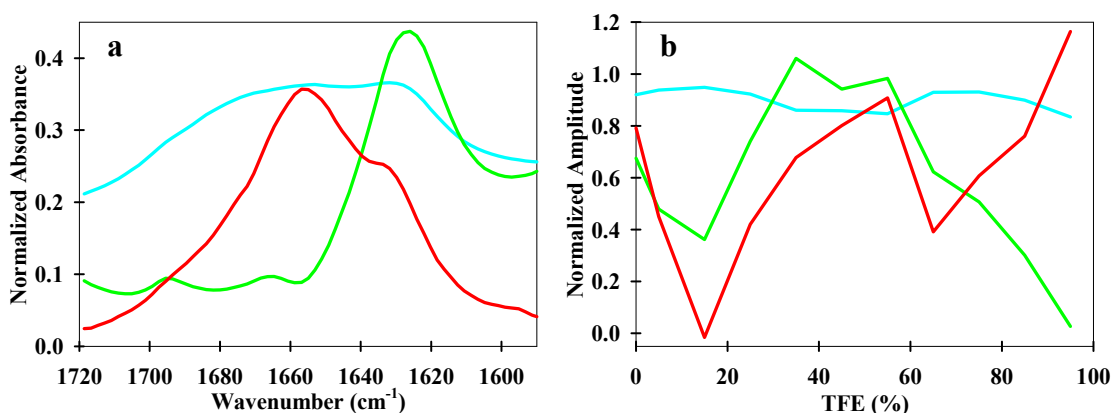
Figure 3.5 Singular Value Decomposition (SVD) of the ATR data. Three main components were obtained the 1st component is the average signal (blue line), the 2nd component corresponds to the β-sheet signal (green line), and the 3rd component to the α-helix signal (red line) (a). The amplitudes for the components is shown in (b).

Figure 3.6 Gaussian fits (red circles) of the ATR spectra (black line) at eleven TFE concentrations indicate three conformational contributions at 1627 cm[-1] (cyan line), 1655 cm[-1] (magenta line) and 1675 cm[-1] (green line) corresponding to β-sheet, α-helix, and, β-turn respectively.

3.3 Conformational analysis of 3R using Fourier Transform Infrared Spectroscopy (FTIR) in Transmission mode

To further investigate the conformation of 3R, we collected the FTIR spectra in transmission mode at eleven TFE concentrations (0, 5, 15, 25, 35, 45, 55, 65, 75, 85, and 95%) and at 200 µM concentration. The amide I spectra are shown in Figure 3.6a. Here, we observed two main band contributions at 1675 and 1645 cm[-1] corresponding the β-turn and α-helical structure in ^2H_2O. By calculating the second derivative of the amide I spectra (Figure 3.7b) and deconvoluting the spectra to a series of Gaussian bands (Figure 3.8), we detect the structural interplay of α-helix and β-turn in the range of TFE concentrations. From zero up to 35% TFE, the spectra
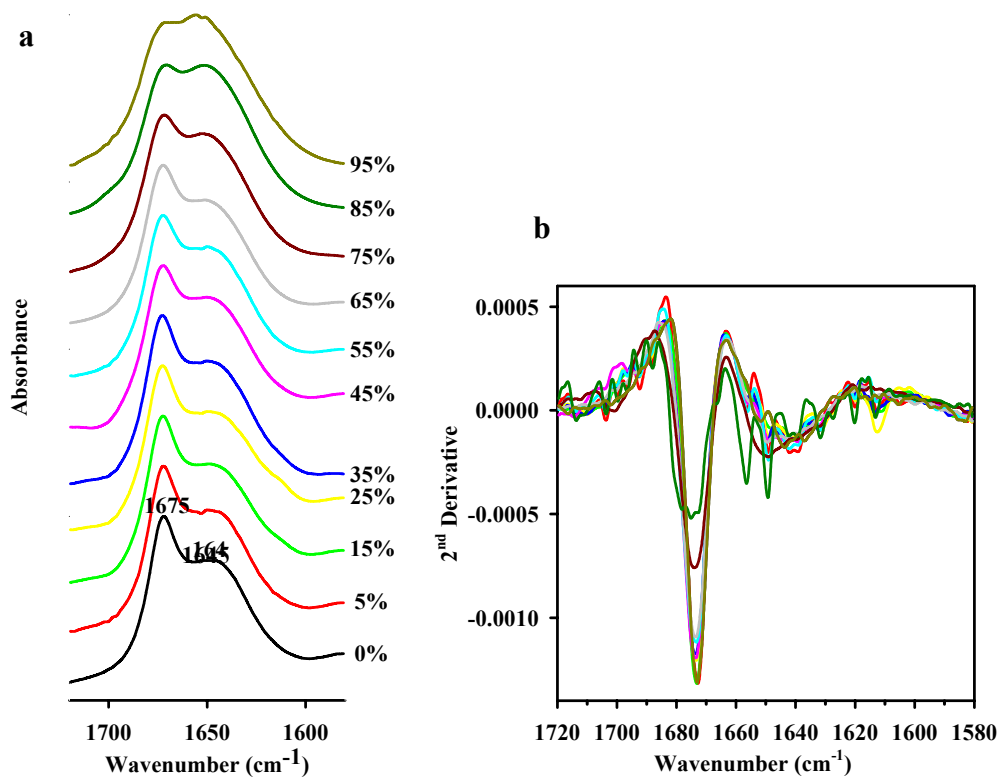
Figure 3.7 The Amide I spectra of 3R at eleven TFE concentrations (a) and the second derivative of the spectra (b). Line colors in (a) correspond to line colors in (b).

consist of a combination of β-turn and α-helix. Upon further increase of TFE concentration we observed an increase in absorbance for both β-turn and α-helix. At TFE concentrations higher than 55% the absorbance of the characteristic α-helix band increases and the 1675 cm^{-1} corresponding to turns decreases. At the highest TFE concentration of 94%, the α-helical structure dominates the spectra and two mimima appear at 1649 and 1657 cm^{-1}.

In Transmission mode we observed that fitting the amide I band into a series of Gaussian equations all the data can be fitted into three Gaussians, which indicates a less complex system for the ATR spectra. The bands correspond to α-helix, β-turn, and a higher wavenumber band. Here, we used even higher sample concentration than the ATR mode (200 μM versus 9 μM in ATR), but we did not observe a β-sheet stabilization as observed by ATR, indicating that the β-sheet component observed in ATR mode is non-specific and probably induced by the slow deposition into a dry film. ATR and Transmission modes have different properties: A transmission measurement collects a spectrum that is an average of the bulk properties of the sample; on the other hand, ATR measurements only measure the surface properties of the deposited sample. These results indicate that non-specific interaction with the crystal surface and film deposition can perturb the conformational properties of natively unfolded proteins.
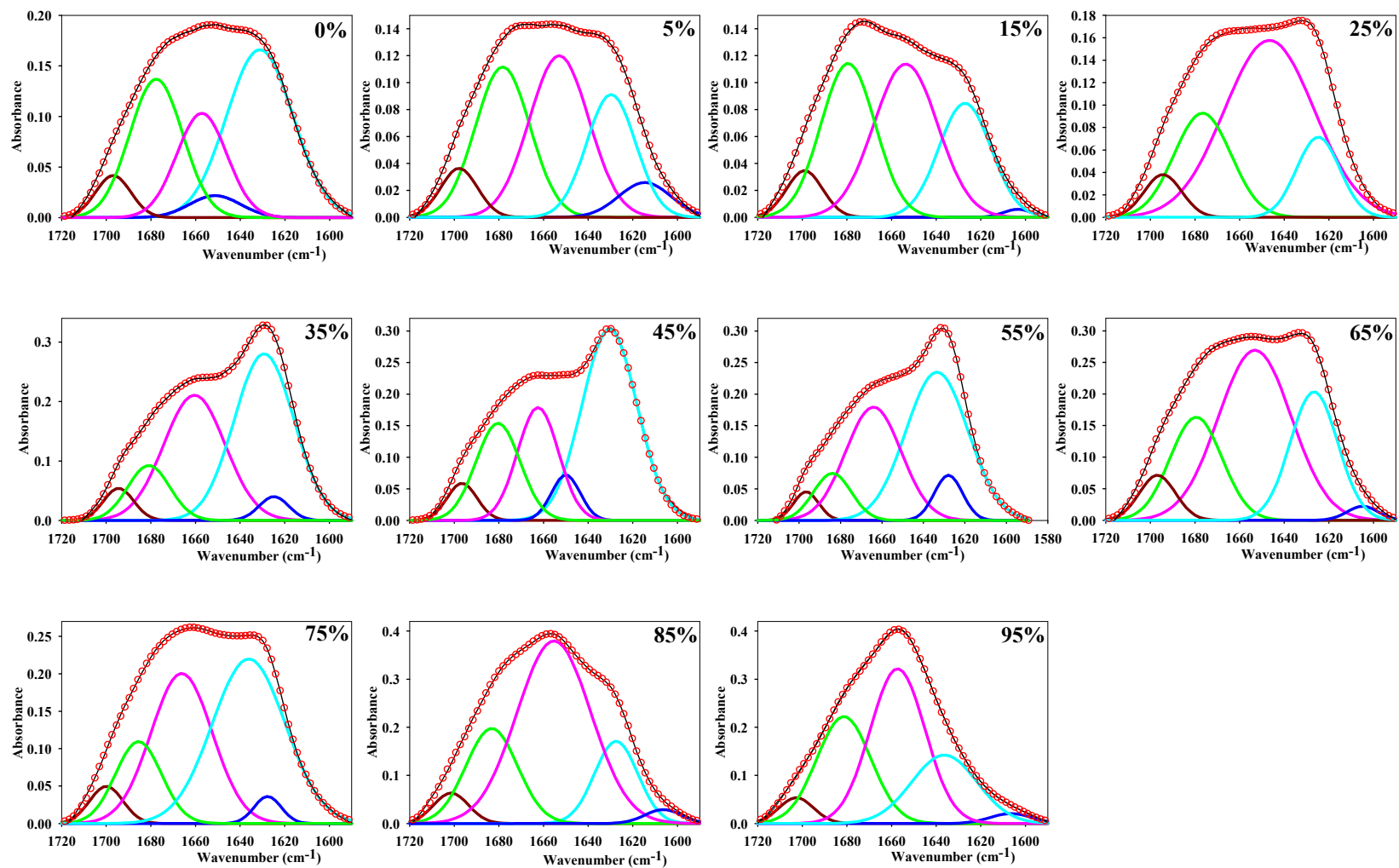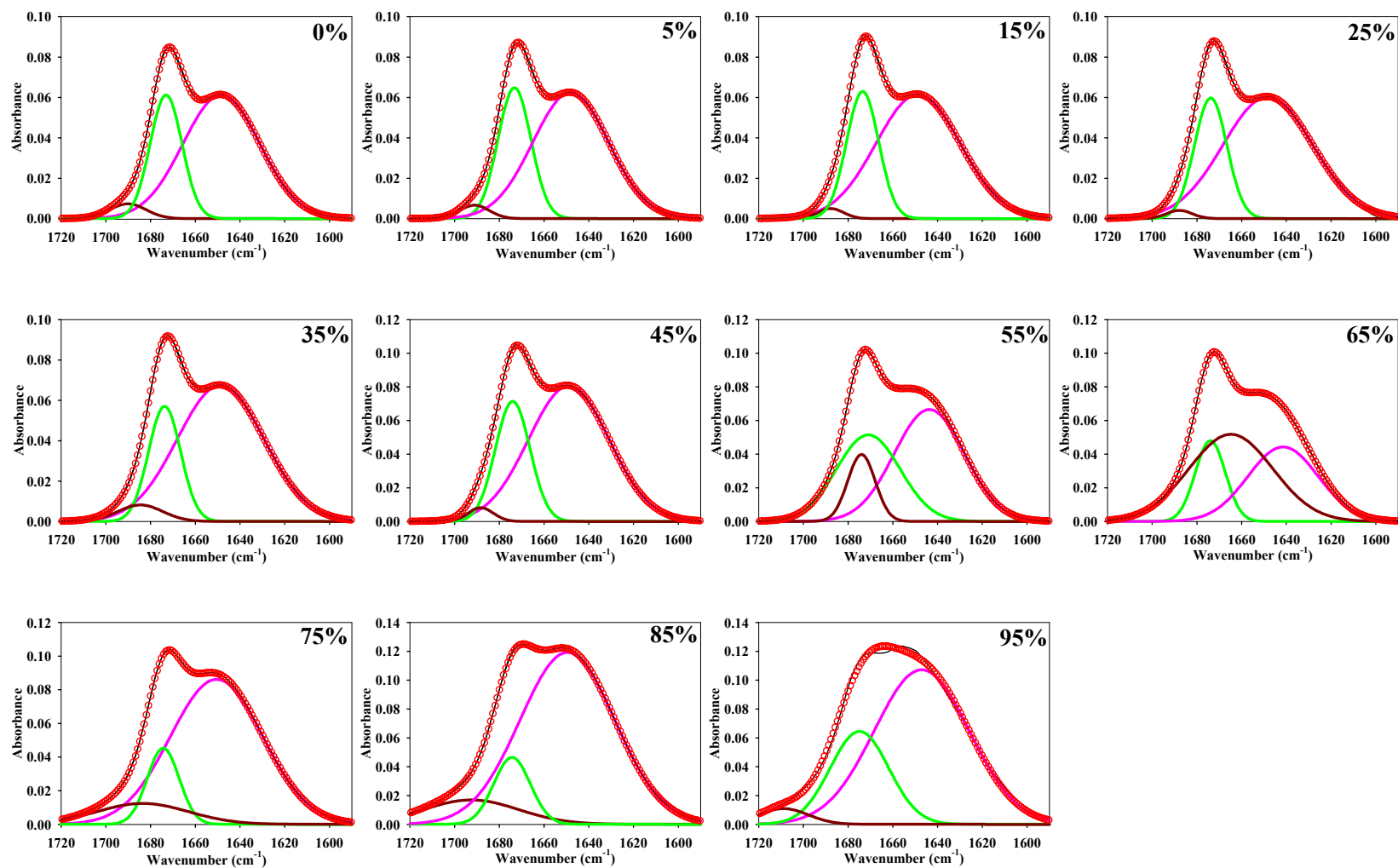
Figure 3.8 Gaussian fits (red circles) of the Transmission spectra (black line) at eleven TFE concentrations indicate two structural contributions at 1645 cm^{-1} (magenta line) and 1675 cm^{-1} (green line) corresponding to α-helix and β-turn respectively.

3.4   Conformation and Assembly Analysis of 3R using Structure

Prediction Programs

To combine the obtained experimental conformation data with theoretical structural and protein aggregation predictions, we used two algorithms: TANGO and COILS.

TANGO is a statistical mechanics algorithm designed to predict protein conformation as well as protein aggregation prone regions [52]. The propensity of conformations for 3R was simulated at 20 mM sodium phosphate buffer (pH 7.0). The TANGO score predicts two predominant structures: α-helix and β-turn with a zero probability for β-sheet (Figure 3.9), in agreement with the obtained CD and FTIR in transmission mode data. The highest α-helical character is localized in the intrarepeat regions for all of the repeats and near the N-terminus of the 1st repeat. These results suggest that β-turns could be an important component in the intrarepeat
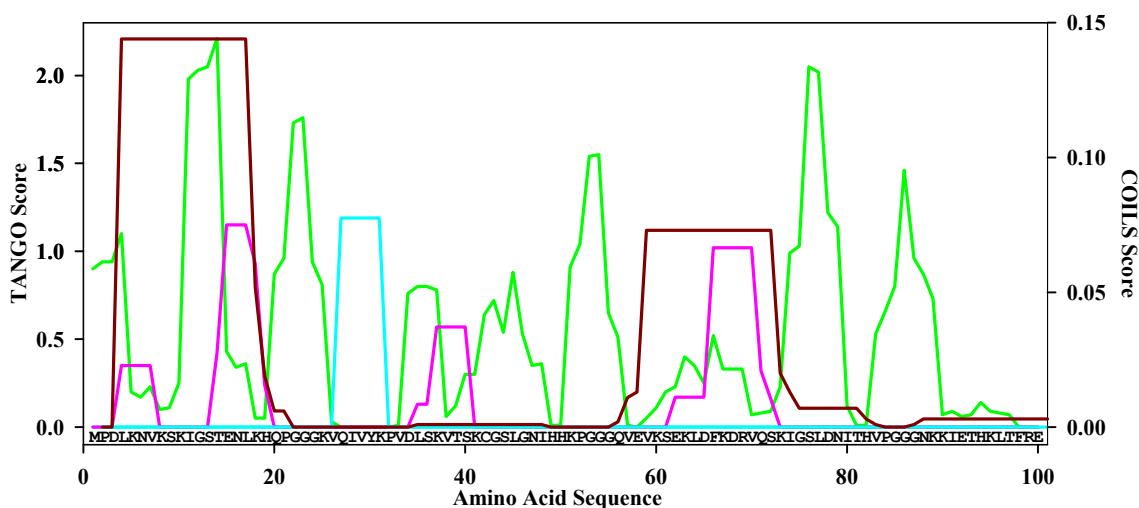
Figure 3.9 Structure and protein aggregation predictions calculated by TANGO and COILS. β-turn (green line), α-helix (magenta line), β-amyloid (cyan line), coiled-coils (brown line).

regions (PGGG) of the conformation of 2R and 3R. This is not surprising. Previous NMR experiments have indeed suggested a *trans* orientation in the Lys311-Pro312 and Lys331-Pro332 [19]. Other studies have suggested the existence of β-turns in the Ser316-Asn317, as well as the typical turn PGGG structure in the C-terminus of all the four individual repeats. Additionally, TANGO predicts a β-sheet aggregation region in the previously identified "hexapeptide region" (^{305}VQIVYK^{311}) but did not indicate α-helical aggregation regions. Alternatively, we used COILS to identify α-helical aggregation regions. COILS is a algorithm that calculates the probability that a protein sequence will adopt a coiled-coil conformation based on sequence similarity with known sequences of parallel two-stranded coiled-coils [53]. COILS indicated a propensity for 3R in the first and fourth repeat for formation of homodimeric coiled coils.

3.5   Assembly analysis of 3R using Atomic Force Microscopy (AFM)

TFE Induces 3R Fibrillization. In our CD and FTIR conformation data we observed a conformational transition from random coil with β-turns into α-helix. To investigate the conformational-assembly interplay of 3R, the endpoint samples after CD were imaged using AFM. Endpoint samples were collected after the far-UV CD titration experiment at 59, 85, and 94% TFE and at ~ 3 μM concentration.

In all three samples, we observed fibrillar aggregates with some PHFs-like characteristics (Figure 3.10-3.12). The morphology of fibrils at 63% TFE show two morphologies (Figure 3.10). A twisted ribbon structure, with a crude periodicity of 50-80 nm and heights ranging from 10 nm at the highest point to 5 nm at the lowest point (image 1a and 3a); and a straight structure, with no regular periodicity and heights ranging from 3 to 5 nm (image 2b and 3b). Lengths for the observed fibrils ranged from 100 to 250 nm. With the increase of TFE concentration at 85%, we observed bundling of three or four fibrils into fibrillar entities (Figure 3.11). The non-assembled fibrils are characterized by heights around 8 nm and with (image 3b) or without PHFs like periodicity (image 1a and 2a); the associated fibrils are characterized by heights up to 25-30 nm at the highest point with no regular periodicity (image 1b, 2b, and 3a). At the highest TFE concentration, 94%, we mostly observed "mature" fibrils with no apparent periodicity and heights ranging from 20 to 30 nm (Figure 3.12).

The 3R *in vitro* aggregation results obtained herein in combination with structural information from spectroscopic techniques show that aggregation is linked to the nature of the secondary structure induced in the presence of the solvent.

However, images were obtained in selected areas in the mica and only represent a small fraction of the total protein population. Conformational change to an α-helix conformation seems to stabilize filament formation and with the progressive increment of α-helix, we also observed bundling of individual filaments. These results contradict previous publications suggesting that aggregation of the microtubule repeat region is little affected by TFE [37] and that assembly only proceeds in the 10-30% TFE only by utilizing heparin. Here, we observed that not only is TFE capable of promoting tau assembly, but there is a TFE concentration dependence of the morphology of the aggregates.
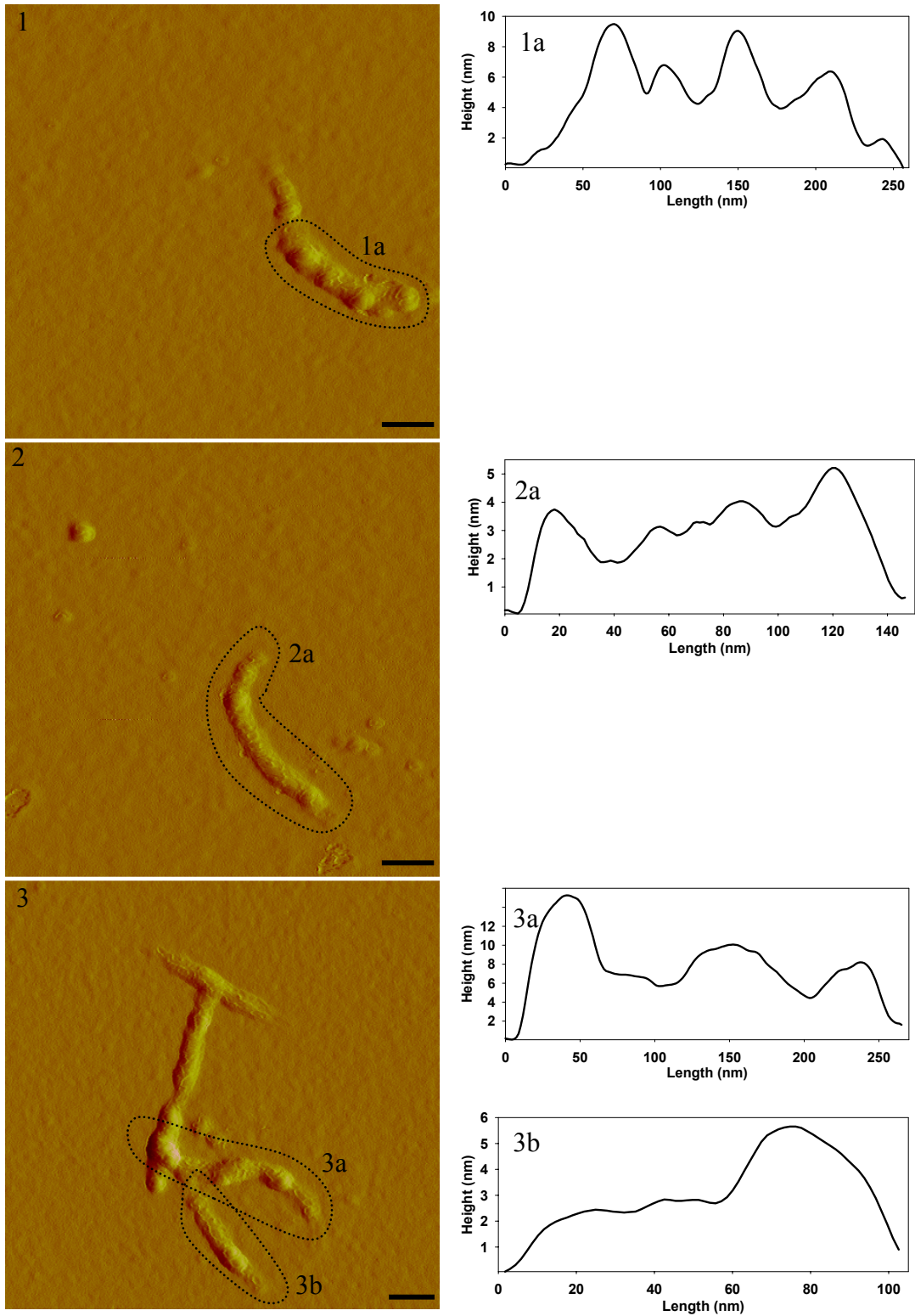
Figure 3.10   *In vitro* aggregation of 3R at 59% TFE. The black bar is 75 nm long.
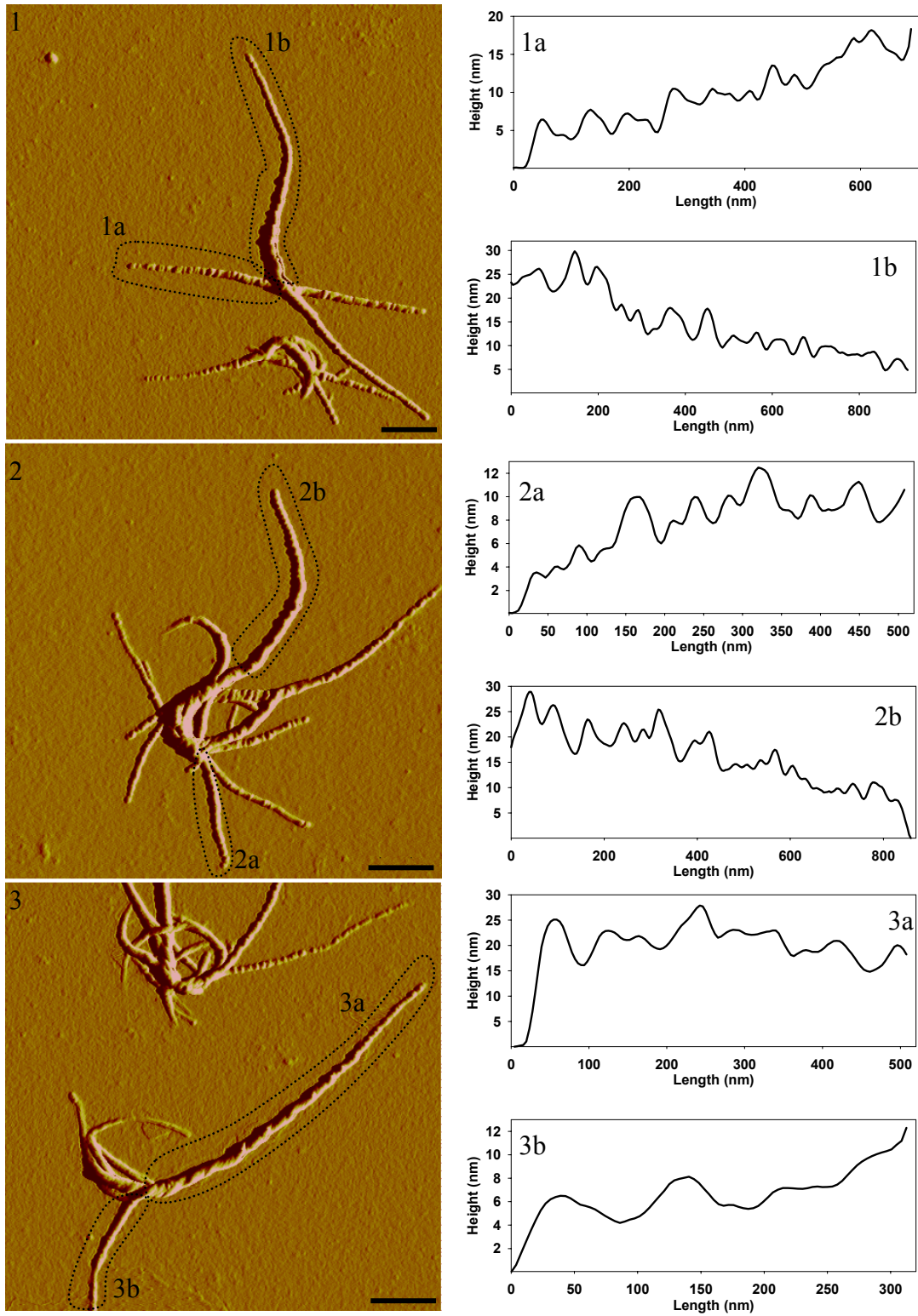
Figure 3.11 In vitro aggregation of 3R at 85% TFE. The black bar is 300 nm long.
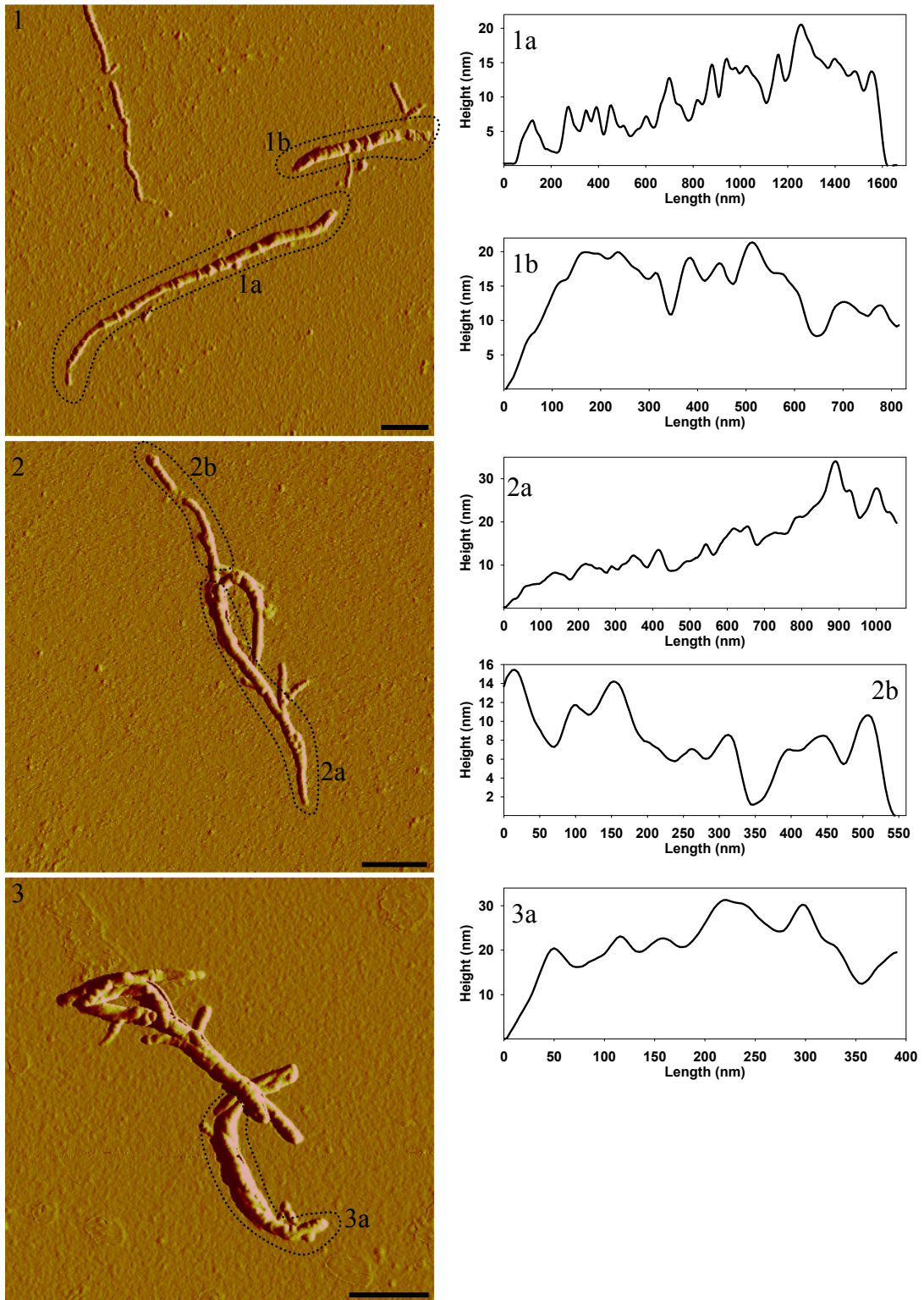
Figure 3.12 In vitro aggregation of 3R at 94% TFE. The black bar is 300 nm long.

TFE induces 3R fibrillization but is irreproducible. In contrast to the above section, where *in vitro* fibrils were observed with PHFs-like characteristics, the experiment was repeated numerous times but it was irreproducible. Instead of the PHFs-like fibrils, we observed amorphous aggregates instead. In attempts to find optimal fibril formation conditions, we first examined if the starting stock concentration played a role. We started from a low 47 µM or high 667 µM 3R concentration stock sample and we diluted down to 7 µM in different TFE concentrations from 0 to 95%. Samples were imaged at day 0, day 1, and day 10 but no fibrils were observed (data not shown). Second, we examined if procedure of preparation samples preparation played a role. We prepared the 7 µM final sample by both titration of TFE to the final concentration and by adding the protein directly to the final TFE concentration. Samples were imaged at day 0, day 1, and day 10 but no fibrils were observed (data not shown).

4. DISCUSSION

Tau is a highly hydrophilic and natively unfolded protein mainly expressed in neurons [13, 16]. The biological role of tau is to stabilize microtubule formation by binding through its MBD region. In certain pathological situations, tau undergoes self-assembly that can result in the formation of aberrant aggregates that are toxic to neurons. The MBD region is known to form the core of the aggregated filament [30]. This process occurs in a number of neurological disorders collectively known as tauopathies, the most commonly recognized of which is Alzheimer's disease. Here, we studied the conformational and assembly properties of two tau fragments, 2R and 3R, which contain the MBD region.

Tau conformation and spectroscopy studies: In our spectroscopic studies, we used TFE as a means to enhance α-helix formation in 2R and 3R and mimic the possible structure upon binding to microtubules. We observed a change from random coil to α-helix conformation in both 2R and 3R. Significant populations of β-turns also appear at intermediate conditions in the titration experiment and these results were obtained using a battery of spectroscopic techniques. Conformational information was resolved, revealing different conformational traits with different techniques. Using CD we observed a transition from mostly random coil to almost 100% of an α-helix conformation that spans the whole sequence. The formation of TFE-induced α-helix displays two clear regimes. The first regime corresponds to a TFE concentration between 0-45% TFE. The second regime goes from 45-94% TFE. Using ATR-FTIR we observe a transition from a heterogeneous mixture of secondary

structures to a highly α-helical conformation, again in a two-step process. The first process corresponds to 0-55% TFE and exhibits an increased content in α-helix and β-sheet. The second step is similar to the one observed by CD, 55-95% TFE, also corresponding to the stabilization of α-helix conformation. Using Transmission-FTIR we observe a transition from a mostly random coil ensemble with small populations of β-turns to an α-helix conformation that occurs again in two steps. Between 0-55% TFE both α-helix and β-turns increase in population, while beyond 55% only α-helix is stabilized. These differences in secondary structure content may be due to the different sensitivity of the spectroscopic techniques used here, which monitor different properties of the molecule. CD spectroscopy is known to be more sensitive to the α-helix conformation. However, an unambiguous identification from CD data of heterogeneous structural ensembles, especially those containing β-sheet and turns, presents significant challenges [48]. Using ATR-FTIR we found a distinctly higher content in β-sheet like structures that seem to arise from the film deposition process [50]. Besides, FTIR of the amide I band results in high spectral overlap, making the deconvolution of experimental spectra difficult [49]. Our spectroscopic results highlight the importance of employing several spectroscopic techniques to study protein conformation in heterogeneous samples.

Tau and its role as a natively unfolded protein: Tau belongs to a relatively new and growing group of proteins classified as "natively unfolded proteins." In standard globular proteins, the folded conformation, which determines the function of the protein, is unique and inherently stable [2, 4]. Three-dimensional structures of globular proteins are stabilized by non-covalent interactions such as hydrogen

bonding, electrostatics, van der Waals, and hydrophobic interactions [2, 4]. In contrast, natively unfolded proteins lack a defined structure in physiological conditions. The absence of regular structure in natively unfolded proteins has raised questions regarding the structure-function paradigm. Unfolded proteins exist as biologically disordered polypeptides and undergo a disorder-to-order transition upon binding to their target to gain function [6]. In general, natively unfolded proteins or domains are characterized by high overall hydrophilicity and net charge [5]. A more recent realization places natively unfolded proteins as not being completely unstructured. In fact, these proteins seem to have a high propensity to become structured upon binding to their biological target [5]. The target is supposed to provide the energy to fold by serving as a template for structure formation. The results obtained here, indicate a high propensity for the MBD regions of tau to become α-helical in presence of the hydrogen-bonding promoting agent TFE. Under native conditions, in 20 mM sodium phosphate buffer (pH 7.0), 2R and 3R have a 20% α-helical content, which is significantly higher than the α-helical content measured in the full length protein, indicating that alpha helical conformation tends to be localized in the repeat regions [34, 37]. *In vivo*, tau binds to the acidic helical C-terminus of α- and β- tubulin through its MBD. Our results strongly support the idea that tau binds to microtubules in an α-helical conformation, thus following the behavior expected for a natively unfolded protein. The comparison between constructs with only two or three adjacent repeats shows a simple additive behavior with respect to α-helix formation. This observation suggests formation of α-helices independently of each other, in agreement with the expectation for a protein in which

69

repeats bind their cellular targets independently. We also identified high turn content in the MBD regions of tau. This local conformation of the polypeptide may also play a recognition role in tau binding to microtubules, and may facilitate the formation of exposed loops that can be phosphorylated, since phosphorylation is important for tau function and is also known to play a role in tau's pathological assembly. Furthermore, hyperphosphorylation of the MBD domain has been previously identified as a microtubule destabilizing factor. However, the exact role that tau phosphorylation plays in promoting self-assembly is currently a topic of much controversy and debate [21, 54].

Tau and assembly: In this project, our initial motivation was to study the assembly properties of tau utilizing 2R and 3R. 2R and 3R have been indicated as the important regions of full length tau to form the core of PHFs [30]. In previous studies that obtained fibrillar aggregates from mutated full length tau, protein samples required long incubation time [34]. To accelerate the fibrillization process, we were limited to the techniques we could use. One procedure involves using high protein concentration, but this is known to induce non-specific β-sheet aggregates [32]. Alternatively, the procedure employed was, where we used centrifugation through filters to periodically increase protein concentration, but filters commercially available could not withstand high TFE concentrations. We obtained fibrillar aggregates at 59, 85, and 94% TFE after the CD experiment, but these results were irreproducible. The *in vitro* grown fibrils were characterized by AFM and were found to have PHFs-like characteristics with a crude periodicity of 80 nm. We also observed that as the α-helical content is increased, individual tau protofilaments bundle to form

mature filaments. The irreproducibility of the assembly experiment could be due to the conditions we used and could indicate some specificity in α-helical fibrillar formation. Specificity lies in the side chain-side chain interactions to stabilize aggregation. Reproducibility may be obtained by using hyperphosphorylated tau or by mutations.

5.   CONCLUSIONS

Tau is an interesting protein to study for its role in stabilizing microtubule assembly and for its implication in diseases by fibrillization. The results of this study suggest that tau stabilizes microtubules through α-helix induction and as other studies also suggest, fibrillization may proceed by inducing α-helical backbone conformation [12, 18, 34]. Here, TFE was used as the cosolvent for stabilization of conformation; *in vivo* microtubules provide the stabilization energy for α-helix, and aggregation may also be enhanced by post-translational modifications such as phosphorylation and glycosylation [21]. Our observation of *in vitro* grown filaments with PHFs-like characteristics further supports the idea that PHFs are an atypical fibrillar aggregate, stabilized through α-helices. To further study the conformational and assembly properties of tau, modifications of the polypeptide such as phosphorylation and glycosylation could be used to enhance the tau assembly process.

GLOSSARY

AD   Alzheimer's Disease

AFM   Atomic Force Microscopy

ATR   Attenuated Total Reflection

BCA   Bicinchoninic Acid Assay

CD   Circular Dichroism

FTIR   Fourier Transform Infrared Spectroscopy

HPLC   High Performance Liquid Chromatography

IPTG   Isopropyl-β-D-thiogalactoside

IR   Infrared Spectroscopy

MALDI-TOF   Matrix Assisted Laser Desorption/Ionization - Time of Flight

MAP   Microtubule Associated Protein

MBD   Microtubule Binding Domain

NMR   Nuclear Magnetic Resonance

PHF   Paired Helical Filament

2R   Repeat 1 and 3 of tau microtubule-binding domain

3R   Repeat 1, 3, and 4 of tau microtubule binding domain

RP   Reverse Phase

SDS-PAGE   Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SF   Straight Filament

SVD   Singular Value Decomposition

TFE   2,2,2-Trifluoroethanol

BIBLIOGRAPHY

1. Anfinsen, C.B., *Principles That Govern Folding of Protein Chains.* Science, 1973. **181**(4096): p. 223-230.

2. Garrett-Grisham, *Biochemistry*. Second ed. 1998, Orlando, FL: Harcourt College Publishers.

3. Dill, K.A., *Dominant Forces in Protein Folding.* Biochemistry, 1990. **29**(31): p. 7133-7155.

4. Stryer, *Biochemistry*. Fifth ed. 2002, New York, NY: W. H. Freeman and Company.

5. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins-Structure Function and Genetics, 2000. **41**(3): p. 415-427.

6. Uversky, V.N., *What does it mean to be natively unfolded?* European Journal of Biochemistry, 2002. **269**(1): p. 2-12.

7. Dobson, C.M., *Experimental investigation of protein folding and misfolding.* Methods, 2004. **34**(1): p. 4-14.

8. Chiti, F., et al., *Designing conditions for in vitro formation of amyloid protofilaments and fibrils.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(7): p. 3590-3594.

9. Dobson, C.M., *Protein misfolding, evolution and disease.* Trends in Biochemical Sciences, 1999. **24**(9): p. 329-332.

10.   Blake, C. and L. Serpell, *Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous beta-sheet helix.* Structure, 1996. **4**(8): p. 989-998.

11.   Sunde, M., et al., *Common core structure of amyloid fibrils by synchrotron X-ray diffraction.* Journal of Molecular Biology, 1997. **273**(3): p. 729-739.

12.   Sadqi, M., et al., *alpha-helix structure in Alzheimer's disease aggregates of tau-protein.* Biochemistry, 2002. **41**(22): p. 7150-7155.

13.   Weingarten, M.D., et al., *A Protein Factor Essential for Microtubule Assembly.* PNAS, 1975. **72**(5): p. 1858-1862.

14.   Cleveland, D.W., S.Y. Hwo, and M.W. Kirschner, *Purification of Tau, a Microtubule-Associated Protein That Induces Assembly of Microtubules from Purified Tubulin.* Journal of Molecular Biology, 1977. **116**(2): p. 207-225.

15.   Noetzel, T.L., et al., *A comparison of the ability of XMAP215 and tau to inhibit the microtubule destabilizing activity of XKCM1.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1455): p. 591-594.

16.   Cleveland, D.W., S.Y. Hwo, and M.W. Kirschner, *Physical and Chemical Properties of Purified Tau Factor and Role of Tau in Microtubule Assembly.* Journal of Molecular Biology, 1977. **116**(2): p. 227-247.

17.   Schweers, O., et al., *Structural Studies of Tau-Protein and Alzheimer Paired Helical Filaments Show No Evidence for Beta-Structure.* Journal of Biological Chemistry, 1994. **269**(39): p. 24290-24297.

18.   Goux, W.J., *The conformations of filamentous and soluble tau associated with Alzheimer paired helical filaments.* Biochemistry, 2002. **41**(46): p. 13798-13806.

19.   Minoura, K., et al., *Amphipathic helical behavior of the third repeat fragment in the tau microtubule-binding domain, studied by H-1 NMR spectroscopy.* Biochemical and Biophysical Research Communications, 2002. **294**(2): p. 210-214.

20.   Arriagada, P.V., et al., *Neurofibrillary Tangles but Not Senile Plaques Parallel Duration and Severity of Alzheimers-Disease.* Neurology, 1992. **42**(3): p. 631-639.

21.   Avila, J., et al., *Role of Tau Protein in Both Physiological and Pathological Conditions.* Physiol. Rev., 2004. **84**(2): p. 361-384.

22.   Couchie, D., et al., *Primary Structure of High Molecular Weight Tau Present in the Peripheral Nervous System.* PNAS, 1992. **89**(10): p. 4378-4381.

23.   Greenberg, S.G. and P. Davies, *A Preparation of Alzheimer Paired Helical Filaments That Displays Distinct Tau-Proteins by Polyacrylamide-Gel Electrophoresis.* Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(15): p. 5827-5831.

24.   Yao, T.M., et al., *Aggregation analysis of the microtubule binding domain in tau protein by spectroscopic methods.* Journal of Biochemistry, 2003. **134**(1): p. 91-99.

25.   von Bergen, M., et al., *Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming*

beta structure. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(10): p. 5129-5134.

26.   Crowther, R.A., *Straight and Paired Helical Filaments in Alzheimer-Disease Have a Common Structural Unit.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(6): p. 2288-2292.

27.   Perez, M., et al., *In vitro assembly of tau protein: Mapping the regions involved in filament formation.* Biochemistry, 2001. **40**(20): p. 5983-5991.

28.   Berry, R.W., et al., *Inhibition of tau polymerization by its carboxy-terminal caspase cleavage fragment.* Biochemistry, 2003. **42**(27): p. 8325-8331.

29.   Kidd, M., *Paired Helical Filaments in Electron Microscopy of Alzheimers Disease.* Nature, 1963. **197**(486): p. 192-.

30.   Wischik, C.M., et al., *Structural Characterization of the Core of the Paired Helical Filament of Alzheimer-Disease.* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(13): p. 4884-4888.

31.   Kirschner, D.A., C. Abraham, and D.J. Selkoe, *X-Ray-Diffraction from Intraneuronal Paired Helical Filaments and Extraneuronal Amyloid Fibers in Alzheimer-Disease Indicates Cross-Beta Conformation.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(2): p. 503-507.

32.   Berriman, J., et al., *Tau filaments from human brain and from in vitro assembly of recombinant protein show cross-beta structure.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(15): p. 9034-9038.

33.   Barghorn, S., P. Davies, and E. Mandelkow, *Tau paired helical filaments from Alzheimer's disease brain and assembled in vitro are based on beta-structure in the core domain.* Biochemistry, 2004. **43**(6): p. 1694-1703.

34.   Kunjithapatham, R., et al., *Role for the alpha-helix in aberrant protein aggregation.* Biochemistry, 2005. **44**(1): p. 149-156.

35.   Inouye, H., et al., *Structure of core domain of fibril-forming PHF/Tau fragments.* Biophysical Journal, 2006. **90**(5): p. 1774-1789.

36.   Eliezer, D., et al., *Residual structure in the repeat domain of tau: Echoes of microtubule binding and paired helical filament formation.* Biochemistry, 2005. **44**(3): p. 1026-1036.

37.   Hiraoka, S., et al., *Conformational transition state is responsible for assembly of microtubule-binding domain of tau protein.* Biochemical and Biophysical Research Communications, 2004. **315**(3): p. 659-663.

38.   Minoura, K., et al., *Different associational and conformational behaviors between the second and third repeat fragments in the tau microtubule-binding domain.* European Journal of Biochemistry, 2004. **271**(3): p. 545-552.

39.   von Bergen, M., et al., *Tau aggregation is driven by a transition from random coil to beta sheet structure.* Biochimica Et Biophysica Acta-Molecular Basis of Disease, 2005. **1739**(2-3): p. 158-166.

40.   Giannetti, A.M., et al., *Fibers of tau fragments, but not full length tau, exhibit a cross beta-structure: Implications for the formation of paired helical filaments.* Protein Science, 2000. **9**(12): p. 2427-2435.

41. Minoura, K., et al., *Solvent-dependent conformation of the third repeat fragment in the microtubule-binding domain of tau protein, analyzed by H-1-NMR spectroscopy and molecular modeling calculation.* Bulletin of the Chemical Society of Japan, 2003. **76**(8): p. 1617-1624.

42. Andrew, C.D., et al., *Effect of phosphorylation on alpha-helix stability as a function of position.* Biochemistry, 2002. **41**(6): p. 1897-1905.

43. Munishkina, L.A., et al., *Conformational behavior and aggregation of alpha-synuclein in organic solvents: Modeling the effects of membranes.* Biochemistry, 2003. **42**(9): p. 2720-2730.

44. Sibille, N., et al., *Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy.* Biochemistry, 2006. **45**(41): p. 12560-12572.

45. Wilson, D.M. and L.I. Binder, *Polymerization of Microtubule-associated Protein Tau under Near-physiological Conditions.* J. Biol. Chem., 1995. **270**(41): p. 24306-24314.

46. Mukrasch, M.D., et al., *Sites of Tau Important for Aggregation Populate {beta}-Structure and Bind to Microtubules and Polyanions.* J. Biol. Chem., 2005. **280**(26): p. 24978-24986.

47. Tinoco, *Physical Chemistry: Principles and Applications In Biological Sciences.* Third Edition ed. 1995, New Jersey, USA: Prentice Hall.

48. Fasman, G., *Circular Dichroism and the Conformational Analysis of Biomolecules.* 1996, New York, NY: Plenum Press.

49. Stuart, B., *Biological applications of infrared spectroscopy*, ed. A. David. 1997, London, England: John Wiley and Sons.

50.   Singh, B.R., *Infrared analysis of peptides and proteins*, ed. S.B. Ram. 2000, Washington, DC: Americal Chemical Society.

51.   Gunning, *Atomic Force Microscopy for Biologists*. 1999, London, UK: Imperial College Press.

52.   Fernandez-Escamilla, A.M., et al., *Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins.* Nature Biotechnology, 2004. **22**(10): p. 1302-1306.

53.   Lupas, A., M. Vandyke, and J. Stock, *Predicting Coiled Coils from Protein Sequences.* Science, 1991. **252**(5009): p. 1162-1164.

54.   Mendieta, J., et al., *Phosphorylation modulates the alpha-helical structure and polymerization of a peptide from the third tau microtubule-binding repeat.* Biochimica Et Biophysica Acta-General Subjects, 2005. **1721**(1-3): p. 16-26.