Synthetic Phosphopeptides and Phosphoproteins as Tools for Studying Peptide and Protein Self-assembly

Dissertation zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

MALGORZATA BRONCEL
aus Glubczyce, Polen

Juli, 2011

1. Gutachter: Prof. Dr. Christian Hackenberger (Freie Universität Berlin)

2. Gutachterin: Prof. Dr. Beate Koksch (Freie Universität Berlin)

Disputation am: July 19, 2011

Declaration

The work presented here was carried out in the research groups of Prof. Dr. Christian Hackenberger and Prof. Dr. Beate Koksch from April 2008 until Jun 2011 at the Institute of Chemistry and Biochemistry in the Department of Biology, Chemistry, and Pharmacy of Freie Universität Berlin.

I herewith confirm that I have prepared this dissertation without the help of any impermissible resources. All citations are marked as such. The present thesis has neither been accepted in any previous doctorate degree procedure nor has it been evaluated as insufficient.

Berlin, June 2011 Malgorzata Broncel

The work on this dissertation resulted so far in the following publications:

- C. Smet-Nocca, M. Broncel, J. M. Wieruszeski, C. Tokarski, X. Hanoulle, A. Leroy, I. Landrieu, C. Rolando, G. Lippens, C. P. R. Hackenberger, Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation, *Mol. Biosyst.* **2011**, *7*, 1420.

- M. Mühlberg, D.M.M. Jaradat, R. Kleineweischede, I. Papp, D. Dechtrirat, S. Muth, M. Broncel, C.P.R. Hackenberger, Acidic and basic deprotection strategies of borane-protected phosphinothioesters for the traceless Staudinger ligation, *Bioorg. Med. Chem.* **2010**, *18*, 3679.

- M. Broncel, J. A. Falenski, S. C. Wagner, C. P. R. Hackenberger, B. Koksch, How post- translational modifications influence amyloid formation: a systematic study of phosphorylation and glycosylation in model peptides, *Chem. Eur. J.* **2010**, *16*, 7881.

- M. Broncel, S. C. Wagner, K. Paul, C. P. R. Hackenberger, B. Koksch, Towards understanding secondary structure transitions: phosphorylation and metal coordination in model peptides, *Org. Biomol. Chem.* **2010**, *8*, 2575.

- M. Broncel, S. C. Wagner, C. P. R. Hackenberger, B. Koksch, Enzymatically trigerred amyloid formation: an approach for studying peptide aggregation, *Chem. Commun.* **2010**, *46*, 3080.

For data security reasons the curriculum vitae has been omitted from the published version.

For data security reasons the curriculum vitae has been omitted from the published version.

For data security reasons the curriculum vitae has been omitted from the published version.

Acknowledgement

I would like to thank Prof. Dr. Christian Hackenberger and Prof. Dr. Beate Koksch for giving me the opportunity to work on the exciting and very fruitful project. I greatly appreciate all the support and mentoring that I was given as well as the freedom in pursuing my own ideas.

In addition I would like to thank all present and former members of Hackenberger and Koksch research groups for helpful discussions and friendly working atmosphere. In particular, I would like to thank Dr. Mario Salwiczek for his continous efforts to keep our analytical equipment in a 'working mode' and for his help with translating my thoughts into German language. I also greatly appreciate the help of Dr. Allison Berger and Dr. Pamela Winchester with proofreading of my manuscripts as well as this dissertation. Last but not least, I want to thank Raheleh Rezaei Araghi and Dr. Cosimo Cadicamo for being good friends and coffee break companions.

I would like to thank Dr. Dirk Schwarzer (Leibniz-Institut für Molekulare Pharmakologie, Berlin Buch) without whom the implementation of the tau semisynthesis project would not have been possible. I greatly appreciate the opportunity to work in Dr. Schwarzer's laboratory and all the help that I have received from him and his co-workers. In particular, I have to thank Rebecca Klingberg, Bernhard Geltinger, and Oliver Jost for introducing me to the molecular cloning and protein expression.

I am grateful to Mrs. Heike Scheffler and Magnus Krüger from Prof. Dr. Ronald Gust's laboratory (Institut für Pharmazie, Freie Universität Berlin) for their help with atomic absorption spectroscopy measurements.

I need to thank Katrin Wittig and Katharina Tebel for their invaluable help with all administrative matters.

I appreciate the financial support from the Deutsche Forschungsgemeinschaft (SFB 765 - Multivalency as chemical organisation and action principle), the Emmy Noether Programm (HA 4468/2-2), and the Freie Universität Berlin within the Innovationfonds.

Finally, I would like to thank my family for their constant support and great faith in me; my parents for making me their life project and my husband Remi for his patience and standing by my side no matter what.

Referat

Die Aggregation löslicher Proteine zu proteolytisch stabilen Fibrillen ist ein Kennzeichen neurodegenerativer Krankheiten. Dieses Phänomen wurde bei Proteinen beobachtet, die unter bestimmten Bedingungen strukturelle Übergänge von einer ungefalteten oder teilweise helikalen Konformation zu einer β-Faltblattstruktur durchlaufen. Umgebungsfaktoren wie der pH-Wert, die Anwesenheit von Metallionen, die Temperatur sowie die Proteinkonzentration selbst wurden als Auslöser der Proteinaggregation identifiziert. In den letzten zwei Jahrzehnten wurde aber auch die Rolle posttranslationaler Modifikationen wie z.B. der Phosphorylierung in diesem Zusammenhang diskutiert. Die Untersuchung der Rolle der Phosphorylierung in natürlichen Systemen sieht sich jedoch mit einigen schwerwiegenden Problemen konfrontiert. Der schnelle Abbau von Phosphoproteinen *in vivo* macht es beispielsweise schwierig, phosphorylierte Reste genau zu identifizieren. Auf der anderen Seite führt die enzymatische Phosphorylierung *in vitro* oftmals zu heterogenen Produkten mit unterschiedlichem Phosphorylierungsgrad.

Im Rahmen der vorliegenden Arbeit werden einige erfolgreiche Anwendungen synthetischer Phosphopeptide als Modelle zum Studium der Proteinaggregation vorgestellt. Verschiedene Aspekte der Fibrillenbildung wurden mit Hilfe unterschiedlicher analytischer Methoden wie CD-Spektroskopie, Thioflavin-T induzierter Fluoreszenz sowie Transmissionselektronenspektroskopie untersucht. Hierbei wurde z.B. der Einfluss ortsspezifischer und multipler Phosphorylierung auf das Aggregationsverhalten eines Amyloid bildenden Coiled-Coil-Modellpeptids untersucht. Die strukturellen Analysen haben gezeigt, dass die Umfaltung zum β-Faltblatt unabhängig von der Anzahl und Position modifizierter Reste durch die Phosphorylierung vollkommen unterdrückt wird. Sämtliche phosphorylierte Peptide zeigten darüber hinaus keine Neigung zur Bildung vom Amyloiden. Um die Prozesse der Umfaltung und Fibrillenbildung, sowie deren Morphologien unter physiologischen Bedingungen besser verstehen zu können, wurde im Weiteren ein Enzym, das Phosphatgruppen abspaltet, als natürlicher Auslöser der Aggregation eingesetzt. Weiterhin wurden durch Metallionen induzierte konformationelle Übergänge an synthetischen Phosphopeptiden untersucht. Dabei wurde gezeigt, dass die Koordination von Magnesium- und Manganionen an Phosphate ausgeprägte strukturelle Veränderungen, wie zum Beispiel die Bildung stabiler α-Helizes bzw. helikaler Fibrillen zur Folge haben. Zusätzlich zu diesen Untersuchungen wurde auch die Phosphorylierung eines natürlichen aggregierenden Systems untersucht. Dabei kam eine Kombination aus rekombinanter Proteinexpression und Peptidsynthese zum Einsatz, um homogene Präparate ortsspezifisch phosphorylierten Tau-Proteins, einer wichtigen Komponente von Neurofibrillen und dem primären Kennzeichen der Alzheimerschen Krankheit, darzustellen.

Abstract

The aggregation of soluble proteins to protease-resistant fibrils is a hallmark of neurodegenerative diseases. This phenomenon has been observed in proteins that under certain conditions undergo a structural transition from an unfolded or partially helical state to a β-sheet structure. Factors such as a solution's pH, protein concentration, presence of metal ions, and temperature have been identified as triggers of protein aggregation. In addition, post-translational modifications, e.g. phosphorylation, have been implicated in aberrant protein fibrillization over the last two decades. However, current studies on phosphorylation patterns in natural aggregating systems suffer from major drawbacks; for instance, rapid dephosphorylation of phosphoproteins *in vivo* renders the identification of key phosphorylated residues difficult, whereas *in vitro* enzymatic phosphorylations generate heterogeneous mixtures of proteins phosphorylated at various sites and to different extent.

This work presents several examples of a successful application of synthetic phosphopeptides as tools for studying amyloid formation. Using such analytical techniques as circular dichroism spectrscopy, Thioflavine T-induced fluorescence, and Transmission Electron Microscopy various aspects of peptide and protein fibrillization were investigated. For example, the influence of site-specific and multiple phosphorylation on the amyloid formation process was examined utilizing an amyloid forming coiled coil peptide model. Structural analysis of phosphorylated peptides revealed that regardless of the quantity and position of modification, the formation of β-sheet structure was completely abolished. Furthermore, all phosphorylated peptides completely lost their amyloid forming potential. Subsequently, with the objective to understand and monitor structure switching as well as fibril formation and morphology under conditions that approximate physiological environment, a dephosphorylating enzyme was applied as a natural aggregation trigger. In addition, synthetic phosphopeptides were applied to investigate metal-induced conformational transitions. Coordination of magnesium and manganese ions to phosphate groups induced vigorous structural changes, resulting in stable α-helices and helical fibers, respectively. Finally, in addition to the aforementioned investigations of coiled coil model systems, also phosphorylation in a natural aggregating system was probed. Here, a combination of synthetic and recombinant approaches (expressed protein ligation) was used to generate homogenous preparations of site-specifically phosphorylated microtubule-assisted tau protein – a major component of neurofibrillary tangles, the primary marker of Alzheimer's disease.

Abbreviations

Abz	ortho-aminobenzoic acid
ACN	acetonitrile
AD	Alzheimer's disease
APP	Amyloid Precursor Protein
Aβ	Amyloid β
Boc	*tert*-butoxycarbonyl
CaMK-II	calcium calmodulin-dependent kinase II
CBD	chitin binding domain
CD	circular dichroism
Cdk5	cyclin-dependent kinase 5
DIC	N,N′-diisopropylcarbodiimide
DIEA	N,N,-diisopropylethylamine
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EPL	expressed protein ligation
eq.	equivalent
ESI	electrospray ionization
Fmoc	9-fluorenylmethyloxycarbonyl
FRET	fluorescence resonance energy transfer
FTIR	Fourier transform infrared spectroscopy
Gsk-3β	glycogen synthase kinase-3β
GTP	guanosine triphosphate
HATU	2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HRP	horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MESNA	sodium 2-sulfanylethanesulfonate
mRNA	messenger ribonucleic acid
Mxe	*Mycobacterium xenopi*
NCL	native chemical ligation
NFT	neurofibrillary tangles
NMR	nuclear magnetic resonance
OD	optical density

O-GlcNAc	O-linked β-N-acetylglucosamine
PCR	polymerase chain reaction
PHF	paired helical filaments
PHF-1	paired helical filament epitope
PIPES	1,4-Piperazinediethanesulfonic acid
PKA	protein kinase A
PP2A	protein phosphatase 2A
PTMs	post-translational modifications
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPPS	solid-phase peptide synthesis
TCEP	tris(2-carboxyethyl)phosphine
TEM	transmission electron microscopy
TEV	Tobacco Etch Virus
TFA	trifluoroacetic acid
ThS	Thioflavin S
ThT	Thioflavin T
TIS	triisopropylsilane
TOF	time of flight
tRNA	transfer ribonucleic acid
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet

Abbreviations of the 20 canonical amino acids are consistent with the biochemical nomenclature proposed by the IUPAC-IUB commission (*Eur. J. Biochem.* **1984**, *138*, 9-37).

Contents

1 Introduction

The neurological pathology known today as Alzheimer's disease (AD) was first recognized in 1906 by the German psychiatrist Alois Alzheimer.[1] Since that time AD has become the most common neurodegenerative disorder with over 20 million cases worldwide, furthermore, this number is expected to double within the next decade.[2] The disease is characterized by the accumulation of two classes of abnormal protein deposits in the brain. The first class consists of extracellular senile plaques composed of amyloid β (Aβ) peptides derived from proteolytic cleavage of amyloid precursor protein (APP).[3] The second class comprises intraneuronal neurofibrillar tangles with paired helical morphology composed of hyperphosphorylated microtubule-assisted tau protein.[4]

Over the years the attention of researchers was predominantly devoted towards Aβ-induced neuropathology, which resulted in the introduction in early 1990s of so-called amyloid hypothesis.[5] It states that the accumulation of amyloid beta is the primary factor that initiates and drives pathological events leading to the clinical manifestation of the disease. Accordingly, amyloid peptides have become targets for many therapeutic strategies. The emphasis has been predominantly put on reducing Aβ production and aberrant self-assembly as well as increasing its clearance. Along these lines inhibitors and modulators of APP processing enzymes which generate Aβ have been extensively tested along with peptide- and small molecule-based aggregation inhibitors and antibody-mediated amyloid clearance approaches.[6, 7] Despite these ongoing efforts, the disease remains undefeated and most of the available therapeutic strategies relay only on the symptomatic treatment.

Initially, tau related research received rather moderate attention. It was after the discovery that mutations in the tau gene cause rare autosomal dominant neurodegenerative diseases collectively known as frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17),[8] in which tau was abnormally phosphorylated and aggregated in the absence of detectable amyloid phatology that this field was brought up to speed. Furthermore, a growing number of sporadic neurodegenerative disorders termed tauopathies was identified.[9] Here lesions of abnormally phosphorylated tau were also the predominant pathology. A prominent addition to aforementioned facts is the discovery that the number of amyloid plaques in the brain does not correlate well with the disease progression[10] and actually it is the distribution and abundance of neurofibrillary degeneration that reflects various AD stages much more reliably.[11] Thus it became evident that tau hyperphosphorylation and self-assembly constitutes a central role in AD pathology and a potential therapeutic target. Unfortunately, tau has received much less attention from the pharmaceutical

industry than amyloid. This is mostly due to both the amyloid cascade hypothesis positioning tau pathology downstream to the amyloid as well as the fact that causative factors in the pathway from tau to tangles have not been clearly identified. Taking into account that in all tau related disorders the protein is found abnormally phosphorylated, the most promising approach would involve selective inhibition of tau kinases and/or activation of tau phosphatases.

Even though significant progress has been made in the identification of disease specific phosphorylation sites on tau and several candidate kinases have been proposed,[12] the question which specific phosphorylation pattern triggers the transformation from functional microtubule-stabilizing protein to pathological aggregate remains open. The difficulty lies in the dynamic nature of phosphorylation *in vivo* and heterogenous character of enzymatic phosphorylations *in vitro*. Additionally, the complex nature of molecular interactions as well as the challenging physicochemical properties of naturally aggregating systems render their detailed structural and functional investigation difficult. A promising strategy to overcome all these obstacles is the application of synthetic phosphopeptides of different length and complexity. On one hand, they can serve as synthetically accessible and easy to handle models for better understanding of molecular events and intermediates that occur during the process of amyloid formation. On the other hand, taking advantage of various chemoselective ligation strategies[13] and recombinant protein expression, synthetic phosphopeptides can be applied to produce homogeneous preparations of site-specifically modified tau protein. This in turn can aid detailed investigation of pathological phosphorylation patterns on tau and thus pave the way for further diagnosis and treatment approaches.

2 The phenomenon of protein aggregation

The ability of most proteins to maintain compact three-dimensional structures is a crucial aspect of their function. The process of protein folding can be defined as the search for the most thermodynamically stable conformation under given conditions and is evolutionarily designed to be rapid and efficient. For these particular reasons, the folding process cannot involve a systematic search between specific partially folded states, but rather a stochastic search of various conformations accessible to a polypeptide chain.[14] Specifically, the intrinsic nature of incompletely folded intermediates allows for a certain level of conformational freedom so that even residues at various positions in the amino acid sequence can come into contact with one another. The stability of native-like interactions is on average higher than the non-native ones, therefore the polypeptide chain can in principle find its lowest energy conformation by simple trial and error. Usually the 'energy landscape,' defined as the free energy of a polypeptide chain as a function of its conformational properties, narrows this search, as the number of possible conformations decreases when the native state is approached.[15] As a result of extensive experimental and computational studies on small proteins, the mechanism of folding can be characterized as a 'nucleation-condensation' event,[16] in which a nucleus comprising a small number of key residues is first formed followed by the generation of the native fold as the reminder of the protein condenses around the nucleus. Larger proteins usually fold in modules, which highlights that folding can occur independently in different segments of the protein.[17]

The acquisition and maintaining of the correct fold is strictly controlled in living organisms. Molecular chaperones and folding catalysts, like peptidylprolyl isomerases and protein disulphide isomerases, as well as co-translational glycosylation events enable and accelerate efficient folding.[18, 19] In addition, newly synthesized proteins are subjected to rigorous 'quality control' tests which involve a series of glycosylation and deglycosylation reactions and allow the incorrectly folded proteins to be distinguished from the properly folded ones.[20] In certain cases, however, unfolded/misfolded proteins escape all the protective mechanisms of the cell. Such a scenario often results in proteins self-assembly leading to the formation of insoluble aggregates, as the normally buried hydrophobic side chains of the protein become surface exposed, and gives rise to disease. Examples of such disorders include so-called amyloidoses, like Alzheimer's, Parkinson's, and Creutzfeldt-Jakob disease as well as type II diabetes.

It is intriguing that the above-mentioned misfolding disorders involve proteins or peptides with no obvious similarities in size, sequence, or structure. Nevertheless, under certain conditions, e.g. the presence of mutation, changes in pH, temperature or local concentration, interaction with metal ions or chemical modification, they aggregate to form a common

amyloid structure. Perhaps even more astonishing is the fact that various non-disease related proteins[21] or even small peptides[22] under experimentally designed conditions that favor destabilization of the native fold can form aggregates indistinguishable in their morphology from those related to disease. This fact suggests that the propensity to form amyloid is a generic feature of a polypeptide chain.[21]

2.1 Structure of amyloid fibrils

Despite the rapid development of new biophysical techniques, revealing the detailed structural features of amyloid fibril in atomic detail has proven difficult. This is mostly due to the fact that amyloids are noncrystalline solids, and are therefore incompatible with X-ray crystallography and liquid state NMR. In addition, factors such as fibril growth conditions and characteristics of the polypeptide chain involved (length, sequence) introduce significant polymorphism in amyloid molecular architecture.[23, 24] However, the application of such techniques as electron and atomic force microscopy, solid state NMR, and X-ray fiber diffraction prompted the emergence of increasingly convincing structural models of amyloid fibers. The most general features of such assemblies include: the binding of specific dyes (Thioflavin T, Thioflavin S, Congo red), unbranched, extended fibril morphology, and 'cross β' X-ray diffraction pattern. In addition, circular dichroism (CD) and Fourier transform infrared (FTIR) spectroscopy provide evidence of predominant β-sheet content.

2.1.1 Macroscopic structure of amyloids

Investigations based on electron and atomic force microscopy revealed that amyloid fibrils are typically long, straight, and unbranched, with the diameter in the range of 70-120 Å. They usually consist of several (up to six) subunit structures, called protofilaments, which twist around each other resulting in supercoiled, rope-like structures.[25, 26]

In the study of Jimenez et al. insulin amyloid fibrils comprising 2, 4, and 6 protofilaments have been investigated in order to gain insight into their common underlying structure.[27] Based on cryo-electron microscopy the 3D maps of protofilament arrangement have been generated (Figure 2.1 A) and subsequently revealed that despite different morphologies, the average size (30-40 Å) and compact shape of individual protofilaments turned out to be the same. Furthermore, a model for protofilament packing was proposed (Figure 2.1 B and C). According to the authors, consistent interactions between protofilaments can occur exclusively in the case if they follow the overall twist of the fibril (left-handed), as depicted in Figure 2.1 B. This particular arrangement would position the noninteracting regions (loops, unordered parts of the polypeptide chain) that can potentially disrupt the packing always on

the surface and away from contacts with other protofilaments. If, however, the twist of the protofilament does not follow the one of the fibril, the regions of the interacting interface rotate progressively around the protofilament, and thus any consistent interactions do not occur (Figure 2.1 C).

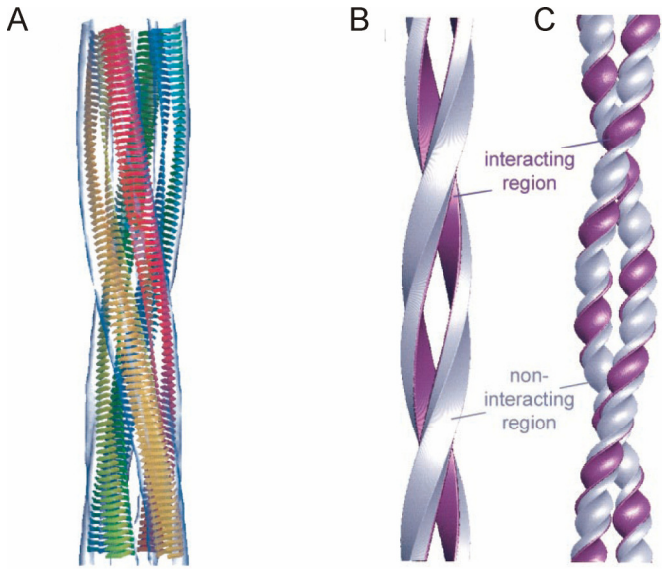


Figure 2.1 *Morphology and protofilament packing of insulin fibers. (A) 3D model of insulin fiber composed of four protofilaments wound around each other. Within each protofilament β-strands are highlighted in color. (B), (C) Models of protofilament packing: (B) A twisted pair of protofilaments with the interacting surface highlighted in purple. (C) A supercoiled pair of protofilaments in which the regions involved in packing interactions rotate around each protofilament (reproduced from J. L. Jimenez et al.[27] Copyright © 2002 National Academy of Sciences, USA)*.

Several investigations addressing other amyloid forming systems deliver structural models resembling the one discussed above. Similar 3D maps have been generated for SH3 domain of phosphatidylinositol-3'-kinase.[28] They reveal the presence of two pairs of thin (20 Å) protofilaments winding around a hollow core. In addition, Kirschner and co-workers proposed a possible model for Aβ(1-40) assembly.[29] In this study an average 70 Å wide fibril was made of three, four, or five ~30 Å wide tubular protofilaments.

Despite the apparent morphological diversity of amyloid fibers, resulting predominantly from different numbers and arrangements of protofilaments, the internal architecture seems rather uniform.

2.1.2 Microscopic structure of amyloids

The first discovery that amyloid fibrils share a common X-ray diffraction signature, known as the cross-β pattern, came in 1968.[30] The most characteristic features of this pattern include a strong 4.7 Å meridional reflection corresponding to the spacing between individual β-strands,

which are perpendicular to the long fibril axis, and a weaker equatorial reflection at 8-12 Å corresponding to the stacking periodicity of β-sheets which propagate along the direction of the fibril. Following investigations of the cross-β spine in atomic detail focused mostly on aggregates derived from short peptides, as they can be assembled *in vitro* and yield well ordered fibrils.

A detailed atomic model of the common fibril spine was described recently by the Eisenberg group.[31] They focused their study on the short peptide (GNNQQNY) derived from the yeast protein Sup35, which readily forms elongated microcrystals enabling X-ray diffraction studies. It was demonstrated that peptide molecules form extended β-strands spaced approximately 4.87 Å apart, hydrogen bonded to each other in a parallel and in register fashion. The strands comprising one β-sheet are antiparallel and vertically shifted with respect to those in the neighboring sheet (Figure 2.2 A). This shift results from the perfect packing of the side chains within the interface of both sheets. Specifically, side chains extending from a strand in one sheet fit between side chains extending from two strands in the next sheet.

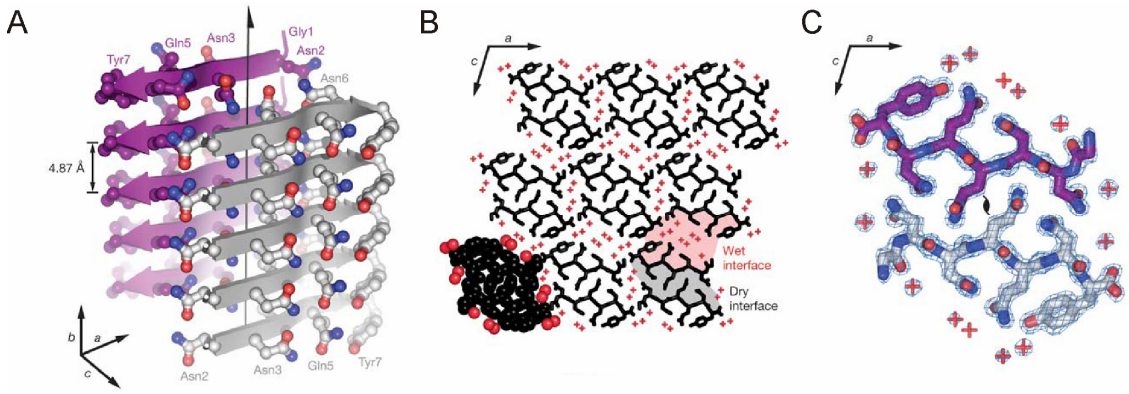


Figure 2.2 *Structure of GNNQQNY peptide. (A) The pair-of-sheets structure with the backbone of each β-strand depicted as an arrow with the side chains protruding. (B) A view down the fibril axis showing dry and wet interface. (C) The 'steric zipper' interaction in a close view (reproduced with permission from R. Nelson et al.[31] Copyright © 2005, Nature Publishing Group).*

The lateral growth, i.e. lamination of more than two pairs of sheets, clearly indicates the presence of two different interfaces in the case of GNNQQNY peptide (Figure 2.2 B). The wet interface contains water molecules, which hydrate polar side chains between the two sheets, rendering the separation of these sheets large, about 15 Å. In contrast, the dry interface does not contain water, and therefore the sheets are closer together, about 8.5 Å apart. Polar side chains within the dry interface do not form hydrogen bonds. They are rather tightly interdigitized and stabilized by van der Waals interactions, highlighting the importance of side chains size and shape complementarity. The authors described this interaction as a 'steric zipper,' as the interdigitized side chains of the dry interface resemble the teeth of a

zipper (Figure 2.2 C). Two fundamental features of the amyloid spine become evident based on the molecular model of GNNQQNY peptide. First, the remarkable complementarity between sheets forming the 'steric zipper' suggests that the basic structural unit of the amyloid spine is a pair of β-sheets. Second, the length of the studied peptide points to the fact that short segments of the polypeptide chain are sufficient for the formation of the core structure.

A similar, detailed atomic model was proposed for the Alzheimer's Aβ(1-42) fibrils.[32] The authors combined the knowledge from previous investigations of Aβ(1-42) fibrils, i.e. the presence of cross-β structure composed of parallel, in register β-sheets,[33, 34] with hydrogen exchange measurements, pairwise mutagenesis, Thioflavin T binding, and cryo-electron microscopy. The proposed 3D structure of Aβ(1-42) protofilament consists of β-strand-loop-β-strand segments aligned in a staggered manner to form two parallel, in register β-sheets which extend along the fibril axis (Figure 2.3 A and B).

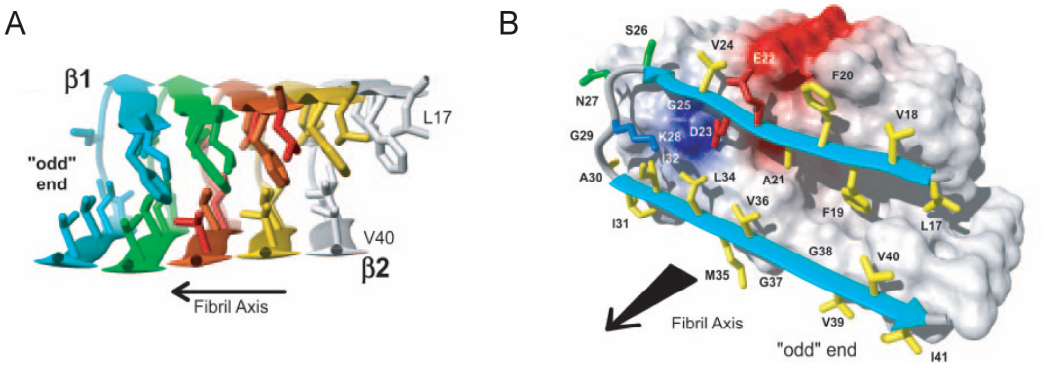


Figure 2.3 *The 3D structure of Aβ(1-42) fibril. (A) Ribbon diagram of the core structure (residues 17-42) illustrating the intermolecular nature of inter-strand interactions. (B) van der Waals contact surface polarity and ribbon representation of the single Aβ(17-42) hairpin (reproduced from T. Lührs et al.[32] Copyright © 2005 National Academy of Sciences, USA).*

Strand (β1) of one Aβ(1-42) molecule interacts with strand (β2) of the preceding one to form an amyloid core with exclusively intermolecular 'steric zipper' packing (Figure 2.3 A). This structural view is distinct from previously reported models that proposed intramolecular interactions in Aβ fibrils.[35] The staggered interaction pattern leads to partially unpaired strands at the fibrillar ends, which could explain the cooperativity and unidirectionality of fibril extension, the sequence selectivity, and finally the structural basis of fibrillization inhibitors.[32]

Even though the cross-β structure has been commonly accepted, it is important to emphasize that there is no indisputable structural model of amyloid fibril, and several alternative models have been proposed. These include β-helix or nanotube and amyloid models retaining native structure.[36]

2.2 Kinetics of fibrillization

Although the amyloidogenic states of peptides and proteins are rich in β-sheet structure, their soluble forms often possess well defined globular fold or consist of a predominantly unfolded structure.[37] In order for the fibrillization to proceed, globular proteins need to partially unfold, so that the polypeptide main chain and hydrophobic residues are exposed. Conversely, the natively random coil amyloid precursors need to adopt partially structured conformations (Figure 2.4 i, ii). The presence of such partially ordered species enables the occurrence of specific non covalent intermolecular interactions that in consequence favor the formation of enriched β-sheet structure via self-assembly.

It is commonly accepted that the process of amyloid formation usually occurs through a nucleation-dependent pathway, which is characterized by a slow formation of nuclei followed by a period of rapid growth.[38] The nucleus formation is the rate limiting step in this mechanism and occurs after a lag phase, which can be eliminated by a local increase in monomer concentration or seeding with a preformed aggregate (Figure 2.4 iii, iv, v).[39] As soon as the nucleus is formed the fibrillization process proceeds fast by a series of elongation steps to yield mature fibrils (Figure 2.4 vi, vii).

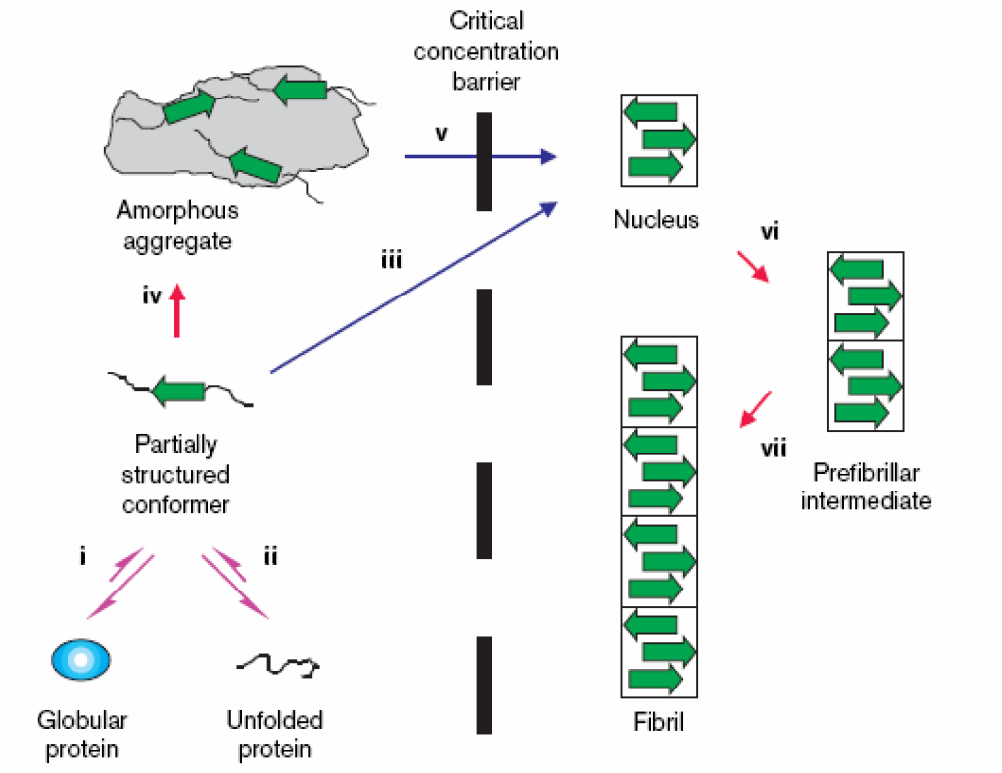


Figure 2.4 *Schematic illustration of the fibrillization process. Partially structured intermediate, which is formed via partial folding or unfolding (i, ii), self-assembles to form a nucleus (iii, iv, v). Growth of the nucleus leads to formation of protofibrils (vi) and finally matured fibrils (vii) (reproduced with permission from J.-C. Rochet et al.[37] Copyright © 2000, Elsevier).*

In its simplest form nucleation-dependent fibrillization is a two-state process, i.e. monomers are converted to mature fibrils in a cooperative fashion without generation of any intermediates, as shown, for example, for Aβ fibrils.[40] However, other investigations of both Aβ fibrils[41] as well as other aggregating systems[23, 24] unambiguously demonstrated the presence of prefibrillar intermediates, called oligomers and protofibrils (Figure 2.4 vi). Although not without controversy,[42] these species are thought to represent on-pathway intermediates of amyloid formation. It is proposed that soluble, spherical or bead-like oligomers usually precede the formation of non-spherical, filamentous protofibrils, which further assemble to form mature fibrils. In addition, as these early aggregates are likely to be disorganized structures exposing various hydrophobic segments to the surrounding environment, they have been suggested to be the toxic agents *in vivo*. This hypothesis was validated by several findings, for example, experiments *in vivo* and in cell culture showed that in the absence of mature fibrils Aβ42 oligomers were neurotoxic.[43] Furthermore, polyclonal antibodies suppress soluble oligomer toxicity, whereas no response was observed for mature fibrils.[44] Finally, it was demonstrated in a mouse model that specific soluble Aβ multimeric species are associated with memory loss in AD.[45] The apparent toxicity of early aggregates appears to be a consequence of their intrinsic ability to impair fundamental cellular processes via interacting with cell membranes or inducing oxidative stress.[26] All this evidence points to the rather controversial conclusion that the mature fibril formation is actually a protective mechanism of the cell.

Initially, it was assumed that the fibrillization can be described by a first order kinetics, which would mean that the rate of fibril elongation is proportional to the concentration of monomers. Recent work indicates, however, that the growth kinetics is more complicated and could be described by a model in which nucleation is followed by a stepwise monomer addition as well as competing fragmentation of fibrils. This reversibility of the elongation phase has been demonstrated for Aβ(1-40) fibrils, where a release of monomers has been observed, and in particular for aggregates of SH3 domains, where fully exchanged fibers were generated by H/D exchange.[46, 47] Together, these results indicate that the process of amyloid formation is not just a 'trivial' polymerization reaction, to which it is often compared, but a complex, multi-step phenomenon with astonishing polymorphism within its intermediates and products.

2.3 Common triggers of amyloid formation

Amyloid formation is usually primed by the presence of misfolded (partially folded/unfolded) species.[18, 37] Conditions that have been associated with generation of such partially unfolded states have therefore been under particular scrutiny.

Mutations in the primary structure[8, 48] that decrease stability of the native state, as well as proteolysis[49] that generates fragmented proteins, have been postulated as causative factors in Alzheimer's pathology. Furthermore, changes in the proteins' environment have been commonly accepted as aggregation triggers. Along these lines, alterations in pH value and/or ionic strength provide a robust strategy for inducing self-assembly. As demonstrated for the SH3 domain of α-subunit of bovine phosphatidylinositol-3'-kinase, changes in the solution's pH and ionic composition not only trigger fibrillization but also generate fibers with distinct morphologies.[23] Studies on β_2-Microglobulin, moreover, illustrated the existence of two competing assembly pathways under varied pH and ionic strength.[24]

As discussed in the preceding section, nucleation is a rate limiting step in amyloid formation. Therefore, an increase in local concentration or 'molecular crowding' can greatly facilitate the generation of the nucleus and the overall propensity to aggregate. The efficiency of this trigger in inducing fibrillization has been demonstrated for protein substrates[24] as well as in model systems.[50] Although considered rather harsh, *in vitro* partial unfolding can also be induced by the increase in temperature[51] as well as addition of an organic solvent, such as trifluoroethanol, which is known to stabilize partially folded states.[21]

Transition metal ions, for example Cu^{2+}, Zn^{2+}, and Fe^{3+}, are yet another example of a common aggregation trigger. It has actually been postulated that metallochemical reactions might be the common denominator in the pathology of Alzheimer's disease.[52] There are two ways by which metals can alter protein conformation. The first involves direct metal coordination to a particular side chain in a peptide or protein, usually cysteine or histidine.[53-55] The second possibility is metal-induced protein oxidation via generation of reactive oxygen species.[52, 56]

Over the last two decades post translational modifications have also been implicated in aberrant protein fibrillization. Especially the role of phosphorylation has been studied extensively, as it has been shown that one of the major brain lesions of Alzheimer's disease is fully composed of hyperphosphorylated and aggregated microtubule-assisted tau protein.[4]

3 Phosphorylation and protein aggregation

Phosphorylation of cellular proteins on their serine, threonine, and tyrosine residues is a major post-translational modification and ubiquitous regulatory mechanism. It is presumably due to the simplicity, flexibility, and reversibility of phosphorylation why it has been selected to control almost all major cellular processes. These include metabolic pathways, kinase cascade activation, gene transcription, cellular proliferation, as well as membrane transport.[57] Phosphorylation is an enzymatic process and two classes of enzymes control the phosphorylation state of a cell. Protein kinases catalyze the covalent attachment of the phosphate moiety, whereas protein phosphatases catalyze its selective removal. The known crystal structures of protein kinases show remarkable sequence similarities and a conserved fold.[58] Phosphatases, in contrast, represent a diverse group of structurally distinct proteins.[59] Both kinases and phosphatases can be classified with respect to their substrate specificity to serine/threonine and tyrosine families.

The control and regulation of various processes via reversible protein phosphorylation is implemented by inducing conformational transitions, which in turn can affect the function of a protein in a variety of ways, e.g. by increasing or decreasing activity, marking it for translocation to another subcellular compartment or for destruction, and allowing association or dissociation from another protein.[60] The ability to induce structural changes is a result of two particular features of the phosphate moiety, namely bulk and charge. Analyses of protein-phosphate interactions revealed two different mechanisms for the structural response to this modification. It was demonstrated that in a dimeric rabbit muscle glycogen phosphorylase phosphorylation results in an allosteric activation as the phosphorylated serine shifts 50 Å to change favorable electrostatic contacts from intra- to intersubunit.[61] On the other hand, in *E. coli* isocitrate dehydrogenase phosphorylation results in inhibition by blocking the access to the active site.[62]

3.1 The impact of phosphorylation on the secondary structure

In majority of cases, the most effective method for elucidating the effects of phosphorylation on proteins' conformation is the crystal structure comparison of both phosphorylated and unphosphorylated variants. This methodology, however, has been severely hampered by the crystallization potential of a particular protein and the required quantities. Due to these obstacles, the influence of phosphorylation on the secondary structure has been a subject of extensive study in smaller, easily accessible model systems.

A great number of investigations have been focused on the impact of phosphorylation on the helical structure. It has been demonstrated that in general phosphorylation of helices is destabilizing either due to electrostatics (especially at the C-terminus) or to the high desolvation penalty associated with the bulky side chain (mainly in the helix interior).[63, 64] Stabilizing interactions involving phosphate moiety are possible only at the positively charged N-terminus of the helix[63, 65] or by the involvement of the phosphate into attractive Coulombic interactions with neighboring residues.[66, 67] Additionally, the consequences of serine, threonine, and tyrosine phosphorylation have also been studied in different structural models. In a designed β-hairpin peptide the phosphorylated residue was juxtaposed with a tryptophan moiety, which was shown to destabilize the original structure due to the electrostatic repulsion between phosphate and electron-rich indole ring.[68] Conversely, in a set of intrinsically unfolded tau protein-derived peptides phosphorylation induced the more ordered polyproline II helix.[69] Furthermore, the capability of phosphorylation to promote β-sheet structure was demonstrated upon full modification of DNA-bound carboxy-terminal domain of histone H1.[70] Finally, accumulated evidence shows that phosphorylation can also impact secondary structure through metal binding.[71-74]

It has thus become apparent that phosphorylation can affect the conformation of peptides and proteins in a variety of ways. Taking into account that about 30 % of all human proteins contain covalently bound phosphate, and that protein kinases and phosphatases constitute about 4 % of the eukaryotic genome,[75] it is not surprising that any imbalance in the otherwise synchronized action of these enzymes would be a cause or a consequence of disease.

3.2 Hyperphosphorylation of tau

Tau is a family of neuronal proteins produced by alternative mRNA splicing of a single gene. Six isoforms of tau are present in an adult brain, they differ form each other by the absence or presence of one or two N-terminal inserts (29 or 58 amino acids) and by the number of tandem repeats in the C-terminal part (Figure 3.1).[76-78] Since its discovery in 1975 tau was associated with microtubules.[79] To this day the best-known biological function of tau is to stabilize microtubule structure and promote their assembly. However, other physiological functions of tau have been suggested as well. For example, tau seems to interfere with the binding of kinesin and kinesin-like motors to microtubules, which leads to the inhibition of kinesin-dependent axonal transport.[80] Furthermore, tau was found to interact with plasma membrane,[81] nucleic acid,[82] mitochondria,[83] and also src-family nonreceptor tyrosine kinases.[84] Although the physiological significance of these interactions remains elusive, possible involvement in signal transduction has been suggested. The binding of tau to

microtubules is implemented via its three or four tandem repeats, whereas possible other functions are probably mediated by the N-terminal part of the protein, which projects away from the microtubule surface.

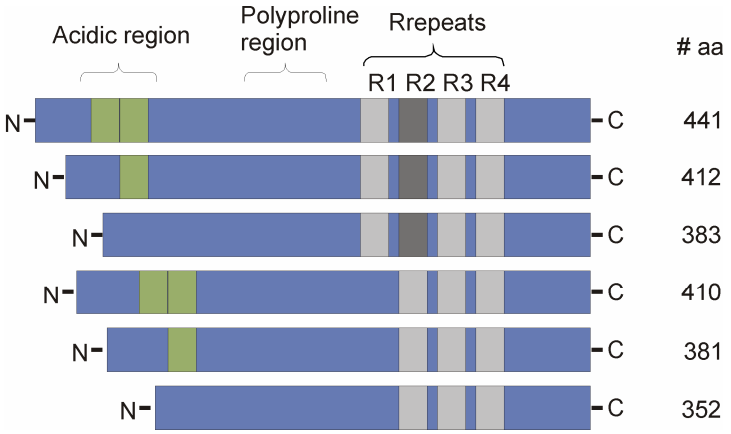


Figure 3.1 *Schematic illustration of six tau isoforms found in human brain. Alternatively spliced regions are shown in green and dark gray.*

From the structural point of view, tau belongs to the growing family of natively unfolded or intrinsically disordered proteins.[85] This classification is based on the fact that in solution tau behaves predominantly like a random coil according to CD, FTIR, and X-ray scattering measurements.[86] Interestingly, recent studies report on the existence of so-called global folding of tau. Mandelkow and co-workers used fluorescence resonance energy transfer (FRET) to investigate possible interactions within the tau molecule.[87] The observed FRET distances were significantly different from those expected for a random coil. Presented data suggest a paperclip folding of tau, in which the C-terminus folds over the repeat domain, whereas the N-terminus approaches the C-terminus. In addition, NMR studies corroborate the presence of transient elements of secondary structure in tau, including β-sheet propensity in the repeat domain, polyproline II helical conformation in the proline rich regions flanking the repeats, and α-helical structure at the end of the N-terminal domain as well as at the extreme C-terminus. [88] Thus, tau appears to be highly dynamic in solution with a distinct domain character and a network of transient long-range interactions.

In 1977 tau was found to be a phosphoprotein.[89] Phosphorylation is one of many post-translational modifications affecting tau[90] and it is necessary for the correct function of the molecule. Due to still largely unknown reasons, however, tau becomes hyperphosphorylated and in that state it loses its biological activity,[91, 92] moreover, as such it is found to make up paired helical filaments (PHFs) that form neurofibrillary tangles (NFTs) in AD[4] and other neurodegenerative disorders collectively called 'tauopathies.'[9] Several lines of evidence suggest that hyperphosphorylation is responsible for the aggregation and

pathological behavior of tau. First of all, it has been observed that in neural cultured cells treated with phosphatase inhibitors, tau hyperphosphorylation inhibits its binding to microtubules, and furthermore induces filamentous aggregates.[93] In contrast, treatment of hyperphosphorylated tau with phosphatases restores its activity[92] and abolishes its self-assembly.[94] Alonso et al. showed, furthermore, that abnormally phosphorylated tau sequesters normal tau molecules, and that this behavior is lost upon dephosphorylation.[95] Taking all this into account, it seems obvious that the hyperphosphorylation event plays a central role in tau loss of function, toxicity, and aggregation, and that the identification of abnormal phosphorylation sites as well as responsible enzymes is imperative.

3.2.1 Phosphorylation sites and responsible enzymes

The longest tau isoform (441 amino acids) contains 80 serine or threonine residues and five tyrosines, which means that about 20 % of the molecule could be phosphorylated. The identification of key phosphorylation sites on tau (Figure 3.2) has been accomplished using a combination of mass spectrometry, phosphopeptide sequencing, and predominantly epitope mapping.[96, 97] To date, more than 30 serine and threonine residues have been found phosphorylated in PHF-tau.[90] Many of these residues are also modified in normal tau, but to a much lesser extent. Some of the phosphorylation sites, however, are only present in the pathological tau, including Thr212/Ser214, Thr231/Ser235,[98] and Ser422[99] (numbering according to 441 amino acid tau isoform). In addition, it was demonstrated that tau phosphorylation at Ser262, Thr231, and Ser235 is responsible for the inhibition of microtubule binding,[100] phosphorylation at Ser199/Ser202/Thr205, Thr212, Thr231/Ser235, and Ser422 is critical for the ability of tau to sequester other microtubule associated proteins,[101] and finally phosphorylation at Thr231, Ser396, Ser400, Ser 404, and Ser422 promotes tau self-assembly to filaments.[12]

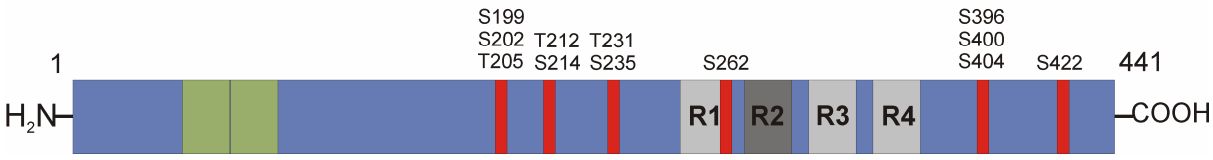


Figure 3.2 *Identified pathological phosphorylation sites on tau.*

The progress in the identification of tau phosphorylation sites was accompanied by extensive studies on the enzymes that modify tau *in vivo*. The kinases that most likely phosphorylate tau in the brain include glycogen synthase kinase-3β (Gsk-3β), cyclin-dependent kinase 5 (cdk5), protein kinase A (PKA), stress activated protein kinases,

and calcium calmodulin-dependent kinase II (CaMK-II).[12] In turn, the major enzyme responsible for tau dephosphorylation was identified as protein phosphatase 2A (PP2A).[102]

Recently, it has been shown that in addition to serine/threonine phosphorylation, tau is also modified on its tyrosine residues by the src-family nonreceptor tyrosine kinase fyn[103] and c-Abl.[104] The tyrosine phosphorylation of tau appears rapid and transient and whether it has any pathophysiological relevance has yet to be elucidated.

3.2.2 Mechanisms leading to hyperphosphorylation

Despite significant progress in tau-related research, the mechanisms leading to tau hyperphosphorylation, which seems to precede its detachment from microtubules and abnormal aggregation, are still unknown. The obvious assumption is the dysregulation of enzymes processing tau. Along these lines the upregulation of tau kinase cdk5 has been proposed by one group but was subsequently argued against by others.[90] On the other hand, the downregulation of brain phosphatases is another possibility. It has been shown that both the activity and expression of PP2A are reduced in selected areas of AD brain which is consistent with the fact that several other neuronal proteins have been found hyperphosphorylated in AD brain.[105] It has also been suggested that downregulation of PP2A may have an indirect role in inducing tau hyperphosphorylation, namely by the activation of several PP2A-regulated protein kinases like CaMK-II.[106]

In addition to tau kinases and phosphatases, alterations of tau protein itself may play a role in the aberrant phosphorylation and PHF formation. Recently, it was discovered that human brain tau is modified by O-linked β-N-acetylglucosamine (O-GlcNAc).[107] This novel class of O-glycosylation catalyzed by O-β-N-acetylglucosaminyltransferases (OGTs) dynamically modifies nucleoplasmic and cytoplasmic proteins on serines and threonines in a manner similar to that of phosphorylation. Due to the fact that both modifications target the same amino acids, and that the reciprocal relationship between them has been shown for many proteins,[108] it was also suggested that O-GlcNAcylation could affect phosphorylation of tau. This assumption was actually proved *in vivo*, where fasting of mice induced decrease of tau O-GlcNAcylation (relating on glucose metabolism) and subsequent hyperphosphorylation.[107] A new hypothesis was thus proposed that impaired glucose metabolism, established decades ago, contributes to the pathological phosphorylation of tau.[109]

Another factor that may influence the phosphorylation level of tau is N-glycosylation. It was demonstrated by specific lectin staining experiments that AD brain tau, but not normal tau, is N-glycosylated[110] with potential modification sites being Asn167, Asn359, and Asn410 residues (numbering according to the longest tau isoform).[111] Furthermore, this modification

was shown to precede and facilitate tau hyperphosphorylation by major tau kinases, including PKA, cdk5, and Gsk-3β.[112, 113]

3.2.3 PHFs structure and assembly

The inevitable consequence of tau hyperphosphorylation, regardless of its origin, is the formation of neurofibrillary deposits. In AD they are predominantly composed of PHFs, whose characteristic morphology was first described in 1963.[114] Electron micrographs show helically twisted filaments 8-20 nm wide with the apparent period of 80 nm. Detailed investigations revealed the double-stranded nature of the packing, furthermore, computed cross section of the filament showed two C-shaped structural units (Figure 3.3 A and B). [115] The internal structure of PHFs has been a matter of debate. Both the cross-β[116, 117] as well as α-helical arrangement[118] has been shown by spectroscopic techniques. Unfortunately, no high resolution data on PHFs core domain has yet been available.

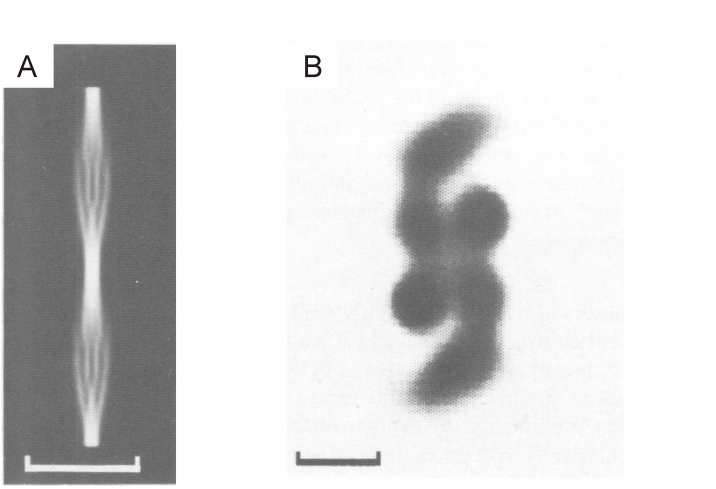


Figure 3.3 *Paired helical filaments. (A) Computed projection normal to the filament axis, scale bar 50 nm. (B) Reconstructed cross-section of (A) showing two C-shaped subunits, scale bar 5 nm (reproduced with permission from R. A. Crowther et al.[115] Copyright © 1985, Macmillan Publishers Ltd: The EMBO Journal).*

The mechanism of tau assembly into PHFs has been a subject of extensive *in vitro* study. Initial experiments have been hampered by the low efficiency of unphosphorylated tau aggregation, the requirement for high protein concentrations, and the lack of suitable detection methods. Several advances have been made, however, which have allowed substantial progress in this area. Specifically, it was demonstrated that the repeat domain of tau aggregates more efficiently than full-length isoforms[119] and that tau assembly proceeds more rapidly if the protein is intermolecularly cross-linked at Cys322 to form dimers.[120]

Finally polyanions, like heparin, RNA, and polyglutamate, further enhance tau aggregation.[121] Taking advantage of these facts Friedfoff et al. developed a fluorescence-based quantitative assay, which allowed monitoring the PHFs assembly in real time.[121] In addition, these authors identified that tau assembly is optimal at physiological pH, low ionic strength, and that it is highly dependent on temperature. These details affecting tau aggregation allowed the same group to subsequently demonstrate that PHFs are generated via the nucleation-dependent mechanism, in which tau dimerization was postulated the necessary prerequisite for nucleation, i.e., the rate-limiting step of PHFs (Figure 3.4) assembly.[122]

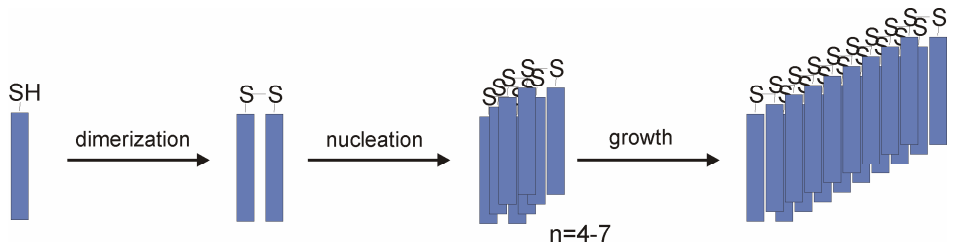


Figure 3.4 *Mechanism of nucleation-dependent PHF assembly. Tau monomers dimerize via oxidation of their SH groups. Four to seven of such dimeric subunits can form a nucleus, which subsequently grows to form mature PHFs.*

3.2.4 Possible mechanism of neurofibrillary degeneration

Data accumulated in the past two decades unambiguously show that the abnormality of tau is central to the pathology characteristic for AD and other tauopathies. Based on this knowledge a possible mechanism of neurofibrillary degeneration has been proposed (Figure 3.5).[12] According to this scenario, tau is a soluble neuronal protein whose main function of promoting and stabilizing microtubule assembly is regulated by phosphorylation. Due to the imbalance in phosphorylation/dephosphorylation events and/or impaired brain glucose metabolism, which downregulates tau O-GlcNAcylation, or possibly increased N-glycosylation, the phosphorylation level of tau drastically increases. This hyperphosphorylation results in the loss of tau activity and subsequent breakdown of microtubule networks, which in turn yields impaired axonal transport and concomitant neuron death. On the other hand, dissociation from microtubules increases the intraneuronal concentration of the free hyperphosphorylated tau, which is not only easier to polymerize to PHFs but also is a toxic molecule capable of sequestering normal tau and other microtubule associated proteins. In contrast to free hyperphosphorylated tau, PHFs appear to be inert, but they rapidly grow in size and thus may physically destroy affected neurons.

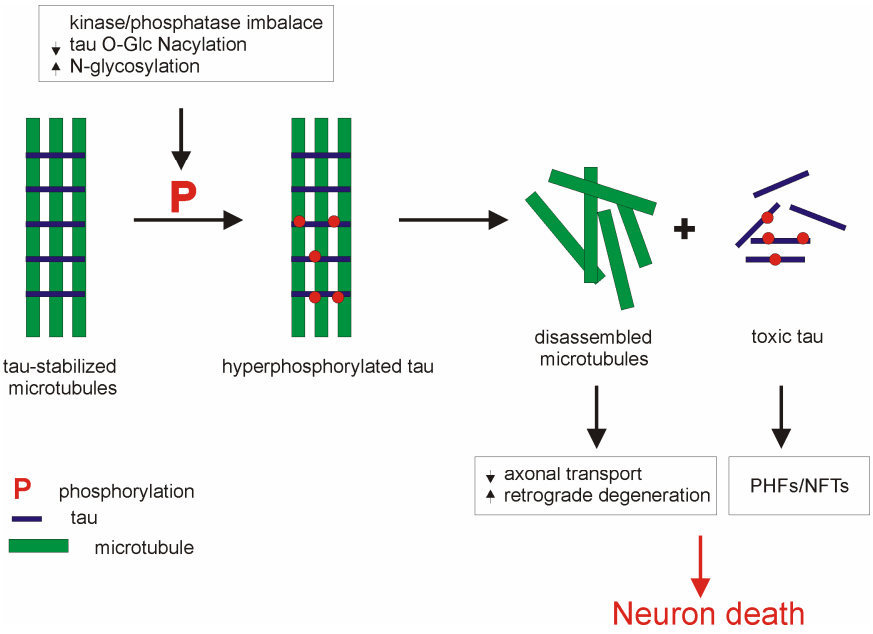


Figure 3.5 *Schematic representation of the neurofibrillary degeneration hypothesis.*

3.3 Phosphorylation of Amyloid Precursor Protein

Amyloid precursor protein (APP) is a single transmembrane protein whose biological function is currently unknown. It comprises a large extracellular N-terminal domain and approximately 55 amino acid long cytoplasmic tail that contains intracellular trafficking signals (Figure 3.6). Alternative splicing of mRNA generates several isoforms of the protein, of which APP695 is the predominant one in neurons. APP undergoes a series of N- and O-glycosylations followed by three major cleavage events catalyzed by a family of aspartyl proteases called α-, β-, and γ-secretase.[3] α-secretase cleavage delivers soluble, non-amyloidogenic products, whereas consecutive action of β- and γ-secretases generates amyloid β peptides, Aβ(1-40) and Aβ(1-42), whose aggregation into senile plaques is one of the characteristic features of AD (Figure 3.6).

Similarly to tau, APP is a phosphoprotein and several phosphorylation sites have been identified.[123-125] Among them Thr668 (numbering according to the 695 amino acid isoform) appears to be the main APP phosphorylation site, as it was found on mature and immature forms of APP as well as on APP originating from different cells.[126] The kinases that phosphorylate APP at Thr668 *in vivo* have been described. They include Gsk-3β,[127] cdk5,[128] cdc2,[48] as well as Jun N-terminal kinase-3 (JNK3).[129] It is interesting to note that some of these kinases have also been implicated in tau hyperphosphorylation.

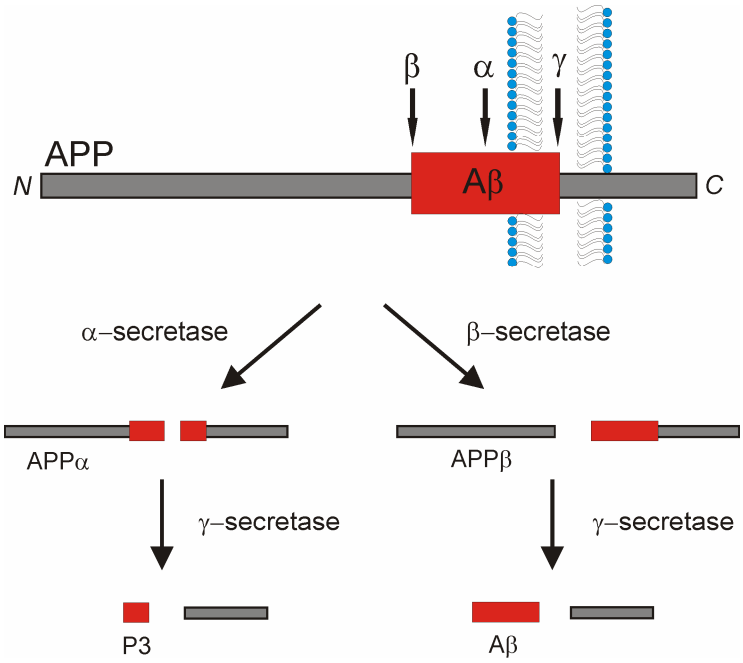


Figure 3.6 *APP processing by secretases.*

APP phosphorylation has a major impact on the regulation of its normal interactions as well as on abnormal events associated with AD. It has been demonstrated that Thr668 phosphorylation induces significant conformational change in the cytoplasmic region, which may affect APP interactions with its binding partners, and thus signal transduction.[126, 130] It was also shown that Thr668 phosphorylation may play an important role in neuronal differentiation.[131] Furthermore, several lines of evidence suggest that Thr668 phosphorylation may alter APP processing by secretases. Lee et al. reported that Thr668 phosphorylated APP is significantly increased in AD brain, that it colocalizes with β-secretase in neurons, and that this colocalization can be diminished by the Thr668 to Ala mutation.[132] Furthermore, these authors raise the possibility that phosphorylation of APP may facilitate the cleavage by β-secretase and thus influence the production of Aβ peptides. Accordingly, they have shown that Aβ generation in neurons was significantly reduced by the Thr668 to Ala mutation or by application of T668 kinase inhibitors. Together, these results suggest that Thr668 phosphorylation is a molecular mechanism that regulates APP cleavage by β-secretase.

In addition to APP the products of its cleavage, Aβ peptides, have also been modified by phosphorylation. Specifically, it was shown that human cdc2 kinase phosphorylates Ser26 in Aβ and this phosphorylation was found in both neuron and AD brain extracts.[133] Furthermore, inhibition of this phosphorylation either by a specific inhibitor or by mutation prevents the cytotoxicity of Aβ peptides, suggesting a role of Aβ phosphorylation in AD pathology.

3.4 Mechanisms linking PHFs and Aβ

Although neurofibrillar tangles and amyloid plaques have been usually regarded as independent pathological entities of AD, a body of evidence suggests that they may indeed be functionally linked. Based on the facts presented above, an obvious connection between tau and Aβ-induced neurodegeneration is the dysregulation of phosphorylation/dephosphorylation events. Several reports highlight that kinases known to be involved in tau hyperphosphorylation also regulate APP processing in a way that ultimately leads to the increase in Aβ levels.[134-136] Increased production of Aβ peptides, in turn, can regulate malprocessing of tau in several ways. First of all, upregulated Aβ(1-42) was shown to induce several tau kinases, including Gsk-3β and cdk5, which leads to tau phosphorylation and apoptosis.[137, 138] Second, Aβ peptides seem to upregulate C-terminal cleavage of tau by caspases, which yields accumulation of proteolytic products with enhanced aggregation kinetics.[49] Third, it has been demonstrated that increased Aβ levels can cause decreased glucose uptake/metabolism.[139] Reduced levels of glucose negatively regulate O-GlcNAcylation of tau, which in turn promotes hyperphosphorylation.[109] Contrary to the above, a key role of tau in mechanisms leading to Aβ-induced neurodegeneration has also been proposed.[140]

Our understanding of molecular mechanisms leading to AD has increased significantly in recent years. The disease appears to have many etiologies, which may cause or contribute to neurodegeneration via different pathways. Based on the accumulated research it is tempting to speculate that phosphorylation imbalance lies in the core of the disease. In order to validate this hypothesis, extensive studies of phosphorylation patterns in disease related peptides and proteins are of paramount importance. Unfortunately, due to the complex nature of molecular interactions as well as the challenging physicochemical properties of naturally aggregating systems their detailed structural and functional investigation is severely hampered. Furthermore, *in vivo* rapid turnover of phosphoproteins renders the identification of key phosphorylated residues difficult, whereas *in vitro* enzymatic phosphorylations generate heterogeneous mixtures of proteins phosphorylated at various sites and to a different extent. In order to overcome these obstacles, two strategies have been particularly useful. The first involves the application of simplified, synthetically accessible model systems. The second utilizes the semisynthetic approach, in which a protein is generated from two fragments of synthetic and recombinant origin to yield homogenous, site-specifically modified sample.

4 Coiled coil-based model of self-assembly

Since the discovery that various non-disease related proteins or even small peptides with different structural origins can self-assemble to form the common amyloid structure,[21, 22] it became evident that the molecular mechanisms that drive fibrillization can be efficiently studied in model systems. Owing to our increasing understanding of sequence-structure relationship,[141-143] various strategies ranging from mutational analysis to de novo approaches guided with computer-based structure predictions have been utilized to model the process of amyloid formation.[144] With respect to de novo design, which usually is both laborious and time-consuming, the most predictable models are based on the versatile coiled coil folding motif.[145] This is due to the fact that coiled coils have been extensively studied, which has allowed their structural features to be recognized and design principles established.[146] Moreover, the folding of coiled coils is based on oligomerization, which is also the foundation of amyloid formation.[147]

Coiled coil folding motif was first recognized and proposed by Crick during the investigation of X-ray diffraction pattern of α-keratin.[148] The first high-resolution crystal structure was solved some 40 years later and it basically confirmed Crick's predictions.[149] In nature coiled coils typically consist of two to five right-handed amphipatic α-helices, which assemble to form helical bundles with left-handed supercoils (Figure 4.1 A). An estimated 5-10 % occurrence in proteins,[150] including skeletal and motor proteins, transcription factors, and viral fusion proteins, highlights the common nature of this folding motif.[151] The primary structure of canonical coiled coils is characterized by a periodicity of seven residues, the so-called heptad, which is traditionally denoted $(abcdefg)_n$ (Figure 4.1 B).

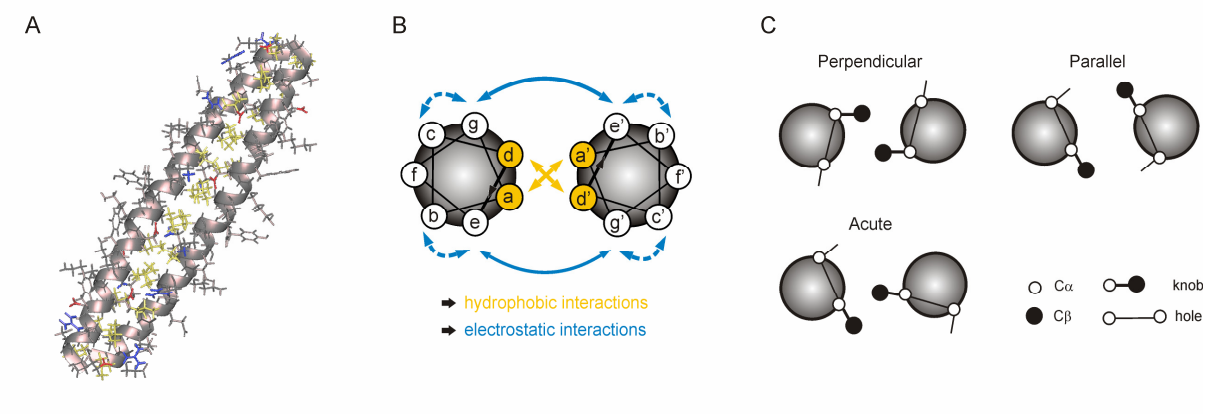


Figure 4.1 *The α-helical coiled coli folding motif. (A) Schematic representation of a dimeric coiled coil. (B) Helical wheel representation of dimeric coiled coil depicting the heptad repeat pattern and major interactions.(C) Different core-packing geometries (based on P. Harbury et al.)*[152]

Positions *a* and *d* are typically occupied by nonpolar residues (Leu, Val, Ile, Ala, Met) that form the first recognition domain by hydrophobic core packing ('knobs-into-holes').[146] Charged amino acids at positions *e* and *g* form the second recognition motif and engage in interhelical electrostatic interactions. Polar residues are often found at the remaining heptad repeat positions *b*, *c,* and *f*, which are solvent exposed and usually amenable for substitutions. The noncanonical coiled coils, in contrast, are built from nonheptad-based repeats, which usually leads to right-handed supercoils.[146]

Efficient packing of residues at *a* and *d* positions along the hydrophobic interface is a factor that provides the major contribution to the coiled coil stability. In addition, the core can also influence the oligomerization state. It has been demonstrated that the packing geometries of residues at *a* and *d* positions are different and called parallel and perpendicular, respectively (Figure 4.1 C).[152, 153] Thus, the identity of residues occupying these sites governs the oligomerization state of parallel coiled coils. Specifically, in dimers Ile and Leu are preferred at *a* and *d* position, respectively. In coiled coil tetramers this geometry is reversed. Whereas trimers, which show so-called acute packing (Figure 4.1 C), do not discriminate between both residues and are generated when *a*=*d*=Ile or Leu. Furthermore, the oligomerization state can also be dictated by other amino acid substitutions in the core positions. For example, placement of polar Asn residues in the hydrophobic environment favors dimer formation.[154] Additionally, Asn potential for hydrogen bonding can aid in specifying the helical alignment.[155]

Electrostatic interactions between positions *e* and *g* mainly provide the specificity of folding, i.e. parallel versus antiparallel arrangement as well as direct the preference for homo- or heterooligomer formation.[156-158] However, the share of attractive Coulombic interactions in stabilizing coiled coil assemblies, both in inter- and intrahelical fashion, has also been demonstrated.[159] In addition, it has been shown that charged amino acids at positions e and g can further aid the coiled coil stability via participation in the hydrophobic core interactions in their nonionic form.[160, 161] Careful adjustment of ionic interactions can thus be a powerful strategy in coiled coil design and application.

Based on the sequence-structure relationships described above as well as computational approaches that allow identification of heptad repeats in proteins and reliably predict the oligomerization state[162] a variety of coiled coil models of self-assembly have been generated in recent years.[145] As the fundamental feature of natural aggregating systems is their ability to change conformation in response to specific stimuli, coiled coil-based models usually follow this so-called 'switch concept.'[163]

The first example of a coiled coil-based model for monitoring α to β conformational transitions was developed by the Mihara group.[164] The design strategy involved a 14 residue coiled coil peptide dimer (Ad-2α) covalently linked via cysteine disulphide at the C-terminus

and equipped with the bulky adamantane groups at the N-terminus (Figure 4.2). The presence of such specific hydrophobic nucleation domain, 'hydrophobic defect', served as a conformational trigger in this system and subsequently yielded amyloid formation under neutral aqueous conditions and relatively short time. Furthermore, during the extensive investigation of transitional properties of the model it was demonstrated that the complexation of adamantanes with β-cyclodextrin inhibits conformational change and stabilizes the system in the α form. Despite the successful implementation of the described model in this investigation, the application of unnatural adamantane as the triggering factor definitely limits biological relevance of this particular design.

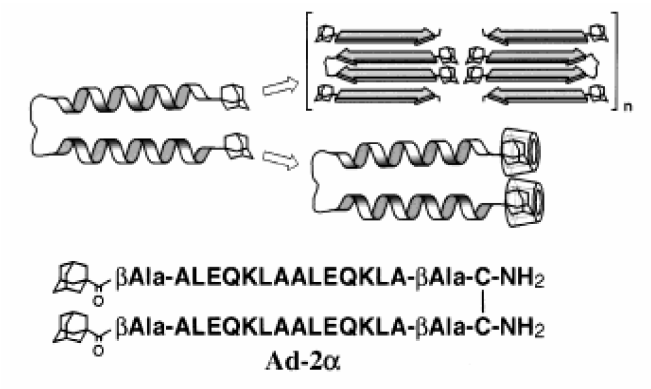


Figure 4.2 *Design of the coiled coil-based model peptide with hydrophobic defect. Amino acid sequence of Ad-2α and schematic representation of the α to β conformational transition as well as its prevention by cyclodextrin (reproduced with permission from Y. Takahashi et al.*[164] Copyright © *1998 WILEY-VCH Verlag GmbH, Weinheim)*.

Several coiled coil models of α to β conformational transitions in response to elevated temperature have been reported.[51, 147, 165] The common feature of these systems is that, unlike in the previous example, the intrinsic propensity for the β-structure is generated via the introduction of hydrophobic/β-sheet promoting residues at the *f* position of the coiled coil, thus eliminating the need for unnatural building blocks. It has been demonstrated that the role of these specific residues at *f* position is not only associated with the non-selective increase in system's overall aggregation propensity, but also with selective stabilization of the competing β-structure through the formation of a hydrophobic pocket comprising also traditionally hydrophobic residues at *a* and *d* position of the coiled coil.[147]

In contrast, Pagel et al. demonstrated that conformational transitions in coiled coils can be initiated without thermal activation.[50, 166] Instead, the authors took advantage of the significant role of intramolecular Coulombic interactions in coiled coil formation and stability. The design strategy involved two basic principles. First was the incorporation of three β-sheet inducing valine residues in solvent exposed positions of the coiled coil (Figure 4.3), which made the system prone to amyloid formation. The second point was careful

positioning of equally charged residues on the same side of the helical cylinder at solvent exposed heptad repeat positions (Figure 4.3), which ultimately rendered the model pH and ionic strength sensitive. Depending on the applied pH, peptide concentration as well as the presence of salts, remarkable structural flexibility of the system was generated. Moreover, two different kinds of aggregates were formed, namely α-helical and amyloid fibers. The same authors have also applied a coiled coil model to monitor metal-induced peptide self-assembly.[54]

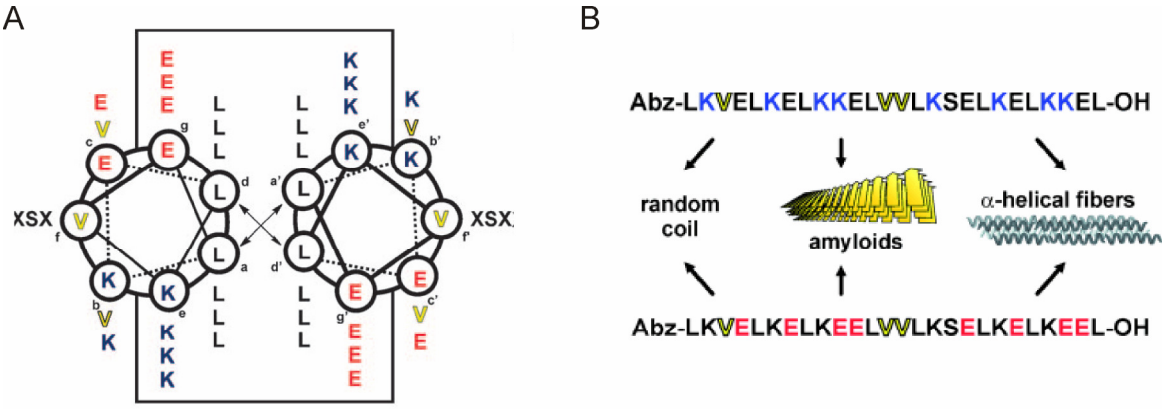


Figure 4.3 *Design of pH sensitive coiled coil model peptides. (A) Helical wheel representation. Residues warranting efficient coiled coil formation are framed. Positions of three valine residues, which make the system prone to amyloid formation, are highlighted in yellow, whereas the accumulation of negative or positive charge which renders the model pH sensitive is shown in red and blue, respectively. Residues denoted X are acidic or basic depending on the identity of the switch. (B) Sequences of switch peptides and environment-dependent structural diversity (reproduced with permission from K. Pagel et al.[166] Copyright © 2008 WILEY-VCH Verlag GmbH, Weinheim).*

Despite both the apparent abundance and reliability of coiled coil models of amyloid formation and the fact that coiled coils have been commonly applied to probe the impact of phosphorylation on the helical structure,[64, 65, 67] the system undergoing conformational transitions to yield β-structure with phosphorylation as a trigger has yet to be developed. This fact unambiguously highlights the great need for further development in this area as both the simplicity and versatility of coiled coil models could be of great help in understanding the basis of phosphorylation-induced structural changes leading to self-assembly.

5 Strategies for the preparation of phosphorylated peptides and proteins

Rapid access to peptides and proteins has been the primary objective for generations of chemists. In contrast to the isolation from natural sources and biochemical expression techniques chemical approach permits site-specific incorporation of certain modifications and the production of pure peptides in large quantities. Since the first chemical synthesis of a dipeptide by Emil Fisher in 1901 peptide synthesis has evolved dramatically.[167] A rapid development of orthogonal protecting groups as well as new coupling strategies, and the introduction of the automated solid-phase peptide synthesis (SPPS) by Bruce Merrifield in 1963[168] revolutionized the field. Nowadays peptides up to 50 amino acids can be routinely synthesized using two major SPPS techniques: original Merrifield's *tert*-butoxycarbonyl (Boc) SPPS methodology and a slightly more recent 9-fluorenylmethoxycarbonyl (Fmoc) SPPS. Due to the particularly harsh acidic conditions utilized by the Boc strategy phosphorylated peptides can be prepared exclusively via Fmoc SPPS.[169]

5.1 Phosphopeptide synthesis

Phosphopeptides are important tools for the investigation of protein phosphorylation and dephosphorylation processes. They have been applied as antigens for the production of phosphorylation-dependent antibodies, signal transduction triggers, and probes for determination of the amino acid sequence specificity of protein phosphatases.[170] Moreover, phosphopeptides have also been applied as model systems to study structural aspects of phosphorylation[63, 65, 66] and its role in the amyloid formation phenomenon.[171-173]

The synthesis of phosphopeptides can be accomplished via two general strategies: the first involves the use of pre-formed building blocks (protected phosphoamino acids) which are incorporated into synthetic peptides during the chain assembly. The second approach, so-called global phosphorylation, involves post-synthetic modification of serine, threonine, or tyrosine hydroxyl groups on the solid support.[174] The phosphorylation strategy for both amino acid building blocks and peptides is quite similar; it utilizes electrophilic phosphorylating agents, most commonly phosphoramidites and phosphorochloridates (Figure 5.1). Whereas the first class represents P^{III} reagents which require mild acidic reaction conditions and further oxidation to yield P^{V} species (Figure 5.1 A), the second group of compounds reacts under alkaline conditions and is already in the P^{V} oxidation state. Therefore it does not need further oxidation (Figure 5.1 B). The significantly lower reactivity of phosphorochloridates, however, limits their use and renders phosphoramidites suitable for most applications.[170]

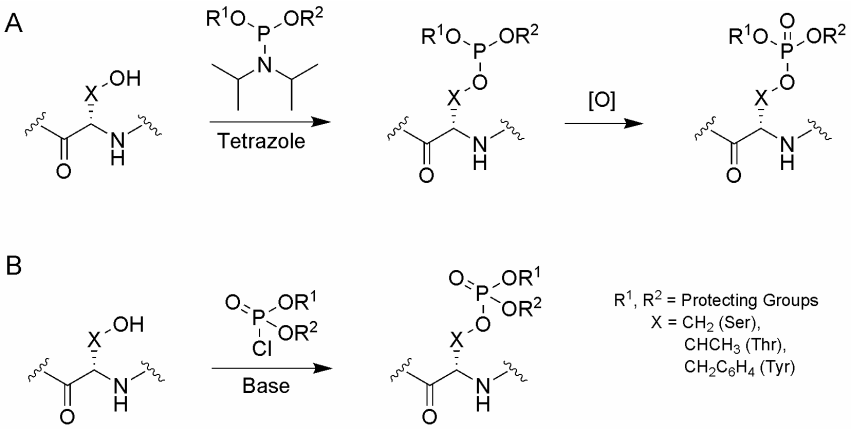


Figure 5.1 *Phosphorylation with (A) phosphoramidites and (B) phosphorochloridates.*

5.1.1 The building block approach

In the building block approach pre-formed, suitably protected phosphorylated amino acids are incorporated into synthetic peptides during the chain assembly utilizing standard Fmoc SPPS protocols (Figure 5.2 A).

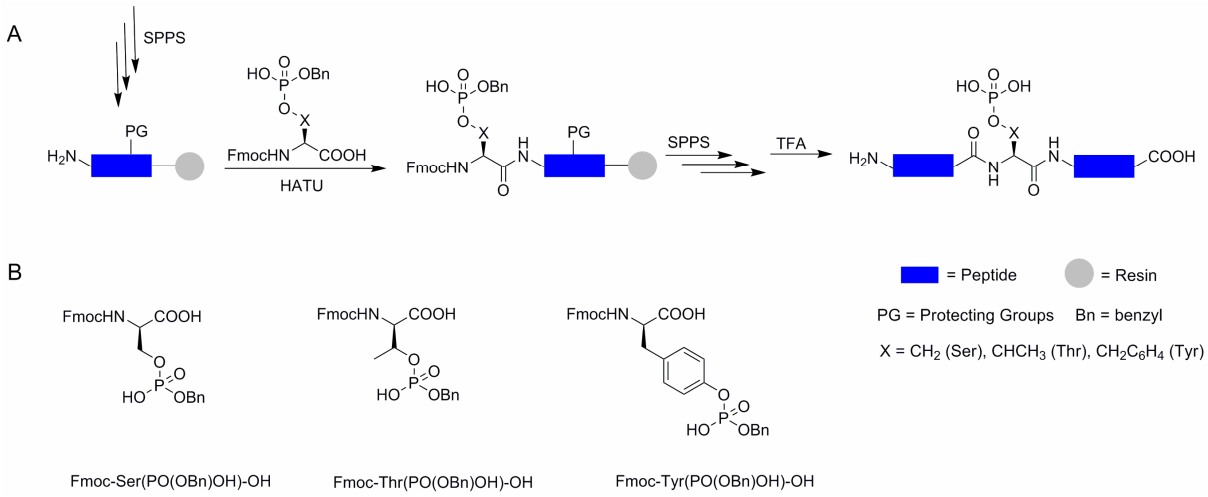


Figure 5.2 *The building block phosphopeptide synthesis. (A) General strategy. (B) Commercially available benzyl protected phosphoamino acid building blocks.*

Initially, it was believed that building block methodology was not compatibile with Fmoc strategy as phosphate triesters of serine and threonine underwent β-elimination upon piperidine treatment resulting in the formation of respective dehydroamino acids. It was the work of Wakamiya et al.[175] demonstrating that the application of mono-protected derivatives can completely eradicate the side-reaction that finally allowed the application of the building block strategy to Fmoc SPPS. Currently mono-protected phosphoamino acids have become

commercially available and mono-benzyl derivatives (Figure 5.2 B) have been particularly useful as they can be deprotected with trifluoroacetic acid (TFA) treatment, which at the same time yields global deprotection as well as cleavage from the solid support.

5.1.2 Global phosphorylation

In the global phosphorylation strategy (Figure 5.3) phosphorylation is performed after the complete peptide assembly. The residues to be phosphorylated may be introduced with or without an orthogonal side-chain protection which prior to the reaction needs to be selectively removed. In order to avoid post-phosphorylation piperidine treatment, which renders phosphate triesters of serine and threonine susceptible to β-elimination, the N-terminal amino acid should be incorporated as an *N*-Boc derivative or alternatively it can be Fmoc deprotected and capped prior to the phosphorylation reaction. The modification of resin-bound hydroxyl groups has been in most cases carried out using di-*t*-butyl and dibenzyl phosphoramidites as their respective protecting groups are TFA labile.[174]

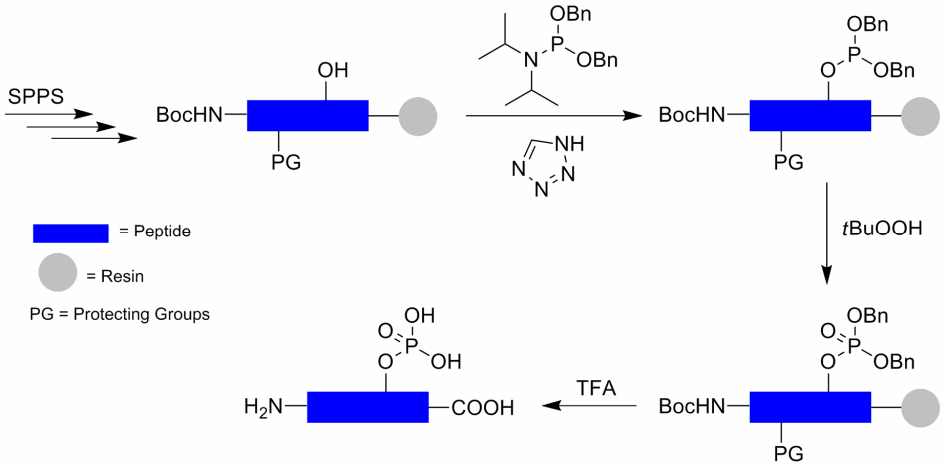


Figure 5.3 *The global phosphorylation approach.*

As described above, phosphorylation with phosphoramidites is a two step procedure, involving initial phosphitylation in the presence of tetrazole and subsequent oxidation utilizing compounds such as *m*-chloroperoxybenzoic acid, aqueous iodine, or *t*-butyl hydroperoxide. The use of *t*-butyl hydroperoxide gives superior results as it does not seem to cause the oxidation of methionine, tryptophan, and cysteine residues.[174] Similarly, benzyl protected phosphoramidites are usually superior to the *t*-butyl protected ones as they are less susceptible towards acid-catalyzed decomposition yielding H-phosphonate formation.[176]

The major advantages of the global phosphorylation approach include the possibility of multiple residue modification in the course of a single phosphorylation procedure and

furnishing both phosphorylated and unphosphorylated peptides in the same synthesis. However, because of the commercial availability of suitably protected phosphoamino acids, this methodology is nowadays a second choice technique and the majority of phosphopeptides are synthesized via the straightforward building block approach.

Even though Fmoc methodology can be readily applied to produce site-specifically phosphorylated peptides in large quantities, the fundamental limitation of SPPS to around 50 amino acids renders chemical access to phosphorylated proteins troublesome. These targets, however, can be accessed via fragment condensation of partially protected peptides or chemoselective ligation of fully unprotected ones. The latter approach has gained particular attention in the area of protein chemistry, as it obviates the necessity for protecting groups on highly complex biological molecules.

5.2 Chemoselective ligation strategies yielding native peptide bond

The interest in chemoselective ligation and modification strategies for peptides and proteins has grown exponentially in recent years. The great advantage of chemoselective ligation reactions is naturally the possibility of joining together two fully unprotected peptide fragments, but also the fact that these reactions proceed rapidly under mild conditions (natural pH, aqueous media, and room temperature). Beginning with the pioneering work of Wieland on the chemical properties of amino acid thioesters[177] and followed by Kemp's prior thiol capture strategy[178, 179] the field evolved dramatically. Various chemoselective ligation strategies for peptides and proteins have been developed,[13] but the most applicable ones are those which yield the native peptide bond.

5.2.1 Native chemical ligation and expressed protein ligation

The most widely used chemoselective ligation technique called native chemical ligation (NCL) was introduced in 1994 by Kent and co-workers.[180] In NCL two fully unprotected peptides, one containing a C-terminal thioester and another with an N-terminal cysteine are joined together via a native peptide bond (Figure 5.4). The reaction mechanism involves the so-called capture/rearrangement concept, where the initial, reversible transthioesterification constitutes the capture step whereas the following spontaneous and irreversible intramolecular S→N acyl shift comprises the rearrangement. Most importantly, the presence of other cysteine residues in the peptide sequence does not interfere with the reaction, as the rearrangement step is only possible at the unique N-terminal one.

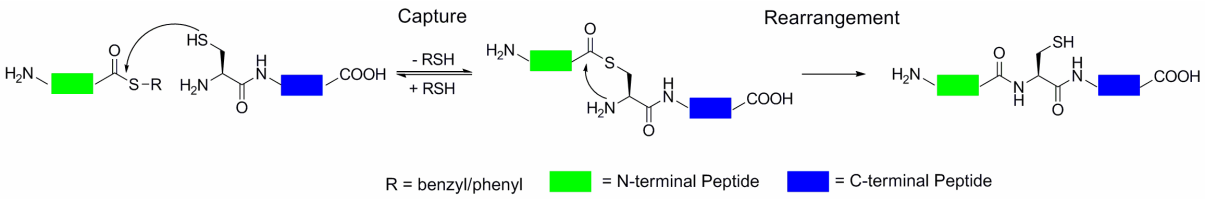


Figure 5.4 *The concept of native chemical ligation (based on C. P. R. Hackenberger et al.)*[13]

NCL is usually performed in buffered aqueous solutions at pH in the range of 7-8. Maintaining this pH regime is important as strong basic conditions favor thioester hydrolysis and render ε-amino groups of lysines nucleophilic, whereas acidic conditions decrease the nucleophilic activity of both cysteine thiol and N-terminal amine. The reaction has a great tolerance towards additives, such as chemical denaturants, detergents, reducing agents, and organic solvents. Furthermore, several factors that affect the rate of NCL have been identified. The work of Dawson and co-workers revealed that the identity of the amino acid bearing the α-thioester group influences the ligation efficiency significantly.[181] The fastest reaction was shown for glycine whereas the slowest transformation was observed for β-branched amino acids and proline. In addition, the nature of the thioester itself has a profound impact on NCL and aryl thioesters have been found more reactive than the alkyl ones.[182]

Despite its utility and robustness, NCL has two main limitations. The first involves the restricted access to peptide thioesters, which due to their base sensitivity were mostly prepared by Boc SPPS. Unfortunately, the strong acid conditions used in this strategy are not compatible with specific side chain modifications like phosphorylation and glycosylation. The second limitation has to do with the fact that cysteine is quite rare amino acid in proteins (1.4% content). In consequence, an artificial cysteine often needs to be introduced to generate a ligation site for NCL, which in turn is not always tolerated in terms of structure and function of the target protein. In order to overcome the above obstacles, several refinements have been developed.

With regard to peptide thioesters synthesis a number of Fmoc based strategies, which mainly relay on the post-synthetic modification of the C-terminus of otherwise fully protected peptide, have been introduced.[13] The most prominent example here is the approach utilizing the alkanesulfonamide safety catch resin, which allows the peptide to be activated with diazomethane or iodoacetonitrile and subsequently cleaved with a specific thiol to yield the corresponding thioester (Figure 5.5 A).[183] Based on this methodology phosphopeptide thioesters have been successfully prepared[184] and a promising thioester SPPS with self-purification effect was introduced by the Seitz laboratory.[185] Another common strategy for Fmoc-based thioester synthesis employs hyper-acid labile 2-chlorotrityl resin, which

allows selective cleavage of protected peptide from the solid support under mild acidic conditions (Figure 5.5 B). The thioesterification is accomplished via the activation of the C-terminal carboxyl group, followed by the thiol coupling.[186] Due to a risk of enantiomerization during the activation step, this approach is mostly suitable for peptides with C-terminal glycine residues. In both mentioned examples the final TFA treatment furnishes unprotected thioesters that can be further used in NCL.

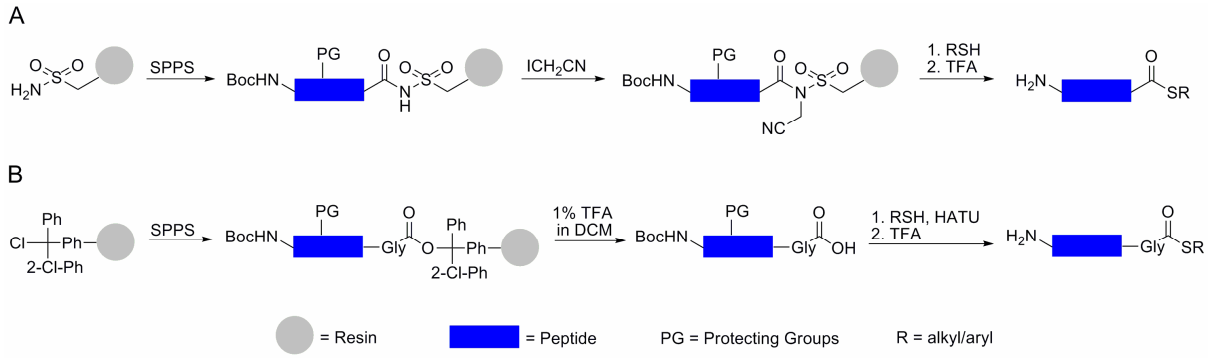


Figure 5.5 *Fmoc-based strategies for generating peptide thioesters using (A) safety catch resin and (B) 2-chlorotrityl resin (based on C. P. R. Hackenberger et al.)*[13]

Significant development has been made in order to overcome the second limitation of NCL, i.e. the necessity for cysteine at the ligation site. The most prominent examples include: (1) the development of acid or UV labile thiol auxiliaries,[187, 188] (2) the application of desulfurization strategy[189, 190] in combination with β-mercapto building blocks[191, 192] that facilitated ligations at alanine, valine, and phenylalanine, and finally (3) alkylation of partially protected NCL products that enabled ligations at serine as well as methionine, lysine, glutamine, and glutamate analogues.[13]

Even though NCL can be applied to both synthetic and recombinant substrates, the lack of chemical approaches to prepare recombinant thioesters forms an obstacle for the use of this methodology in protein semisynthesis. This problem was elegantly solved by Muir et al. who inspired by the process of protein splicing developed semisynthetic variant of NCL called expressed protein ligation (EPL).[193]

Protein splicing is a post translational modification in which a precursor protein undergoes a series of intramolecular rearrangements that result in the excision of the internal domain (an intein) and concomitant ligation of the flanking segments (exteins) via a native peptide bond.[194] The process itself is autocatalytic and the underlying mechanism involves in principle three steps (Figure 5.6A). First the cysteine residue located at the N-terminus of the intein undergoes an N→S acyl shift to yield thioester linkage between N-extein and the intein. Next the entire N-extein is transferred to a downstream cysteine at the intein-C-extein boundary in a transthioesterification step. Finally, the intein excises itself

as a C-terminal succinimide derivative via a cyclization reaction involving a conserved asparagine at its C-terminus, and subsequently the two exteins are joined by the native peptide bond.

In order to take advantage of inteins' catalytic activity for the recombinant thioesters production the target protein is usually expressed as in-frame N-terminal fusion to a specifically mutated intein that catalyzes only the first step of the splicing reaction. Subsequent thiol addition furnishes corresponding protein thioester in a simple transthioesterification step. The whole procedure is usually performed on an affinity resin as commercially available intein fusions are equipped with an affinity tag, such as chitin binding domain (CBD) for purification purposes (Figure 5.6 B). Following the cleavage, the protein thioester is isolated by straightforward elution from the resin and can undergo NCL with a synthetic N-terminal cysteine-containing peptide usually bearing the desired modification, e.g. phosphorylation. Alternatively, the ligation product can be generated *in situ* during the thiol cleavage as originally proposed by Muir.[193]

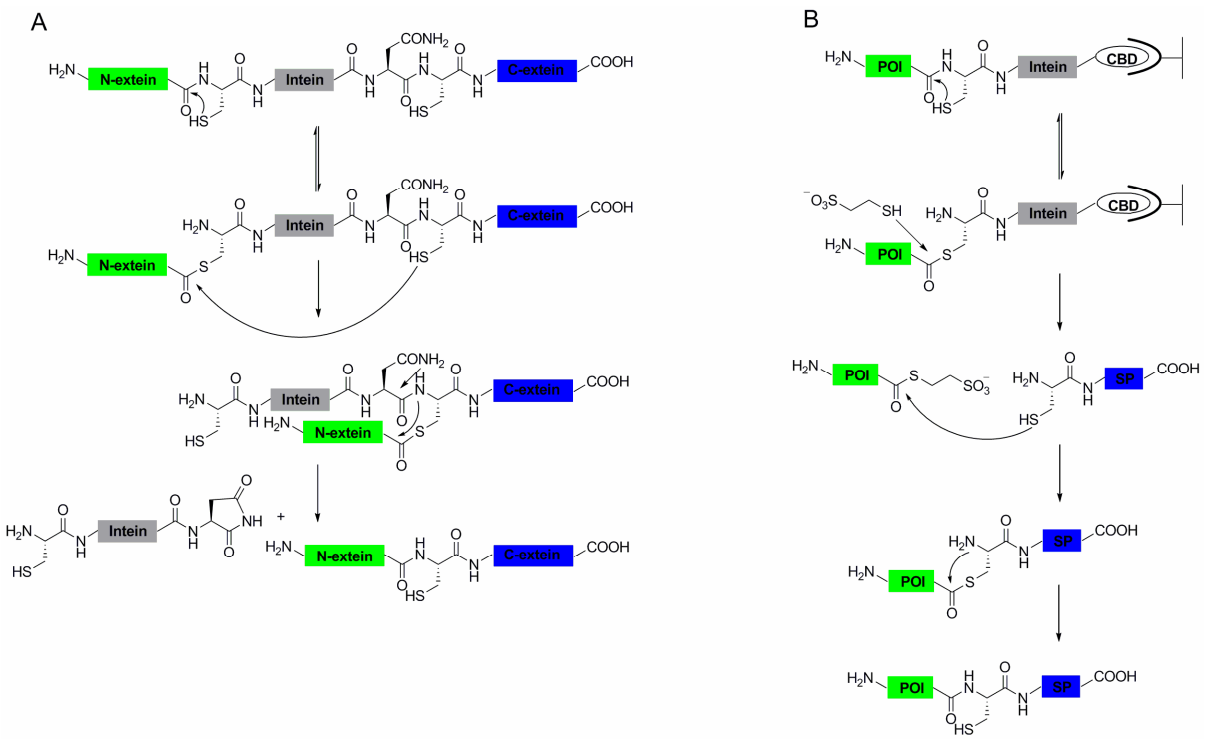


Figure 5.6 *Schematic illustration of (A) protein splicing and (B) expressed protein ligation (based on C. P. R. Hackenberger et al.)[13] POI = protein of interest, SP = synthetic peptide, CBD = chitin binding domain*

Specifically engineered inteins are also applied as tools for the preparation of recombinant proteins with N-terminal cysteines.[195] Here, the target protein is expressed as C-terminal intein fusion, and pH and temperature are usually used as cleavage inducers. On the other hand, these recombinant targets can also be generated via enzymatic

processing of a genetically introduced leader sequence (initiating methionine, protease recognition motif).[195]

In its simplest form EPL is applied to produce regioselectively modified semisynthetic targets from two fragments of recombinant and synthetic origin. This approach allows for the desired modification to be introduced only within the flanking regions of the protein. However, more sophisticated strategies for three fragment ligations have been developed and thus enabled the whole protein modification. In these so-called tandem ligations the middle segment of the target protein needs to be equipped with a temporary protection of the N-terminal cysteine in order to avoid self-ligation in an inter- or intramolecular fashion. [196, 197]

Since its introduction in 1998 EPL has been applied for the semisynthesis of proteins containing a variety of modifications,[195, 198] e.g. post-translational modifications,[199-201] unnatural amino acids,[202, 203] and spectroscopic probes.[197, 204] It is noteworthy however that EPL is actually the first choice technique for the preparation of site-specifically phosphorylated proteins. The most prominent examples include: the proof of principle experiment of Muir, where tyrosine phosphorylated semisynthetic Crk kinase was generated;[193] the synthesis of hyperphosphorylated version of the type 1 transforming growth factor β receptor (TβR-I),[184] which allowed the molecular mechanism of receptor activation to be elucidated; the generation of phosphorylated Smad transcription factors[205, 206] together with the preparation of photocaged derivatives[207] that revealed the structural basis for Smads oligomerization and allowed for the spatiotemporal control of this process, respectively. Finally taking into account the dynamic nature of phosphorylation, EPL was also utilized to introduce non-hydrolysable mimics of phosphates into proteins.[208, 209]

One particular drawback of EPL applications is the difficulty of performing this reaction *in vivo* due to side reactions of thioester and N-terminal cysteine with endogeneous metabolites. However, the related and also intein-based strategy called protein *trans*-splicing (PTS)[210] holds great potential for a selective chemical modification in living systems.

5.2.2 Other chemoselective ligation strategies yielding native peptide bond

Even though NCL and EPL are the most useful strategies for selective protein modification and engineering, they both have their flaws and therefore other approaches that yield native peptide bond have been developed.[13] Although these ligation strategies have not been used so far as tools for the generation of phosphorylated peptides and proteins, they may offer such possibility in the future and therefore should be mentioned here. The two most promising approaches include traceless Staudinger ligation and decarboxylative amide ligation (Figure 5.7).

The first strategy, based on the known Staudinger reaction,[211] was introduced independently by Raines[212] and Bertozzi[213] in 2000. It involves ligation reaction between two fragments equipped with an azide and a phosphinothioester, respectively, to yield an iminophosphorane (Figure 5.7 A). This intermediate subsequently undergoes an intramolecular S→N acyl shift to generate a native peptide bond. Although initially the chemoselectivity of this approach was not often addressed, this strategy was successfully applied for the chemoselective cyclization of unprotected peptides, in a recent study.[214] Moreover it was demonstrated that the reaction can proceed in aqueous media.[215] Another advantage of this strategy is that it does not require cysteine at the ligation site, but efficient ligations have only been demonstrated for non-sterically hindered junctions.[212, 213]

The second methodology, the decarboxylative amide ligation, was recently introduced by Bode et al.[216] It involves a decarboxylative condensation of *N*-alkylhydroxylamines with α-ketoacids, which chemoselectively furnishes a native peptide bond (Figure 5.7B). Intensive studies demonstrated that the reaction proceeds efficiently at 40 ºC in polar solvents, does not require added reagents and produces no by-products. Most importantly, unprotected peptides containing potential nucleophilic side chains can be chemoselectively ligated at Phe-Ala, Pro-Ala, Val-Gly, and Ala-Ala junctions.

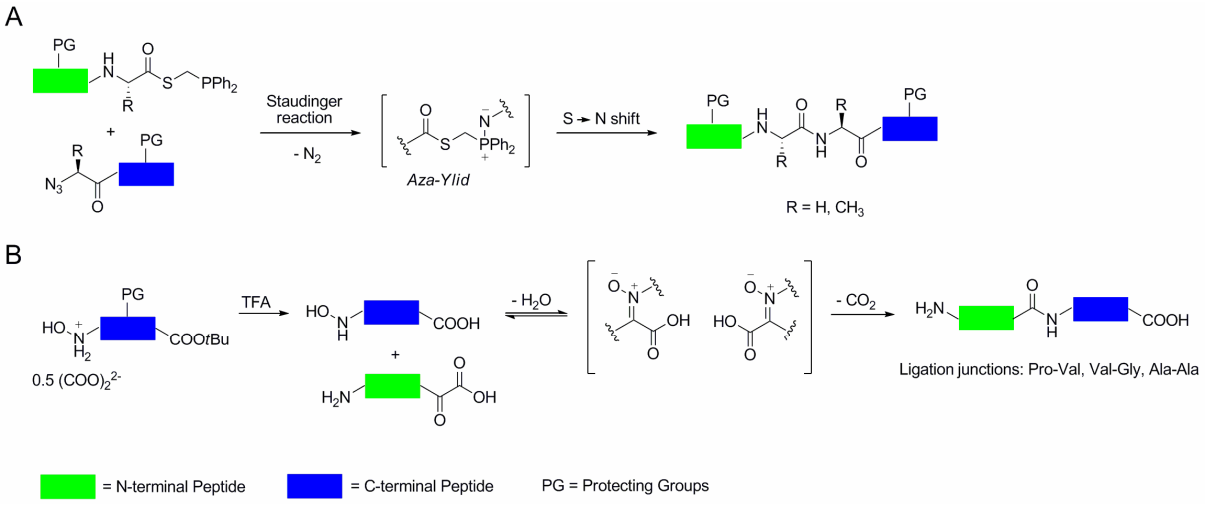


Figure 5.7 *Schematic illustration of (A) traceless Staudinger reaction and (B) decarboxylative amide ligation (besed on C. P. R. Hackenberger et al.)*[13]

6 Aim of the work

Over the last two decades it has become apparent that phosphorylation imbalance plays a central role in Alzheimer's disease (AD) pathology. Despite the significant progress that has been made in elucidating the underlying structural transformations as well as identifying responsible enzymes and aberrant phosphorylation sites, the detailed characterization of molecular events and specific phosphorylation patterns that trigger the transformation from functional protein to pathological aggregate still remains a challenge. The difficulty lies in the complex nature of molecular interactions as well as low solubility and poor synthetic access to naturally aggregating systems. Additionally, the dynamic nature of phosphorylation *in vivo* and heterogeneous character of enzymatic phosphorylations *in vitro* renders the identification of key phosphorylated residues difficult. The development of strategies that help to overcome these obstacles represents one of the biggest targets of current AD-related research.

The aim of this work was to apply synthetic phosphopeptides and phosphoproteins as reliable tools for detailed investigation of peptide and protein aggregation utilizing a battery of analytical techniques. In the first part of the project, coiled coil-based phosphopeptides served as synthetically accessible and easy to handle models to study the impact of phosphorylation on amyloid formation and as substrates for phosphatase triggered self-assembly to monitor discreet structural transformations and aggregation kinetics. Furthermore, coiled coil-based phosphopeptides were also utilized to investigate structural consequences of interactions between peptides and metal ions. In the second part of the project, synthetic phosphopeptides were applied to produce homogeneous preparations of site-specifically phosphorylated tau protein via expressed protein ligation. Specifically, this investigation involved the semisynthesis of tau protein phosphorylated in its C-terminal part, precisely at the so-called PHF-1 epitope (Ser396/Ser400/Ser404), which is characteristic for AD tau.

7 Applied analytical methods

Peptides and proteins involved in the process of amyloid formation can be investigated by a variety of bioanalytical and biophysical techniques, including solid state NMR, X-ray fiber diffraction, CD spectroscopy, FTIR, dye-induced fluorescence, as well as electron and atomic force microscopy. Techniques routinely applied by the author for monitoring structural transitions of investgated systems and for detailed kinetic analysis of amyloid formation, i.e. CD spectroscopy and ThT-induced fluorescence, respectively, are briefly described below.

7.1 Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is one of the most widely applied techniques for determining protein conformation in solution.[217] In contrast to X-ray crystallography and NMR, which can provide structural information in atomic detail, CD is a low resolution technique that describes overall structural features of a protein. It is however non-destructive and definitely less demanding in terms of time and sample quantities. Moreover, CD allows determination of protein structure under a variety of experimental conditions as well as detailed monitoring of conformational transitions of the studied system. Both of these features were extensively explored for experimental work described in this thesis.

The physical principle of the method is the differential absorption of left and right circularly polarized light by chromophores in a chiral sample. This phenomenon can be mathematically described according to the Lambert-Beer's Law as follows:

$$\Delta A = A_L - A_R = \varepsilon_L c l - \varepsilon_R c l = \Delta \varepsilon c l \qquad (1)$$

where A is the absorption, ε is the molar extinction coefficient, c is the molar concentration of the sample, and l is the path length. The term $\Delta\varepsilon$ is the circular dichroism.

Left and right circularly polarized light are equal magnitude components of plane polarized light. When they are not absorbed or absorbed to an equal extent by the analyte, the recombination of both beams will yield a straight line and thus the radiation will be polarized in the original plane (Figure 7.1 A). If, however, both components are absorbed to different extent, the elliptically polarized radiation will be generated instead and give rise to circular dichroism (Figure 7.1 B).

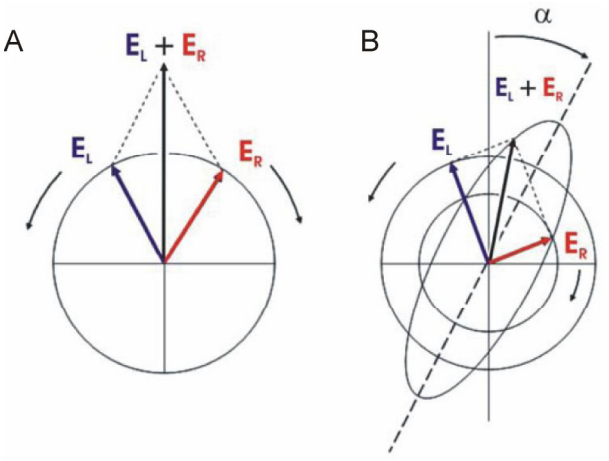


Figure 7.1 *Field vectors of left (blue) and right (red) circularly polarized light and their superposition (black), before (A) and after (B) passage through an optically-active sample. The angle α is the optical rotation, i.e. the angle between the major axis of the ellipse and the initial plane-polarized light beam (reproduced form K.Pagel)*[218]

Even though CD instruments record the difference in absorption, CD is usually reported in degrees of ellipticity θ, which is defined as the angle between the minor and major axis of the ellipse. The relation of ΔA and θ can be described by a simple mathematical equation:

$$\theta = 32.98\ \Delta A \qquad (2)$$

This formula in turn can be converted to mean molar residue ellipticity $[\theta]$, which removes the linear dependence of the signal on path length, sample concentration, and chain length and thus allows comparison of CD spectra generated for different proteins. The resulting equation looks as follows:

$$[\theta] = \theta/10000\ c\ l\ n \qquad (3)$$

where θ is the measured ellipticity in mdeg, c is the sample concentration in mol/l, l is the path length in cm, and n is the number of residues.

Proteins and peptides are composed of enantiomeric amino acids (except for glycine) which are connected by chromophoric peptide bonds. These two particular features render them most suitable targets for analysis by CD spectroscopy. Peptide bonds display two characteristic absorption bands in the far-UV region (240 nm and below) of electromagnetic spectrum: a weak but broad n→π* transition around 220 nm and a more intense π→π* transition around 190 nm.[219] Specific arrangement of peptide bonds in various types of secondary structure gives rise to characteristic CD signatures in far-UV (Figure 7.2).

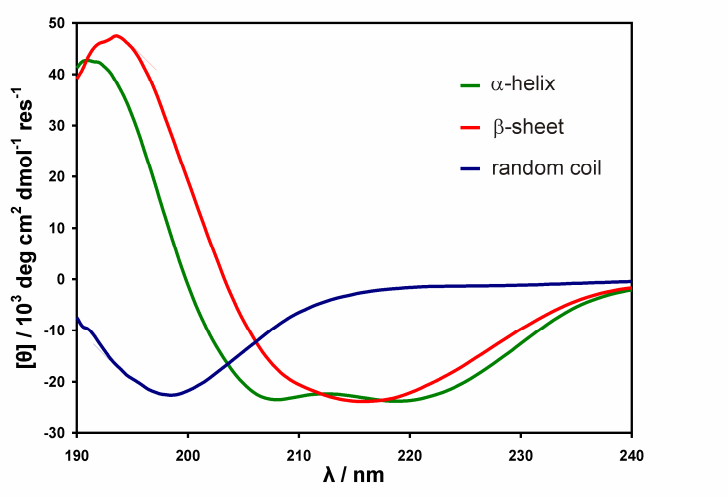


Figure 7.2 *Characteristic far-UV spectra of α-helix, β-sheet and random coil.*

The spectrum of α-helix is characterized by two minima at 208 nm (π→π*) and 222 nm (n→π*) as well as one maximum at 190 nm (π→π*). β-sheets possess one broad minimum around 215 nm (π→π*) and a maximum at 195-198 nm (n→π*), whereas random coils are usually recognized by one minimum at approximately 200 nm (n→π*). This standard set of CD spectra of isolated α-helices, β-sheets, and coils was experimentally obtained utilizing a variety of systems, including pH dependent conformations of poly-L-lysine,[220] globular proteins,[221] and model peptides.[222] In order to obtain structural characterization of more complex systems, the deconvolution, i.e. individual contribution of each particular structure is necessary. Deconvolution is usually accomplished by the application of sophisticated computer-based approaches.[223]

7.2 Dye binding studies

The ability to bind various dye molecules is one of the characteristic features of amyloid (see Section 2.1). Accordingly, it has been frequently applied as a diagnostic tool for amyloid identification *in vivo* but also for monitoring amyloid fibril formation *in vitro*. One of the most frequently utilized dyes is Thioflavine T (ThT) – a cationic benzothiazole dye (Figure 7.3 A), which upon amyloid binding and excitation at λ_{ex} = 450 nm shows enhanced fluorescence at λ_{em} = 485 nm. Even though these fluorescence properties of ThT are quite well understood,[224] the mechanism of dye binding to amyloid has been a subject of debate. In the study of Khurana at al. it was suggested that due to its amphiphilic structure, i.e. hydrophobic dimethylamino group attached to an aromatic ring and a more polar benzothiazole group (Figure 7.3 A), ThT can form micelles in aqueous solutions and bind to amyloid fibrils in this form.[225] Krebs et al. however proposed that dye molecules can bind to

amyloids in channels formed by side chains protruding form hydrogen bonded β-strands in β-sheets (Figure 7.3 B).[226] This binding occurs parallel to the long fibril axis and perpendicular to the β-strands (Figure 7.3 C). The proposed model also explains the increase in ThT fluorescence upon amyloid binding, the specificity of this effect and its varying efficiency as a result of steric interactions of the dye with the neighboring side chains.

Despite the still not fully understood nature of ThT interaction with amyloid as well as the dye's quite hydrophobic character and reported aggregation tendency,[225] the application of Thioflavin T for amyloid studies has become a standard and much appreciated technique. The great advantage of this method is that it allows both kinetic and quantitative data acquisition due to the fact that the fluorescence emission of the dye strongly depends on the amount of amyloid present. Monitoring the increase in ThT fluorescence in time is the fastest, most convenient, and most frequently used approach for recording fibril formation kinetics. For these particular reasons this method was extensively applied in the majority of studies performed herein.

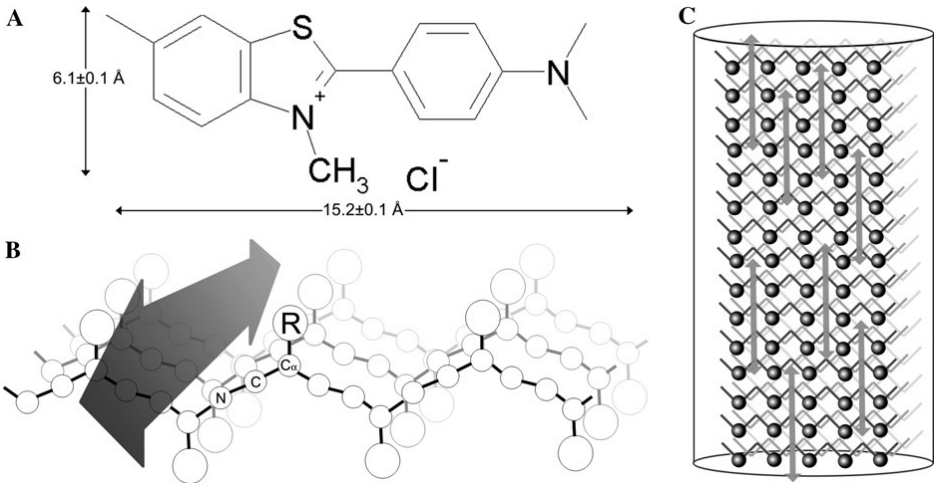


Figure 7.3 *ThT binding to amyloid fibers. (A) Structure and dimensions of ThT molecule (the thickness is 4.3 ± 0.1 Å). (B) Schematic representation of a β-sheet with backbone atoms (N, C and C_α) and a side chain group (R) indicated. ThT binding is represented by a double headed arrow and occurs in a channel between the side chains. (C) Schematic representation of a protofilament with ThT molecules (double headed arrows) bound parallel to the fibril axis and perpendicular to the β-strands (reproduced with permission from M. R. H. Krebs et al.[226] Copyright © 2005, Elsevier).*

8 Results and discussion

The results presented in this chapter have been published in the following peer-reviewed reports:

- M. Broncel,* J. A. Falenski,* S. C. Wagner, C. P. R. Hackenberger, B. Koksch, How post- translational modifications influence amyloid formation: a systematic study of phosphorylation and glycosylation in model peptides, *Chem. Eur. J.* **2010**, *16*, 7881.
  * (co-first authors)

- M. Broncel, S. C. Wagner, C. P. R. Hackenberger, B. Koksch, Enzymatically triggered amyloid formation: an approach for studying peptide aggregation, *Chem. Commun.* **2010**, *46*, 3080.

- M. Broncel, S. C. Wagner, K. Paul, C. P. R. Hackenberger, B. Koksch, Towards understanding secondary structure transitions: phosphorylation and metal coordination in model peptides, *Org. Biomol. Chem.* **2010**, *8*, 2575.

Results that are not part of these publications are presented in section 8.4.

8.1 How post-translational modifications influence amyloid formation: a systematic study of phosphorylation and glycosylation in model peptides

The results presented in this section have been originally published as: M. Broncel,* J. A. Falenski,* S. C. Wagner, C. P. R. Hackenberger, B. Koksch, How post- translational modifications influence amyloid formation: a systematic study of phosphorylation and glycosylation in model peptides, *Chem. Eur. J.* **2010**, *16*, 7881. * (co-first authors)

The original paper with supporting information is available at:

http://dx.doi.org/10.1002/chem.200902452

Individual contributions: M. Broncel synthesized all phosphorylated peptides and performed their CD spectroscopic and ThT fluorescence analysis, J. A. Falenski synthesized all glycosylated peptides and performed their CD spectroscopic and ThT fluorescence analysis, S. C. Wagner performed the TEM analysis.

Paper summary

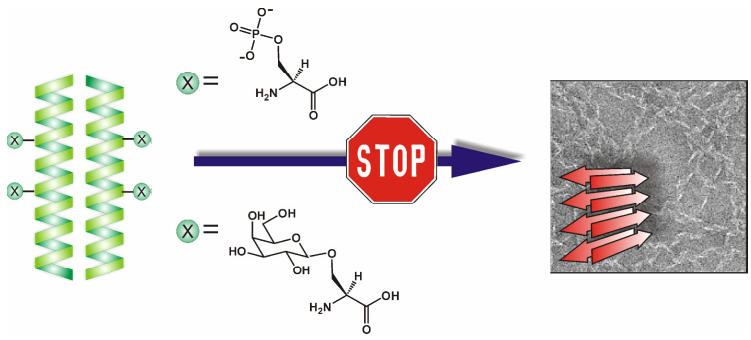
A reciprocal relationship between phosphorylation and O-glycosylation has been reported for many cellular processes and human diseases. Furthermore, accumulated evidence points to a significant role of these post-translational modifications (PTMs) in aberrant protein aggregation. Elucidating how phosphorylation and glycosylation influence amyloid formation is therefore of great interest.

The objective of this study was to apply an amyloid forming coiled coil peptide as a simplified system that allows detailed investigation of the impact of site-specific and multiple phosphorylation and O-glycosylation on the amyloid formation process. Accordingly, a set of model phosphopeptides and glycopeptides modified at the same solvent exposed *f* positions of the coiled coil was synthesized via SPPS. Phosphorylation was introduced in the form of phosphorylated serine whereas O-glycosylation in the form of synthetically accessible β-galactose also appended to the serine side chain. All peptides were analyzed in terms of their structural properties using CD spectroscopy and in terms of their aggregation propensity using Thioflavine T fluorescence. Furthermore TEM was applied to characterize amyloid fibers generated in this study.

It was demonstrated that phosphorylation and glycosylation affect folding and aggregation kinetics differently. Incorporation of phosphoserines, regardless of their quantity

and position, turned out to be most efficient in preventing β-sheets formation and their lamination towards amyloid. In contrast, O-glycosylation has a more subtle effect. The introduction of a single β-galactose does not change the folding behavior of the model peptide, but it does alter the aggregation kinetics in a site-specific manner. The presence of multiple galactose residues has an effect similar to that of phosphorylation.

In conclusion, the application of a simplified and reliable coiled coil model helped elucidate the influence of two most ubiquitous post-translational modifications on amyloid formation process. Taking into account that the specific role of PTMs in aberrant protein aggregation remains a subject of debate, the results presented have shed new light on this complex phenomenon.



GRAPHICAL ABSTRACT: An important step towards the understanding of the impact of phosphorylation and glycosylation on amyloid formation has been made. It has been demonstrated that even single phosphorylation is sufficient to completely inhibit fibrilization, whereas glycosylation showed more diverse effects.

8.2 Enzymatically triggered amyloid formation: an approach for studying peptide aggregation

The results presented in this section have been originally published as: M. Broncel, S. C. Wagner, C. P. R. Hackenberger, B. Koksch, Enzymatically triggered amyloid formation: an approach for studying peptide aggregation, *Chem. Commun*. **2010**, *46*, 3080.
Reproduced by permission of The Royal Society of Chemistry.
The original paper with supporting information is available at:

http://dx.doi.org/10.1039/c001460e

Individual contributions: M. Broncel synthesized all peptides, performed the CD spectroscopy, ThT fluorescence and HPLC-based kinetic measurements, S. C. Wagner performed the TEM analysis.

Paper summary

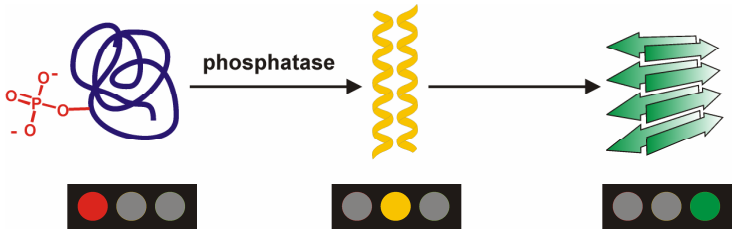
It is commonly accepted that amyloid formation is primed by the presence of partially folded intermediates. Accordingly, factors known to alter protein secondary structure, such as a solution's pH, presence of metal ions, and temperature, have been identified as protein aggregation triggers. Due to the rather harsh nature of all these stimuli, the search for milder and more biologically relevant aggregation inducers is ongoing.

The objective of this study was to utilize an enzyme as a natural tool for the triggering and control of amyloid formation. Based on the previous investigation (Section 8.1), which demonstrated that phosphorylation is an efficient aggregation inhibitor in amyloid forming coiled coil system, one of the unstructured monophosphorylated peptides was applied here as a substrate for protein phosphatase lambda. The structural switching was recorded by CD spectroscopy, the progress of enzymatic dephosphorylation was monitored via analytical HPLC, the kinetics of enzyme-induced aggregation was followed by ThT fluorescence, and finally the produced amyloid fibers were characterized by TEM.

It was demonstrated that enzymatic dephosphorylation effectively restores the β-sheet forming tendency of the model peptide and allows detailed monitoring of discrete conformational transitions of this system. Furthermore, the progress of enzymatic reaction, which indicates that the initial rapid phase is followed by a period of slower transformation,

helped a better understanding and rationalizing of the nucleation-dependent kinetics of self-assembly as well as of the morphology of produced fibers.

In conclusion, a phosphatase has been successfully applied as a natural aggregation trigger in a coiled coil peptide model. The strategy presented here does not require elevated temperatures, drastic pH, or concentration changes. On the contrary, it allows investigating of peptide aggregation in a dynamic way under mild, biologically relevant conditions.



GRAPHICAL ABSTRACT: A strategy has been demonstrated that utilizes a phosphatase as a natural tool for the triggering and control of amyloid formation in a coiled coil peptide model under conditions that closely approximate a physiological environment.

8.3 Towards understanding secondary structure transitions: phosphorylation and metal coordination in model peptides

The results presented in this section have been originally published as: M. Broncel, S. C. Wagner, K. Paul, C. P. R. Hackenberger, B. Koksch, Towards understanding secondary structure transitions: phosphorylation and metal coordination in model peptides, *Org. Biomol. Chem.* **2010**, *8*, 2575.

The original paper with supporting information is available at:

http://dx.doi.org/10.1039/c001458c

Individual contributions: M. Broncel synthesized investigated peptides and performed their complete CD and size exclusion chromatography (SEC) characterization. Atomic absorption spectroscopy analysis was performed in Prof. Gust laboratory (Institut für Pharmazie, Freie Universität Berlin), S. C. Wagner performed the TEM analysis, K. Paul carried out preliminary studies regarding manganese coordination during her bachelor internship in the group of Prof. B. Koksch under the supervision of the author.

Paper summary

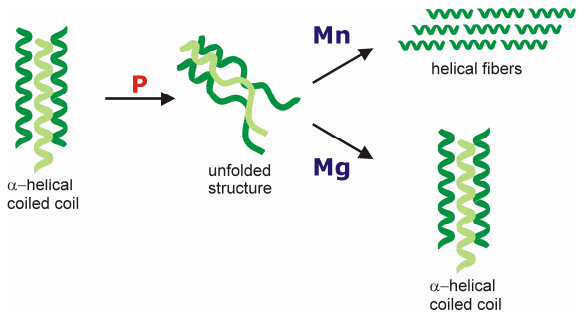
Secondary structure transitions are important modulators of signal transduction and protein aggregation. Phosphorylation is a well known post-translational modification capable of dramatic alteration of protein secondary structure. Additionally, phosphorylated residues can induce structural changes through metal binding. Despite the fact that data derived from the Protein Data Bank demonstrate that magnesium and manganese are metal ions most favored by phosphate, no previous reports describing structural consequences of the coordination of these metal ions to phosphoproteins are available.

The objective of this study was a careful CD spectroscopic analysis of structural consequences of phosphorylation and subsequent magnesium and manganese ions coordination in a coiled coil peptide model. Accordingly, a control, helical coiled coil peptide was first modified by double phosphorylation at the solvent exposed *f* position and subsequently titrated with the two tested metal ions.

The obtained results point to a significant structure switching ability of both tested factors, with phosphorylation being highly destabilizing and metal ions possessing structure

inducing properties. In the course of the experiment a remarkable switch cascade was obtained. It originates from a stable helical conformation of the control peptide, then continues through the phosphorylation-induced unfolded structure, and finally ends with a metal-stabilized α-helix in the case of magnesium or helical fibers in the case of manganese. Furthermore, each of the obtained new metal-stabilized structures could be transferred back to the unfolded form upon EDTA chelation.

In conclusion, this study demonstrated how small peptide models can aid in the evaluation and better understanding of interactions between metal ions and phosphopeptides. Owing to the intrinsic simplicity of the model, discrete conformational transitions can be studied in detail and further carried over to more complex systems, thus providing a great help in the understanding of protein structure and function.



GRAPHICAL ABSTRACT: *Structural consequences of phosphorylation and subsequent magnesium and manganese ion coordination were investigated in a coiled coil-based peptide model. It was demonstrated that these biologically relevant factors have significant molecular switching ability, with phosphorylation being highly destabilizing and metals possessing structure-inducing properties.*

8.4 Semisynthesis of site-specifically phosphorylated tau protein

Tau is a family of neuronal proteins whose primary function is to stabilize microtubules and maintain their function.[79] In Alzheimer's disease (AD) tau becomes abnormally phosphorylated, and in consequence dissociates from microtubules and aggregates to form insoluble lesions called paired helical filaments (PHFs) and neurofibrillary tangles (NFTs) (see section 3.2).[92, 94] Factors associated with the hyperphosphorylation of tau are not fully understood, but kinase/phosphatase imbalance seems to be at the very core of this phenomenon.[12] Accordingly, significant progress has been made in the identification of pathological phosphorylation sites on tau and several candidate kinases have been suggested (see section 3.2.1).[90] Current tau phosphorylation studies are mainly based on the application of phosphorylated tau-derived peptides,[171-173] pseudophosphorylation approach,[227, 228] and enzymatically generated phosphorylated tau.[96, 229] Unfortunately, these strategies suffer from major drawbacks. First, peptide-based studies may not necessarily reflect the behavior of the full length protein; second, pseudophosphorylation, i.e. substitution of serine and threonine residues with aspartic and glutamic acid, respectively, represents a poor approximation of the phosphorylated state; and finally *in vitro* enzymatic phosphorylations generate heterogeneous mixtures of tau phosphorylated at various sites and to a different extent. An alternative strategy could be envisioned, where the combination of synthetic and recombinant approaches, i.e. expressed protein ligation (EPL),[193] is applied for homogenous preparations of site-specifically phosphorylated tau, and thus greatly facilitate the investigation of phosphorylation patterns leading to neurofibrillary pathology.

EPL is a semisynthetic version of the immensely successful native chemical ligation (NCL)[180] in which two unprotected peptides, one containing a C-terminal thioester and another with N-terminal cysteine, are joined together via a native peptide bond (see section 5.2.1). Since its discovery in 1998 EPL has been successfully applied for the installation of a variety of site-specific modifications into proteins,[13, 195, 198] including post-translational modifications,[199-201] unnatural amino acids,[202, 203] and spectroscopic probes.[197, 204] Most importantly, however, EPL is actually the first choice technique for the preparation of site-specifically phosphorylated proteins.[184, 193, 205, 207]

Experiment design

As a 'proof of principle' experiment, a semisynthesis of the longest isoform of tau (441 amino acids) triply phosphorylated at so-called PHF-1 epitope (Ser396/Ser400/Ser404) was designed (Figure 8.1). It is important to emphasize that phosphorylation at this particular eptiope has been recognized exclusively in tau from AD brain and not in the healthy

tissue.[230] Furthermore it has also been shown to stimulate the rate of tau polymerization *in vitro*.[231]

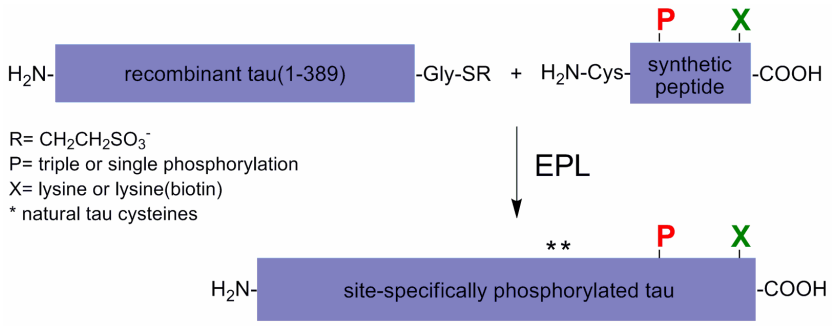


Figure 8.1 *General strategy undertaken in this study: site-specifically phosphorylated tau is generated via expressed protein ligation (EPL) from two fragments of recombinant and synthetic origin.*

The study involved solid phase peptide synthesis (SPPS) of 52-amino acid long C-terminal peptide bearing the triple phosphate modification in its N-terminal part and the preparation of the recombinant N-terminal tau (1-389) fragment comprising the α-thioester. Taking into account possible difficulties with both SPPS of triply phosphorylated peptide of such length as well as purification of semisynthetic tau from the crude ligation mix, an alternative design strategy was also envisioned. It involved first of all the preparation of singly phosphorylated tau at Ser404, an important, aggregation-related position (see section 3.2.1) that is a priming site for sequential phosphorylation at Ser400 and Ser396 by Gsk-3β, which subsequently yields AD-characteristic PHF-1 epitope;[232, 233] and second of all the installation of an affinity handle, i.e. biotin moiety, in the C-terminal part of the synthetic peptide which should enable efficient separation of the ligation product from the unreacted/hydrolyzed recombinant α-thioester (Figure 8.1).

Due to the fact that NCL requires a cysteine residue at the ligation site, and both natural tau cysteines are located in the central part of the protein (sequence positions 291 and 322, Figure 8.1), a conservative Ala(390)/Cys mutation was introduced at the N-terminus of the synthetic peptide to generate a non-hindered Gly-Cys ligation site (Figure 8.1). Importantly, the presence of glycine at the ligation junction has been shown to significantly accelerate native chemical ligation.[182]

In order to validate that the designed mutation does not affect tau's behavior, the full length wild type tau (WT-tau) and the Ala(390)/Cys mutant of full length tau (M-tau) were prepared as controls. The undertaken cloning strategy resulted in appending an additional glycine residue at the extreme N-terminus of these proteins. Due to its peripheral location and small size, this modification is not likely to interfere with tau structure and function.

Peptide synthesis

Taking into account that the designed peptide necessary for EPL constitutes 52 amino acids and as such practically reaches the limit of standard SPPS (~50 amino acids), preliminary studies were focused on extensive optimization of synthetic protocols, including testing of resins (Wang, 2-chlorotrityl, NovaPEG), couplings (single, double, 1-6 h), activation strategies (DIC/HOBt, DIC/HOAt, HBTU/DIEA, HATU/DIEA) as well as the application of pseudoproline derivatives. Employment of optimized conditions (Figure 8.2 A) first to the unmodified, i.e. non-phosphorylated and non-biotinylated test C-terminal peptide yielded a fully automated protocol which delivered the desired sequence in excellent yield (25 mg, 8.5 %) and purity (Figure 8.2 B).

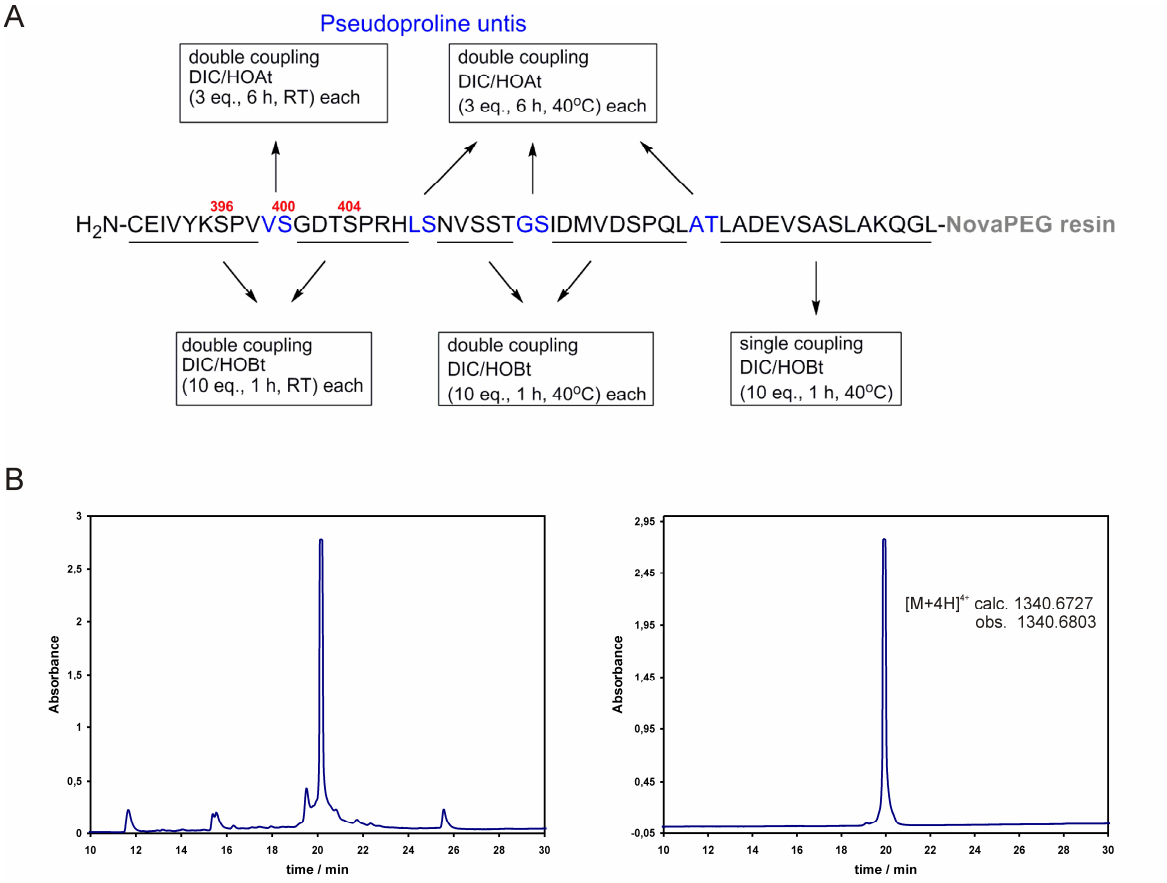


Figure 8.2 *SPPS of the unmodified 52 amino acid test C-terminal peptide. (A) Optimized SPPS conditions. Due to the risk of histidine racemization heating was aborted starting from this residue. Serine residues constituting the PHF-1 epitope are indicated by their numbers according to the 441 amino acid tau. (B) Analytical HPLC of crude (left) and purified (left) peptide.*

Unfortunately, despite the application of the same strategy and even manual synthesis with extended coupling times (double coupling, 4-6 h each) for the critical residues, i.e. Ser404 through Ser396 (Figure 8.2 A), the triply phosphorylated C-terminal peptide turned out to be inaccessible by means of building block phosphopeptide SPPS.

An attempt to produce this target via global phosphorylation strategy (see section 5.1.2) using dibenzyl diisopropylphosphoramidite delivered a complex mixture of peptides among which the desired product was not observed; instead a product that could be identified as H-phosphonate was found. Taking into account that the formation of this particular by-product is associated with the acidic conditions applied in this strategy, another class of phosphorylating reagents, phosphorochloridates (see section 5.1), which work under basic conditions was probed. Unfortunately, global phosphorylation using dibenzyl phosphorochloridate was unsuccessful as well and in this case only dehydroalanine by-products could be identified.

Considering the length of the target peptide as well as the fact that all three phosphorylation sites are located in the N-terminal part of its sequence, the peptide could be in principle obtained also via segment ligation utilizing NCL (Figure 8.3 A). The reasoning behind this strategy was simply to shorten (52→21 amino acids) the target phosphorylated peptide which, at least in theory, should facilitate its SPPS. The possible drawback on the other hand would be the necessity to prepare this fragment as an α-thioester. The new approach involved generation of the C-terminal triply phosphorylated 52 amino acid peptide from two separate fragments that would be subsequently joined together at the Asn-Val junction applying NCL (Figure 8.3 A). The choice of the ligation site was not random, as in the absence of internal cysteines, the commercially available penicillamine (β-mercaptovaline) can be utilized instead and upon desulfurization recreate original Val at the ligation site. As the application of the desulfurization protocol will address all thiol containing amino acids throughout the sequence, the N-terminal cysteine in fragment 2 would be introduced as commercially available thiazolidine derivative which upon treatment with methoxyamine HCl yields the native Cys residue (Figure 8.3 A). Initial results within this project were quite encouraging as SPPS of both fragments (fragment 2 in the non-phosphorylated and α-carboxylic acid form) revealed that they can be efficiently prepared applying just single couplings (10 eq., 1h) without any need for pseudoproline units (Figure 8.3 B). However, when an attempt was made to generate fragment 2 in the triply phosphorylated form, the synthesis failed despite the application of manual couplings (HATU/DIEA) with extended times (2-6 h). Thus, it can be concluded that even significant shortening of the target phosphopeptide length does not facilitate its synthesis and that the cause of the unsuccessful SPPS in case of phosphorylated fragment 2 is most probably the nature of the involved sequence that actually contains several β-branched amino acids (Val, Thr) in the direct vicinity of also quite bulky phosphoserines (Figure 8.3 A top). This particular residue arrangement can be a significant prerequisite for the phenomenon known as on-resin peptide aggregation resulting in the poor solvation within the peptide-polymer matrix and concomitantly SPPS failure.[174]

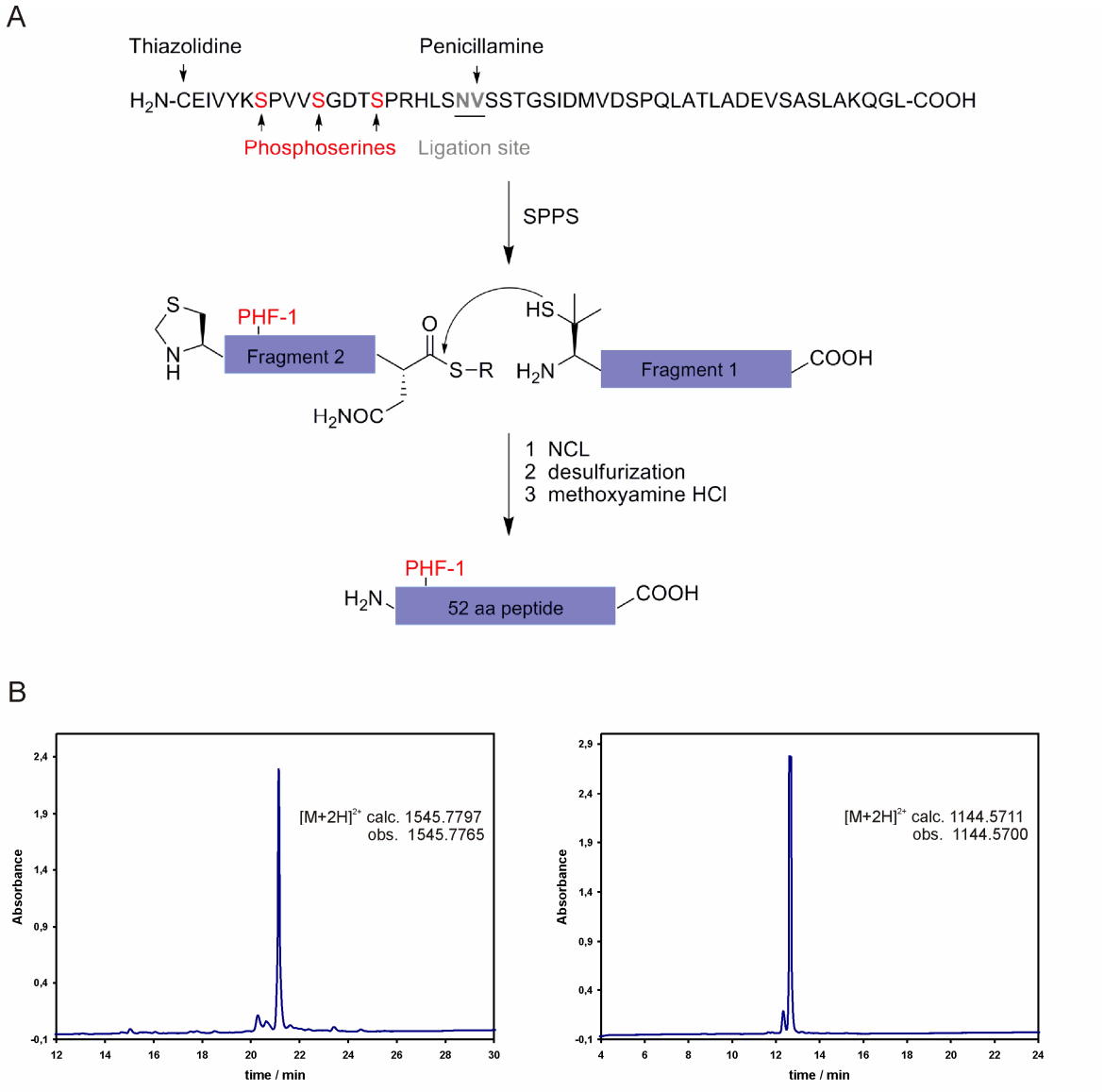


Figure 8.3 *Segment ligation approach for the preparation of the triply phosphorylated 52 amino acid C-terminal peptide. (A) General strategy and (B) analytical HPLC and HRMS (ESI) of crude peptides: fragment 1 (left) and fragment 2 in its non-phosphorylated form (right).*

In consequence, and according to the earlier envisioned scenario, the SPPS of a peptide bearing a single phosphate moiety at Ser404 was subsequently probed (Figure 8.4 A). In this case the synthesis was rather successful and the 52 amino acid long C-terminal fragment for EPL experiment was delivered in both non-phosphorylated (Figure 8.4 B left) as well as phosphorylated (Figure 8.4 B right) form in overall yields of 5 mg (2 %) and 6 mg (3 %), respectively. Taking into account the purification issue discussed above, both peptides were additionally equipped with the biotin affinity handle appended to a lysine side chain in order to facilitate post-ligation processing.

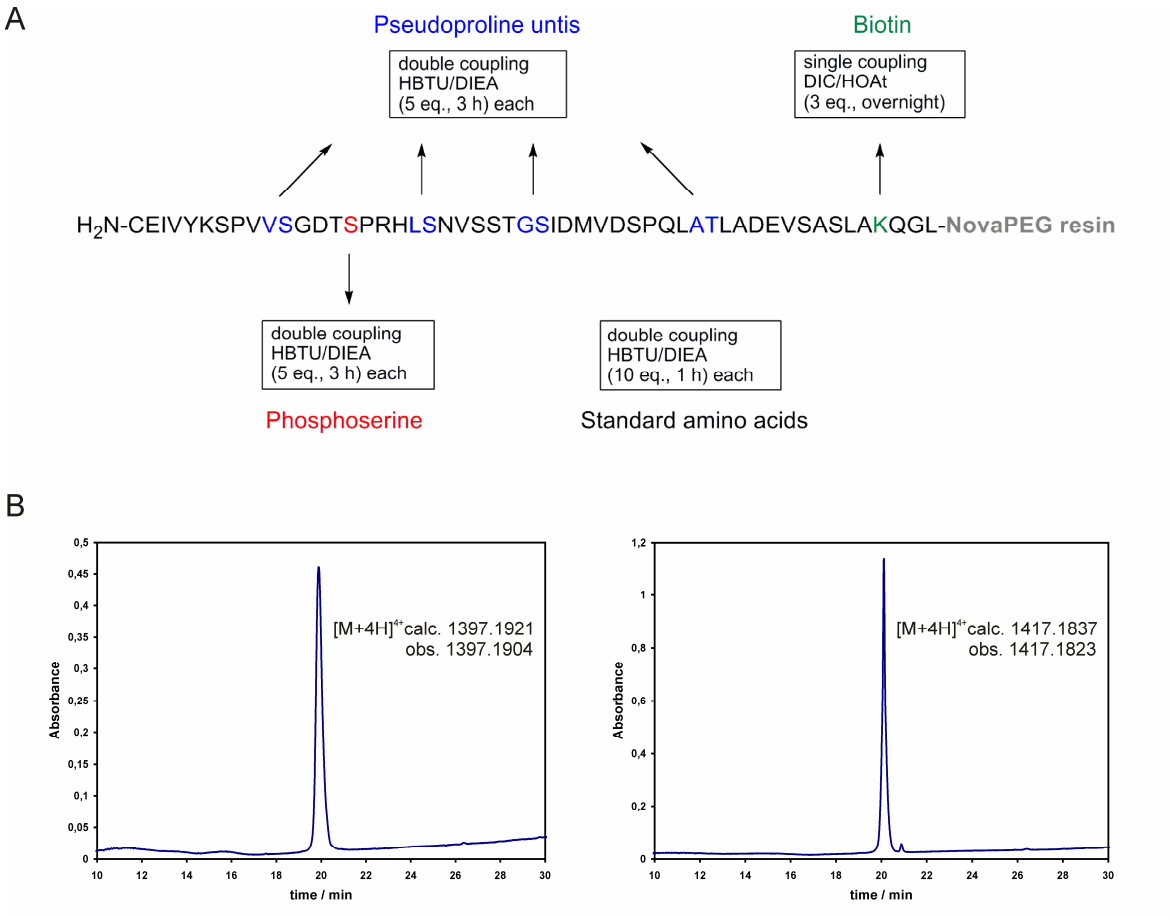


Figure 8.4 *SPPS of the biotinylated 52 amino acid C-terminal peptides. (A) Optimized SPPS conditions. In case of non-phosphorylated peptide Ser404 was introduced using standard coupling protocol. (B) Analytical HPLC and HRMS (ESI) of non-phosphorylated (left) and phosphorylated (left) biotinylated peptide.*

Preparation of the recombinant thioester and EPL

In order to deliver the recombinant α-thioester, DNA encoding tau (1-389) fragment was subcloned into pTXB1 vector which upon expression generates tau(1-389)/*Mxe* GyrA-intein/chitin binding domain (CBD) fusion protein (Figure 8.5). This construct was subsequently loaded on chitin beads, where thiol-induced cleavage activity of the intein afforded the reactive N-terminal tau (1-389) α-thioester. To avoid decomposition, the thioester was not isolated; instead the *in situ* EPL with an excess of synthetic peptide was performed (Figure 8.5).

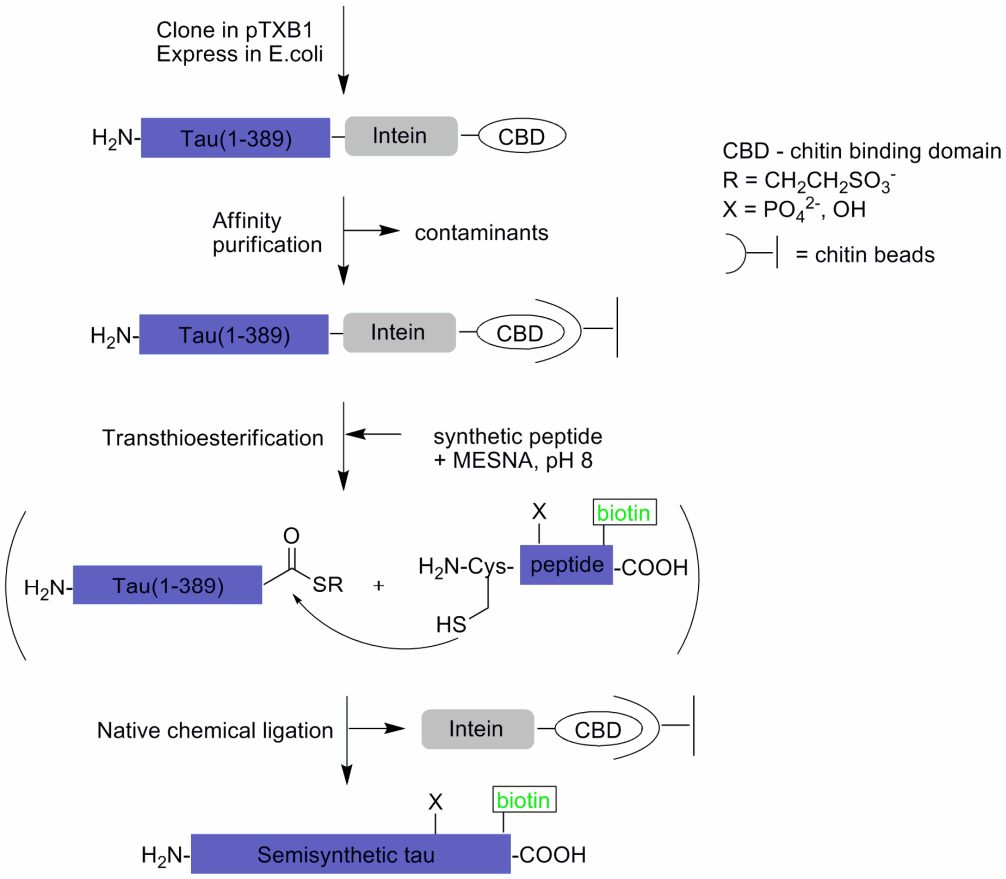


Figure 8.5 *The principle of expressed protein ligation. N-terminal tau fragment is expressed as N-terminal fusion to intein/CBD. Following affinity purification this construct is cleaved with thiols to yield recombinant α-thioester, which subsequently reacts in situ with a synthetic fragment to furnish semisynthetic tau protein.*

As estimated from SDS-PAGE (Figure 8.6 A) the ligation reaction proceeded quite efficiently, generating the desired product on average in 70-80% yield. In order to remove two main contaminants present in the ligation mixture, i.e. the excess of peptide and the unreacted/hydrolyzed thioester (Figure 8.6 A lane 4), a double purification strategy was applied. First, spin filtration utilizing 30 kDa cutoff yielded the peptide-free sample and subsequently affinity purification on the monomeric avidin resin removed the unreacted thioester furnishing pure semisynthetic site-specifically phosphorylated and unmodified proteins, S-tau(P) and S-tau, respectively (Figure 8.6 B). Although ligations were not fully optimized, good average yields of 200 µg ligation products/L of bacterial culture were obtained.

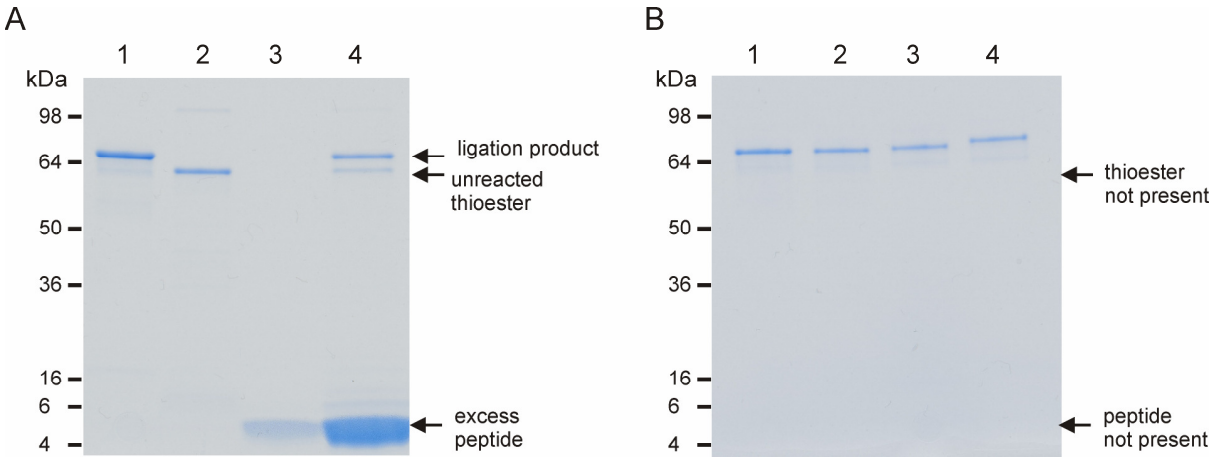


Figure 8.6 *SDS-PAGE analysis (Coomassie stained proteins) of (A) representative ligation reaction: lane 1: WT-tau, lane 2: separated thioester, lane 3: synthetic peptide, lane 4: crude ligation of S-tau; and (B) purified proteins: lane 1: WT-tau, lane 2: M-tau, lane 3: S-tau, lane 4: S-tau(P). Molecular weight markers are shown to the left. Note, tau runs retarded on SDS/PAGE gels, calc. MW 46 kDa, obs. MW 67 kDa.*[234]

Protein analysis

Semisynthetic proteins and both negative controls were analyzed in terms of immunochemical recognition and functional activity. Initial Western blot experiments with the monoclonal tau 46 antibody raised against human recombinant tau delivered an unexpected result, as both semisynthetic proteins were not recognized (Figure 8.7 A). Taking into account that the antibody targets the C-terminal part of tau protein, the presence of biotin moiety in this particular region apparently hindered the recognition event. Therefore, another tau-specific antibody which targets the central region of tau was subsequently applied. As shown in Figure 8.7 B in this case all tested tau variants, including EPL derived proteins, were successfully recognized. Moreover, antibody sensitive towards site-specific tau phosphorylation at Ser404 selectively recognized S-tau(P) (Figure 8.7 C).

Functional studies were implemented via straightforward *in vitro* assay which measures the ability of tau to promote tubulin polymerization. It was demonstrated that all tested proteins induced vigorous tubulin self-assembly, as illustrated by the rapid sigmoidal increase in optical density (OD) at 350 nm (Figure 8.7 E).

These initial results demonstrate that the introduced A(390)/C mutation does not change tau's antigenic properties and activity (comparison of WT-tau and M-tau), and thus validate the implemented design strategy. In addition, analysis of S-tau unambiguously proves that the protein's behavior is not altered by the applied semisynthesis conditions, and therefore confirms the utility of EPL methodology for the preparation of tau. Finally, investigation of S-tau(P) illustrated that single phosphorylation at Ser404 apparently does not induce any significant changes in tau function. Phosphorylation at this particular position, however, is

expected to have an effect on the aggregation potential of tau. Intensive examination of such scenario is ongoing.

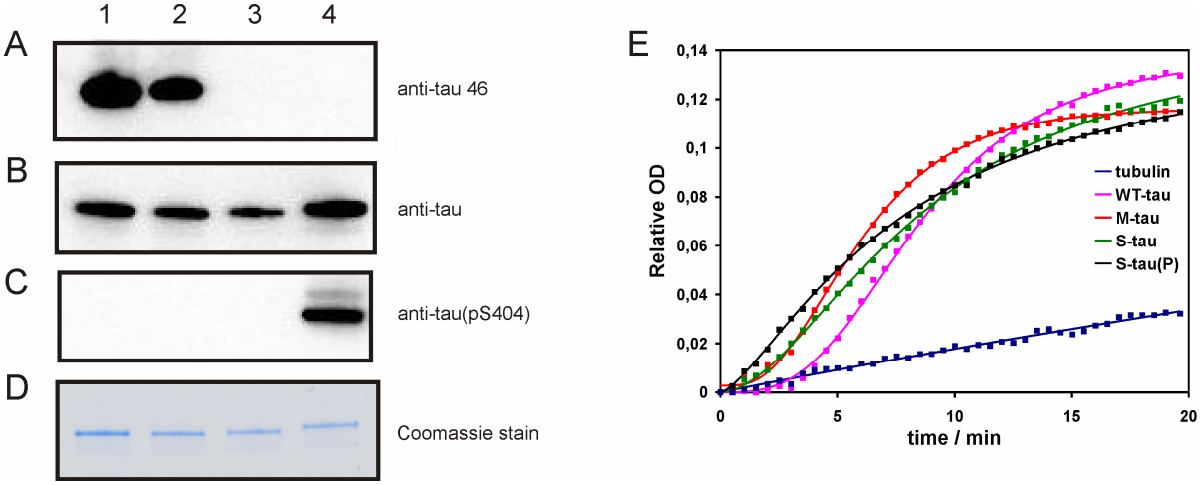


Figure 8.7 *Antigenic properties and functional activity of analyzed proteins. Western blots with (A) tau 46, (B) tau, and (C) phosphorylation-dependent antibodies as well as Coomassie stain (D) used as a control for protein loading (0.5 µg/lane). Lane1: WT-tau, lane 2: M-tau, lane 3: S-tau, lane 4: S-tau(P). (E) Tubulin polymerization assay in the absence and presence of prepared tau proteins. Normalized* OD_{350nm} *was fitted to a sigmoidal growth model.*

In conclusion, a semisynthetic strategy was successfully applied for the preparation of a homogenous, site-specifically phosphorylated tau protein. Implementation of this methodology allows unambiguous verification of individual phosphorylation sites on tau, and consequently greatly facilitates investigations of complex phosphorylation patterns on this protein.

Experimental procedures

Materials

Fmoc-L-amino acids, HBTU and HOBt were purchased from Fa. Gerhardt (Wolfhagen, Germany). HOAt was from Iris Biotech. Fmoc-Ser(PO(OBzl)OH)-OH, Fmoc-Lys(biotin)-OH, all pseudoproline dipeptides and resins were from Novabiochem. All peptide synthesis chemicals were purchased from Acros. Expression vectors and chitin beads were from New England Biolabs. Monomeric avidin agarose and biotin were from Pierce. Ni-NTA agarose was from Qiagen. HiTrap SP Sepharose FF column was supplied by Amersham Biosciences. Amicon Ultra centrifugal filter devices were from Millipore. BC Assay for protein quantitation was from Uptima. SDS-PAGE 4-20 % pre-cast gels were obtained from Bio-Rad. Antibodies were purchased from Santa Cruz Biotechnology. Tubulin was obtained from Cytoskeleton. Microtiter plates were from Corning.

Synthesis and purification of biotinylated peptides for EPL

Peptides were synthesized on Activo-P11 LS automated peptide synthesizer (Activotec) according to Fmoc-based protocols using HBTU/DIEA or DIC/HOAt and pre-loaded H-Leu-HMPB-NovaPEG resin. Double coupling was applied for standard amino acids (10 eq., 1h each), pseudoproline dipeptides (5 eq., 3h each) and Fmoc-Ser(PO(OBzl)OH)-OH (5 eq., 3h each). Fmoc-Lys(biotin)-OH (3eq.) was coupled manually with DIC/HOAt overnight. Fmoc deprotection was accomplished with 25% piperidine in DMF. Peptides were cleaved from the resin by treatment with 2 mL TFA/TIS/DTT/thioanisole (95/2/2/1) for 3 h followed by precipitation with cool diethyl ether. Purification was carried out by preparative reversed phase HPLC on a Knauer Smartline system (Knauer GmbH) equipped with a Luna C8 (10 µm, 250×21.20 mm) column (Phenomenex) running with water/0.1% TFA and ACN/0.1% TFA gradient at 20 mL min^{-1}. Purified peptides were characterized by analytical HPLC and HRMS (ESI). The analytical HPLC was carried out with a VWR-Hitachi Elite LaChrome system (VWR International) equipped with a Luna C8 (5µ, 250×4.6 mm) column (Phenomenex). The flow rate was 1 mL min^{-1}. Peptide mass to charge ratios were measured on an Agilent 6210 ESI-TOF (Agilent Technologies).

Cloning, expression and purification of recombinant proteins

Restriction sites (*Nde*I with downstream TEV protease recognition sequence and *Xho*I) were introduced into the tau gene (encoding 441 aa) using standard PCR protocols. Following cleavage with *Nde*I and *Xho*I the resulting DNA fragment was subcloned downstream T7 RNA polymerase promoter into *Nde*I/*Xho*I cut expression vector pET 15b. This cloning strategy placed the TEV cleavage site and tau encoding sequence downstream of a hexahistidine tag. DNA encoding the mutant tau A(389)/C was prepared from the wild type tau DNA by standard site-directed mutagenesis. Following DNA sequencing constructed plasmids were transformed into *E. coli* BL 21 (DE3) cells for expression. Bacterial cultures were grown at 37 ºC in LB medium containing 100 µM/mL ampicilin. When the OD_{600nm} reached 0.6 protein expression was induced with IPTG (1 mM). After 3 h incubation at 37 ºC cells were collected by centrifugation and the cell pellet was resuspended in 20 mL of the extraction buffer (50 mM sodium phosphate pH 7.4, 300 mM NaCl, 0.25 mM DTT, 1% Triton X100, 6 mM imidazole) complemented with a protease inhibitor cocktail. Following sonication the cell lysate was heated (75 ºC, 15 min.), clarified by centrifugation, and loaded onto Ni-NTA agarose equilibrated in 50 mM sodium phosphate pH 8, 6 mM imidazole, 300 mM NaCl. The protein of interest was eluted with 150 mM imidazole in 50 mM sodium phosphate pH 7.6, 300 mM NaCl. Following dialysis (50 mM sodium phosphate pH 6.8, 1 mM EDTA, 2 mM DTT) the His-tag was removed by TEV protease (1:100 OD_{280nm} ratio of enzyme to protein) during an overnight reaction at 8 ºC. Cleaved protein was directly loaded

onto 5 mL HiTrap SP Sepharose FF column and subjected to cation exchange chromatography. The column was equilibrated in 50 mM sodium phosphate pH 6.8, 1 mM EDTA, 2 mM DTT and the protein of interest was eluted with the linear gradient of salt (0-100% 50 mM sodium phosphate pH 6.8, 0.5 M NaCl, 1 mM EDTA, 2 mM DTT in 20 min.). One milliliter fractions were collected and analyzed on SDS-PAGE. Appropriate fractions were pooled, concentrated (Amicon Ultra-15 10 kDa), and following buffer exchange (50 mM Ammonium bicarbonate) proteins were aliquoted, lyophilized, and stored at -20 ºC until use. Protein concentration was determined using BC Assay. On average 15-20 mg protein/L bacterial culture was obtained.

Preparation of the recombinant α-thioester and EPL

Primers containing *Nde*I (upstream) and *Xho*I (downstream) sites were used to PCR amplify tau(1-389) DNA fragment for in frame insertion upstream *Mxe*-Gyr A intein/CBD sequence in the pTXB1 expression vector. As this cloning strategy results in the presence of five additional amino acids at the tau(1-389)-intein boundary, a mutagenesis protocol including the PCR amplification of the whole plasmid with the exclusion of the mutation site, followed by standard DNA phosphorylation and re-ligation was applied. Following DNA sequencing plasmid was expressed in *E.coli* ER2566 cells grown at 37 ºC in LB medium containing 100 µM/mL ampicilin. When OD_{600nm} reached 0.6 protein expression was induced at 18 ºC with IPTG (0.4 mM). After overnight incubation at 18 ºC, cells were collected by centrifugation and resuspended in 20 mM Tris HCl pH 7.2, 500 mM NaCl, 1mM EDTA, 0.1% Triton X100, 0.1 mM TCEP complemented with a protease inhibitor cocktail. Cells were disrupted by sonication and after centrifugation the clarified supernatant was slowly loaded onto chitin beads (2-3 mL) washed with equilibration buffer (20 mM Tris HCl pH 8.0, 500 mM NaCl, 1mM EDTA, 0.1% Triton X100, 0.1 mM TCEP). Following extensive wash with equilibration buffer, the beads were quickly flushed with cleavage buffer (20 mM Tris HCl pH 8.0, 500 mM NaCl, 1mM EDTA, 0.1 mM TCEP, 50 mM MESNA) and finally the corresponding peptide (1 mM) was applied in 0.5 mL of cleavage buffer complemented, in the case of the phosphopeptide, with Na_3VO_4 (0.1 mM) to prevent any phosphatase activity that might have co-purified with the recombinant construct. After 24 h standing at RT an additional portion of thiol (4-mercaptophenyl acetic acid, 50mM) was added and following concomitant 24 h the column was eluted with equilibration buffer (-Triton X100). The eluate was then concentrated and buffer exchanged to avidin buffer (20 mM Tris, pH 7.2, 150 mM NaCl, 1mM DTT) using centrifugal filter units (Amicon Ultra-4, 30 kDa cutoff). This procedure yields simultaneous removal of the unligated peptide (MW~6 kDa). In order to remove unreacted thioester, monomeric avidin agarose was applied. Pure ligation product was eluted with 2 mM biotin in avidin buffer. Following buffer exchange (50 mM Ammonium

Bicarbonate), semisynthetic proteins were aliquoted, lyophilized and stored at -20 ºC until use. Protein concentration was determined as described above.

Western blot

Proteins (0.5 µg) were separated using SDS-PAGE (4-20% gradient gels) and wet blotted onto a nitrocellulose membrane (max. voltage, 200 mA, 2 h). After blocking with 2 % milk powder the blot was incubated overnight at 8 ºC with anti-tau rabbit polyclonal antibody (1:200) and subsequently with a goat anti-rabbit HRP-conjugated secondary antibody (1:1000) for 1 h at RT. Following chemiluminescent detection in Lumi-Imager F1 (Boehringer Mannheim) the blot was stripped (50 mM Tris HCl, pH 6.8, 2 % SDS, 0.1 M 2-Mercaptoethanol, 15 min at 50 ºC) and re-probed with p-Tau (Ser 404) rabbit polyclonal antibody (1: 500) for 1 h at RT.

Tubulin polymerization assay

Tubulin polymerization was monitored in the presence or absence of tau proteins over 20 min at 37 ºC by the change in optical density at 350 nm using a 96 half area well plate and a Tecan spectrofluorimeter (Crailsheim, Germany). Tubulin (3 mg/mL) and various forms of tau (2.5 µM) were mixed in 60 µL of ice cold assembly buffer (80 mM PIPES pH 7.0, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, and 1 mM GTP) and the reaction was initiated by the increase in temperature to 37 ºC. Optical density at 350 nm was normalized (OD at the start point of the kinetic trace was set to zero) and fitted to a sigmoidal growth model using SigmaPlot 10.0 software.

9 Summary and outlook

During the course of this work peptide and protein aggregation was investigated using synthetic phosphopeptides and phosphoproteins. Two main strategies were applied. The first involved the application of phosphopeptides as reliable coiled coil-based models for a better understanding of the influence of phosphorylation on amyloid formation, for studying phosphatase induced self-assembly and detailed characterization of structural transformations and aggregation kinetics, as well as for investigating phosphopeptide-metal ion interactions. The second strategy involved the application of phosphopeptides as reagents for protein semisynthesis via expressed protein ligation in order to produce site-specifically modified tau – a protein that in a hyperphosphorylated form constitutes one of the major brain lesions of Alzheimer's disease.

Model studies

Initially, the impact of phosphorylation on the amyloid formation process was investigated utilizing a coiled coil model peptide which transformed from α-helix to β-sheet and further to amyloid in a time-dependent manner. Phosphoserine residues were introduced at the solvent exposed *f* position of the coiled coil in a site-specific and multiple fashion to mimic phosphorylation patterns observed *in vivo*. Structural analysis of phosphorylated coiled coil-based peptides revealed that regardless of the quantity and position of modification, the formation of β-sheet structure was completely abolished. Furthermore, all phosphorylated peptides completely lost their amyloid forming potential as demonstrated by time-dependent measurements as well as seeding experiments. Finally, TEM analysis confirmed the lack of any type of aggregates in the analyzed samples. It has been concluded, therefore, that phosphorylation is a very efficient aggregation breaker in this particular coiled coil model. This behavior could be predominantly attributed to the charged nature of the phosphate group, specifically to intramolecular Coulomb repulsions between glutamate (at position *b*) and phosphate (at position *f*) that destabilize the initial helical structure of the model, as well as to the excess of uncompensated charge that renders the formation and further lamination of β-sheets energetically unfavorable.

Based on the aforementioned results indicating that even a single phosphorylation was sufficient to completely abolish amyloid formation in a coiled coil model, a strategy that utilized a phosphatase as a natural aggregation trigger was subsequently pursued. It was demonstrated that enzymatic dephosphorylation effectively restored β-sheet forming tendency of the model peptide and allowed detailed monitoring of discrete conformational transitions of this system. Furthermore, the progress of enzymatic reaction, which indicated that the initial rapid phase is followed by a period of slower transformation, led to a better

understanding and rationalization of the nucleation-dependent kinetics of self-assembly, as well as of the morphology of the fibers generated. Importantly, the application of enzyme as peptide aggregation trigger obviated the employment of harsh environmental changes, and allowed investigation of amyloid formation in a dynamic way under mild, biologically relevant conditions.

Taking into account that conformational transitions constitute a central prerequisite for amyloid formation and that both phosphorylation as well as metal ions have been implicated in this phenomenon, the helical coiled coil model was applied in the following investigation to elucidate structural consequences of phosphorylation and subsequent magnesium and manganese ions coordination. The results obtained demonstrated a significant structure switching ability of both tested factors. Double phosphorylation at the *f* position of the coiled coil turned out to be highly destabilizing whereas subsequent titration with metal ions demonstrated their structure-inducing properties. In the course of the experiment a remarkable switch cascade was obtained. It originated from a stable helical conformation of the control peptide, continued through the phosphorylation-induced unfolded structure, and finally ended with a metal-stabilized α-helix in case of magnesium or α-helical fibers in case of manganese. Together these results highlight the role of metal ions in inducing conformational transitions yielding ordered structures in phosphopeptides/phosphoproteins, and following this line of reasoning possibly also in phosphorylation dependent aberrant self-assembly.

Future investigation using the coiled coil model could focus on:

- Further exploitation of enzymatic approaches for studying amyloid formation. Based on the finding that phosphorylation is very efficient in preventing aggregation in the studied model and the successful application of a phosphatase as the aggregation trigger, a complementary strategy that utilizes a kinase as an aggregation inhibitor can be envisioned. This particular approach would be divided into two general parts. The first would include the design and synthesis of amyloid forming coiled coil peptides comprising a kinase recognition sequence as well as their detailed characterization in terms of structural behavior, amyloid forming properties, and fibril morphology. The second part would be focused on experiments with the kinase. The progress and efficiency of enzymatic phosphorylation would be monitored by analytical HPLC measurements, structural transitions would be recorded by a real-time CD spectroscopy, aggregation kinetics would be followed by ThT-induced fluorescence, and finally time-dependent TEM would provide evidence of any changes in fiber morphology. Implementation of this strategy would be highly dependent on the kinase ability to effectively phosphorylate a dynamic, aggregating

system, and if successful could be a valuable addition to the commonly pursued peptide- and small molecule-based approaches for the inhibition of *in vitro* peptide self-assembly.

- Further investigation of phosphopeptide-metal ions interactions. Specifically, the metal ions that are known to be involved in the process of amyloid formation, e.g. copper (Cu^{2+}) and zinc (Zn^{2+}) could be applied. It would be quite interesting to elucidate whether the interaction of phosphopeptides with these particular ions could generate amyloid fibers in the studied coiled coil model. If so, it could be speculated that the reported high levels of copper and zinc ions in the brain might constitute yet another factor contributing to tau pathology.

Protein semisynthesis

Upon completion of model studies, attention was shifted towards studying phosphorylation in natural aggregating systems. Accordingly, a combination of synthetic and recombinant approaches, i.e. expressed protein ligation was utilized to produce site-specifically phosphorylated tau protein in its longest isoform. Even though the initially designed protein bearing triple phosphorylation at the AD-characteristic PHF-1 epitope (Ser396/Ser400/Ser404) was not delivered due to the extremely challenging synthesis of the necessary phosphoryated 52 amino acid long C-terminal peptide, a monophosphorylated (Ser404) semisynthetic tau was successfully produced. Owing to application of *in situ* EPL, which minimized recombinant thioester decomposition, as well as the double purification strategy applied, which efficiently removed the ligation contaminants, the site-specifically phosphorylated tau was obtained in good overall yield and high purity. Immunochemical analysis of the product revealed selective recognition by both tau- and phosphorylation-specific antibodies. Furthermore functional studies demonstrated that site-specific phosphorylation at Ser404 did not affect the protein's activity of inducing tubuline polymerization. The implemented strategy thus paves the way towards unambiguous verification of individual phosphorylation sites on tau, and consequently should greatly facilitate investigation of complex phosphorylation patterns on this protein.

Future work within this project could involve:

- Analysis of aggregation behavior of produced monophosphorylated tau utilizing ThS-induced fluorescence followed by TEM characterization of generated fibers. Based on the investigation of Abraha et al.[231] who observed that mimicking the phosphorylation at Ser396 and Ser404 (pseudophosphorylation approach) stimulates

the rate of tau self-assembly, it is probable that even single phosphorylation at Ser404 should increase the the protein's aggregation propensity.

- Investigation of additional aberrant phosphorylation sites on tau and detailed evaluation of their effect on tau structure (NMR), function (tubulin polymerization), and aggregation (ThS-induced fluorescence). These studies would be based on the application of tandem EPL that enables the introduction of site-specific phosphorylation(s) in the middle region of the protein, which harbors such AD-related sites as Ser202/Thr205 (AT8 epitope), Thr212/Ser214 (AT100 epitope) and Thr231/Ser235 (AT180 epitope). Implementation of this strategy would involve the insertion of the corresponding site-specifically phosphorylated peptide equipped with both an N-terminal cysteine and an α-thioester moiety between two recombinant fragments containing complementary reactive groups at their respective termini. The crucial aspect of such sequential ligation would be the reversible protection (TEV protease recognition sequence) of the N-terminal cysteine in the central peptide fragment which prevents its uncontrolled intra- and inter-molecular self-ligation. The N-terminal recombinant fragment containing an α-thioester would be generated via standard intein-based expression and subsequent thiolysis. The C-terminal protein part would be N-terminally fused to the TEV protease recognition sequence, which upon cleavage generates the N-terminal cysteine. In order to facilitate purification and minimize handling losses, sequential ligation steps would be performed on the solid support. To execute this strategy the C-terminal recombinant segment should be additionally equipped with an affinity handle appended to its C-terminus, e.g. His-tag. The first step of the procedure would involve the loading of the C-terminal tau fragment on the affinity column followed by the TEV protease-induced liberation of the N-terminal cysteine. Next the synthetic peptide α-thioester would be added to initiate the first EPL. The formed ligation intermediate would be again treated with TEV protease to reveal N-terminal cysteine for the second EPL, this time with the N-terminal tau α-thioester segment. Upon completion of the ligation procedure the phosphorylated semisynthetic protein could be eluted form the Ni-NTA beads either with an imidazole gradient (final product with C-terminal His-tag) or alternatively it could be cleaved from the solid support, provided that a protease recognition site would be introduced on the genetic level upstream to the tag (His-tag free ligation product).

- Site-specific introduction of O-GlcNAc residues at the previously identified modification positions (Thr123, Ser208, Ser238, and Ser400) and evaluation of their

effect on tau structure (NMR experiments), function (tubulin polymerization assay), and self-assembly (ThS-induced fluorescence). Following these initial experiments the project could move towards the investigation of the reported dynamic interplay between phosphorylation and O-GlcNAcylation. Here the produced site-specifically O-GlcNAcylated tau proteins would be subjected to enzymatic phosphorylation by AD-related tau kinases, for instance Gsk-3β, cdk5, or PKA. Thereby obtained phosphorylation patterns could be subsequently compared with those of the non-glycosylated control sample, thus providing a direct evaluation of the effect of each individual O-GlcNAcylation site on tau phosphorylation.

- Production of segmental isotope labeled tau in order to enable and simplify heteronuclear NMR-based structural investigations. This strategy could address tau protein generated via both semisynthetic and recombinant approaches. In case of the former, EPL would be used to couple a synthetic peptide bearing the respective modification (phosphorylation, glycosylation) with a stable isotope (^{15}N, ^{13}C) labeled recombinant thioester generated by intein-based expression. In case of the latter, recombinant isotope labeled tau fragment(s) could be coupled with non-labeled recombinant counterpart(s) via either EPL or protein trans-splicing (PTS). The advantage of PTS is that it additionally enables performing the ligation *in vivo* thereby facilitating protein handling and reducing the number of purification steps. The application of segmental isotopic labeling strategy should significantly decrease the complexity of NMR spectra and allow detailed examination of discrete polypeptide regions, both of which are of paramount importance in NMR investigations of large and intrinsically disordered proteins such as tau.

In conclusion, the application of phosphopeptides as reliable models and reagents that facilitate production of site-specifically modified tau protein overcomes severe obstacles in current AD-related research, namely the complexity and challenging physicochemical properties of natural aggregating systems as well as the difficult access to homogenous samples of phosphorylated tau. Consequently, the work presented in this thesis as well as future results that might be obtained from the proposed follow up experiments should without a doubt create better characterization and understanding of the specific role that phosphorylation plays in aberrant peptide and protein aggregation and the pathology of Alzheimer's disease.

10 Literature

1. M. Goedert, M. G. Spillantini, *Science* **2006**, *314*, 777.

2. K. A. Jellinger, *J. Neural. Transm.* **2006**, *113*, 1603.

3. C. A. Wilson, R. W. Doms, V. M. Lee, *J. Neuropathol. Exp. Neurol.* **1999**, *58*, 787.

4. I. Grundke-Iqbal, K. Iqbal, Y. C. Tung, M. Quinlan, H. M. Wisniewski, L. I. Binder, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 4913.

5. J. A. Hardy, G. A. Higgins, *Science* **1992**, *256*, 184.

6. R. Jakob-Roetne, H. Jacobsen, *Angew. Chem. Int. Ed.* **2009**, *48*, 3030.

7. M. Citron, *Nat. Rev. Neurosci.* **2004**, *5*, 677.

8. M. G. Spillantini, T. D. Bird, B. Ghetti, *Brain Pathol.* **1998**, *8*, 387.

9. D. Yancopoulou, M. G. Spillantini, *Neuromolecular Medicine* **2003**, *4*, 37.

10. P. V. Arriagada, J. H. Growdon, E. T. Hedley-Whyte, B. T. Hyman, *Neurology* **1992**, *42*, 631.

11. H. Braak, E. Braak, *Acta Neuropathol.* **1991**, *82*, 239.

12. C. X. Gong, F. Liu, I. Grundke-Iqbal, K. Iqbal, *J. Biomed. Biotechnol.* **2006**, *2006*, 31825.

13. C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 10030.

14. K. A. Dill, H. S. Chan, *Nat. Struct. Biol.* **1997**, *4*, 10.

15. A. R. Dinner, A. Sali, L. J. Smith, C. M. Dobson, M. Karplus, *Trends Biochem. Sci.* **2000**, *25*, 331.

16. C. M. Dobson, *Semin. Cell. Dev. Biol.* **2004**, *15*, 3.

17. C. M. Dobson, A. Sali, M. Karplus, *Angew. Chem. Int. Ed.* **1998**, *37*, 868.

18. C. M. Dobson, *Nature* **2003**, *426*, 884.

19. R. Glozman, T. Okiyoneda, C. M. Mulvihill, J. M. Rini, H. Barriere, G. L. Lukacs, *J Cell. Biol.* **2009**, *184*, 847.

20. C. Hammond, A. Helenius, *Curr. Opin. Cell. Biol.* **1995**, *7*, 523.

21. F. Chiti, P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, C. M. Dobson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3590.

22. M. Lopez De La Paz, K. Goldie, J. Zurdo, E. Lacroix, C. M. Dobson, A. Hoenger, L. Serrano, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16052.

23. J. Zurdo, J. I. Guijarro, J. L. Jimenez, H. R. Saibil, C. M. Dobson, *J. Mol. Biol.* **2001**, *311*, 325.

24. W. S. Gosal, I. J. Morten, E. W. Hewitt, D. A. Smith, N. H. Thomson, S. E. Radford, *J. Mol. Biol.* **2005**, *351*, 850.

25. L. C. Serpell, M. Sunde, M. D. Benson, G. A. Tennent, M. B. Pepys, P. E. Fraser, *J. Mol. Biol.* **2000**, *300*, 1033.

26. M. Stefani, C. M. Dobson, *J. Mol. Med.* **2003**, *81*, 678.

27. J. L. Jimenez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, H. R. Saibil, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9196.

28. J. L. Jimenez, J. L. Guijarro, E. Orlova, J. Zurdo, C. M. Dobson, M. Sunde, H. R. Saibil, *Embo J.* **1999**, *18*, 815.

29. S. B. Malinchik, H. Inouye, K. E. Szumowski, D. A. Kirschner, *Biophys. J.* **1998**, *74*, 537.

30. E. D. Eanes, G. G. Glenner, *J. Histochem. Cytochem.* **1968**, *16*, 673.

31. R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe, D. Eisenberg, *Nature* **2005**, *435*, 773.

32. T. Luhrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Doeli, D. Schubert, R. Riek, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17342.

33. J. J. Balbach, A. T. Petkova, N. A. Oyler, O. N. Antzutkin, D. J. Gordon, S. C. Meredith, R. Tycko, *Biophys. J.* **2002**, *83*, 1205.

34. D. A. Kirschner, C. Abraham, D. J. Selkoe, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 503.

35. A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, R. Tycko, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16742.

36. O. S. Makin, L. C. Serpell, *FEBS J.* **2005**, *272*, 5950.

37. J. C. Rochet, P. T. Lansbury, Jr., *Curr. Opin. Struct. Biol.* **2000**, *10*, 60.

38. R. Wetzel, *Acc. Chem. Res.* **2006**, *39*, 671.

39. J. T. Jarrett, P. T. Lansbury, *Cell* **1993**, *73*, 1055.

40. B. P. Tseng, W. P. Esler, C. B. Clish, E. R. Stimson, J. R. Ghilardi, H. V. Vinters, P. W. Mantyh, J. P. Lee, J. E. Maggio, *Biochemistry* **1999**, *38*, 10424.

41. J. D. Harper, S. S. Wong, C. M. Lieber, P. T. Lansbury, *Chem. Biol.* **1997**, *4*, 119.

42. R. Kodali, R. Wetzel, *Curr. Opin. Struct. Biol.* **2007**, *17*, 48.

43. M. P. Lambert, A. K. Barlow, B. A. Chromy, C. Edwards, R. Freed, M. Liosatos, T. E. Morgan, I. Rozovsky, B. Trommer, K. L. Viola, P. Wals, C. Zhang, C. E. Finch, G. A. Krafft, W. L. Klein, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6448.

44. R. Kayed, E. Head, J. L. Thompson, T. M. McIntire, S. C. Milton, C. W. Cotman, C. G. Glabe, *Science* **2003**, *300*, 486.

45. S. Lesne, M. T. Koh, L. Kotilinek, R. Kayed, C. G. Glabe, A. Yang, M. Gallagher, K. H. Ashe, *Nature* **2006**, *440*, 352.

46. B. O'Nuallain, S. Shivaprasad, I. Kheterpal, R. Wetzel, *Biochemistry* **2005**, *44*, 12709.

47. N. Carulla, G. L. Caddy, D. R. Hall, J. Zurdo, M. Gairi, M. Feliz, E. Giralt, C. V. Robinson, C. M. Dobson, *Nature* **2005**, *436*, 554.

48. T. Suzuki, M. Oishi, D. R. Marshak, A. J. Czernik, A. C. Nairn, P. Greengard, *Embo J* **1994**, *13*, 1114.

49. T. C. Gamblin, F. Chen, A. Zambrano, A. Abraha, S. Lagalwar, A. L. Guillozet, M. Lu, Y. Fu, F. Garcia-Sierra, N. LaPointe, R. Miller, R. W. Berry, L. I. Binder, V. L. Cryns, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10032.

50. K. Pagel, S. C. Wagner, K. Samedov, H. von Berlepsch, C. Bottcher, B. Koksch, *J. Am. Chem. Soc.* **2006**, *128*, 2196.

51. R. A. Kammerer, D. Kostrewa, J. Zurdo, A. Detken, C. Garcia-Echeverria, J. D. Green, S. A. Muller, B. H. Meier, F. K. Winkler, C. M. Dobson, M. O. Steinmetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4435.

52. A. I. Bush, *Curr. Opin. Chem. Biol.* **2000**, *4*, 184.

53. J. W. Karr, L. J. Kaupp, V. A. Szalai, *J. Am. Chem. Soc.* **2004**, *126*, 13534.

54. K. Pagel, T. Seri, H. von Berlepsch, J. Griebel, R. Kirmse, C. Bottcher, B. Koksch, *ChemBioChem* **2008**, *9*, 531.

55. E. Cerasoli, B. K. Sharpe, D. N. Woolfson, *J. Am. Chem. Soc.* **2005**, *127*, 15008.

56. T. Lynch, R. A. Cherny, A. I. Bush, *Exp. Gerontol.* **2000**, *35*, 445.

57. L. N. Johnson, R. J. Lewis, *Chem. Rev.* **2001**, *101*, 2209.

58. D. Bossemeyer, *FEBS Lett.* **1995**, *369*, 57.

59. D. Barford, Z. C. Jia, N. K. Tonks, *Nat. Struct. Biol.* **1995**, *2*, 1043.

60. P. Cohen, *Eur. J. Biochem.* **2001**, *268*, 5001.

61. L. N. Johnson, D. Barford, *Annu. Rev.Biophys. Biom. Struct*. **1993**, *22*, 199.

62. R. M. Stroud, *Curr. Opin. Struct. Biol.* **1991**, *1*, 826.

63. C. D. Andrew, J. Warwicker, G. R. Jones, A. J. Doig, *Biochemistry* **2002**, *41*, 1897.

64. L. Szilak, J. Moitra, D. Krylov, C. Vinson, *Nat. Struct. Biol.* **1997**, *4*, 112.

65. R. S. Signarvic, W. F. DeGrado, *J. Mol. Biol.* **2003**, *334*, 1.

66. N. Errington, A. J. Doig, *Biochemistry* **2005**, *44*, 7553.

67. L. Szilak, J. Moitra, C. Vinson, *Prot. Sci.* **1997**, *6*, 1273.

68. A. J. Riemen, M. L. Waters, *J. Am. Chem. Soc.* **2009**, *131*, 14081.

69. A. A. Bielska, N. J. Zondlo, *Biochemistry* **2006**, *45*, 5527.

70. A. Roque, I. Ponte, J. L. R. Arrondo, P. Suau, *Nucleic Acids Res.* **2008**, *36*, 4719.

71. S. Balakrishnan, N. J. Zondlo, *J. Am. Chem. Soc.* **2006**, *128*, 5590.

72. L. Settimo, S. Donnini, A. H. Juffer, R. W. Woody, O. Marin, *Biopolymers* **2007**, *88*, 373.

73. L. L. Liu, K. J. Franz, *J. Biol. Inorg. Chem.* **2007**, *12*, 234.

74. N. L. Huq, K. J. Cross, E. C. Reynolds, *J. Pept. Sci.* **2003**, *9*, 386.

75. T. Hunter, *Cell* **1995**, *80*, 225.

76. M. Goedert, C. M. Wischik, R. A. Crowther, J. E. Walker, A. Klug, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4051.

77. M. Goedert, M. G. Spillantini, M. C. Potier, J. Ulrich, R. A. Crowther, *Embo J.* **1989**, *8*, 393.

78. M. Goedert, M. G. Spillantini, R. Jakes, D. Rutherford, R. A. Crowther, *Neuron* **1989**, *3*, 519.

79. M. D. Weingarten, A. H. Lockwood, S. Y. Hwo, M. W. Kirschner, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1858.

80. Y. Tatebayashi, N. Haque, Y. C. Tung, K. Iqbal, I. Grundke-Iqbal, *J. Cell. Sci.* **2004**, *117*, 1653.

81. R. Brandt, J. Leger, G. Lee, *J. Cell. Biol.* **1995**, *131*, 1327.

82. Q. Hua, R. Q. He, N. Haque, M. H. Qu, A. del Carmen Alonso, I. Grundke-Iqbal, K. Iqbal, *Cell. Mol. Life. Sci.* **2003**, *60*, 413.

83. A. Rendon, D. Jung, V. Jancsik, *Biochem. J.* **1990**, *269*, 555.

84. K. Bhaskar, S. H. Yen, G. Lee, *J. Biol. Chem.* **2005**, *280*, 35119.

85. A. L. Fink, *Curr. Opin. Struct. Biol.* **2005**, *15*, 35.

86. O. Schweers, E. Schonbrunn-Hanebeck, A. Marx, E. Mandelkow, *J. Biol. Chem.* **1994**, *269*, 24290.

87. S. Jeganathan, M. von Bergen, H. Brutlach, H. J. Steinhoff, E. Mandelkow, *Biochemistry* **2006**, *45*, 2283.

88. M. D. Mukrasch, S. Bibow, J. Korukottu, S. Jeganathan, J. Biernat, C. Griesinger, E. Mandelkow, M. Zweckstetter, *PLoS Biol.* **2009**, *7*, e34.

89. D. W. Cleveland, S. Y. Hwo, M. W. Kirschner, *J. Mol. Biol.* **1977**, *116*, 227.

90. C. X. Gong, F. Liu, I. Grundke-Iqbal, K. Iqbal, *J. Neural. Transm.* **2005**, *112*, 813.

91. K. Iqbal, I. Grundke-Iqbal, T. Zaidi, P. A. Merz, G. Y. Wen, S. S. Shaikh, H. M. Wisniewski, I. Alafuzoff, B. Winblad, *Lancet* **1986**, *2*, 421.

92. A. C. Alonso, T. Zaidi, I. Grundke-Iqbal, K. Iqbal, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5562.

93. M. Perez, F. Hernandez, A. Gomez-Ramos, M. Smith, G. Perry, J. Avila, *Eur. J. Biochem.* **2002**, *269*, 1484.

94. A. Alonso, T. Zaidi, M. Novak, I. Grundke-Iqbal, K. Iqbal, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6923.

95. A. C. Alonso, I. Grundke-Iqbal, K. Iqbal, *Nat. Med.* **1996**, *2*, 783.

96. M. Goedert, R. Jakes, R. A. Crowther, P. Cohen, E. Vanmechelen, M. Vandermeeren, P. Cras, *Biochem. J.* **1994**, *301 (Pt 3)*, 871.

97. M. Goedert, R. Jakes, E. Vanmechelen, *Neurosci. Lett.* **1995**, *189*, 167.

98. R. Hoffmann, V. M. Lee, S. Leight, I. Varga, L. Otvos, Jr., *Biochemistry* **1997**, *36*, 8114.

99. T. Bussiere, P. R. Hof, C. Mailliot, C. D. Brown, M. L. Caillet-Boudin, D. P. Perl, L. Buee, A. Delacourte, *Acta Neuropathol.* **1999**, *97*, 221.

100. A. Sengupta, J. Kabat, M. Novak, Q. Wu, I. Grundke-Iqbal, K. Iqbal, *Arch. Biochem. Biophys.* **1998**, *357*, 299.

101. C. Alonso Adel, A. Mederlyova, M. Novak, I. Grundke-Iqbal, K. Iqbal, *J. Biol. Chem.* **2004**, *279*, 34873.

102. M. Goedert, R. Jakes, Z. Qi, J. H. Wang, P. Cohen, *J. Neurochem.* **1995**, *65*, 2804.

103. G. Lee, R. Thangavel, V. M. Sharma, J. M. Litersky, K. Bhaskar, S. M. Fang, L. H. Do, A. Andreadis, G. Van Hoesen, H. Ksiezak-Reding, *J. Neurosci.* **2004**, *24*, 2304.

104. P. Derkinderen, T. M. Scales, D. P. Hanger, K. Y. Leung, H. L. Byers, M. A. Ward, C. Lenz, C. Price, I. N. Bird, T. Perera, S. Kellie, R. Williamson, W. Noble, R. A. Van Etten, K. Leroy, J. P. Brion, C. H. Reynolds, B. H. Anderton, *J. Neurosci.* **2005**, *25*, 6584.

105. C. X. Gong, S. Shaikh, J. Z. Wang, T. Zaidi, I. Grundke-Iqbal, K. Iqbal, *J. Neurochem.* **1995**, *65*, 732.

106. M. Bennecib, C. X. Gong, I. Grundke-Iqbal, K. Iqbal, *FEBS Lett.* **2001**, *490*, 15.

107. F. Liu, K. Iqbal, I. Grundke-Iqbal, G. W. Hart, C. X. Gong, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10804.

108. F. I. Comer, G. W. Hart, *J. Biol. Chem.* **2000**, *275*, 29179.

109. C.-X. Gong, F. Liu, I. Grundke-Iqbal, K. Iqbal, *J. Alzheimers Dis.* **2006**, *9*, 1.

110. J. Z. Wang, I. Grundke-Iqbal, K. Iqbal, *Nat. Med.* **1996**, *2*, 871.

111. Y. Sato, Y. Naito, I. Grundke-Iqbal, K. Iqbal, T. Endo, *FEBS Lett.* **2001**, *496*, 152.

112. F. Liu, T. Zaidi, K. Iqbal, I. Grundke-Iqbal, R. K. Merkle, C. X. Gong, *FEBS Lett.* **2002**, *512*, 101.

113. F. Liu, K. Iqbal, I. Grundke-Iqbal, C. X. Gong, *FEBS Lett.* **2002**, *530*, 209.

114. M. Kidd, *Nature* **1963**, *197*, 192.

115. R. A. Crowther, C. M. Wischik, *EMBO J.* **1985**, *4*, 3661.

116. J. Berriman, L. C. Serpell, K. A. Oberg, A. L. Fink, M. Goedert, R. A. Crowther, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9034.

117. S. Barghorn, P. Davies, E. Mandelkow, *Biochemistry* **2004**, *43*, 1694.

118. M. Sadqi, F. Hernandez, U. Pan, M. Perez, M. D. Schaeberle, J. Avila, V. Munoz, *Biochemistry* **2002**, *41*, 7150.

119. H. Wille, G. Drewes, J. Biernat, E. M. Mandelkow, E. Mandelkow, *J. Cell. Biol.* **1992**, *118*, 573.

120. O. Schweers, E. M. Mandelkow, J. Biernat, E. Mandelkow, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8463.

121. P. Friedhoff, A. Schneider, E. M. Mandelkow, E. Mandelkow, *Biochemistry* **1998**, *37*, 10223.

122. P. Friedhoff, M. von Bergen, E. M. Mandelkow, P. Davies, E. Mandelkow, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15712.

123. S. Gandy, A. J. Czernik, P. Greengard, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6218.

124. M. Oishi, A. C. Nairn, A. J. Czernik, G. S. Lim, T. Isohara, S. E. Gandy, P. Greengard, T. Suzuki, *Mol. Med.* **1997**, *3*, 111.

125. P. E. Tarr, R. Roncarati, G. Pelicci, P. G. Pelicci, L. D'Adamio, *J. Biol. Chem.* **2002**, *277*, 16798.

126. T. Suzuki, T. Nakaya, *J. Biol. Chem.* **2008**, *283*, 29633.

127. A. E. Aplin, G. M. Gibb, J. S. Jacobsen, J. M. Gallo, B. H. Anderton, *J. Neurochem.* **1996**, *67*, 699.

128. K. Iijima, K. Ando, S. Takeda, Y. Satoh, T. Seki, S. Itohara, P. Greengard, Y. Kirino, A. C. Nairn, T. Suzuki, *J. Neurochem.* **2000**, *75*, 1085.

129. C. L. Standen, J. Brownlees, A. J. Grierson, S. Kesavapany, K. F. Lau, D. M. McLoughlin, C. C. Miller, *J. Neurochem.* **2001**, *76*, 316.

130. T. A. Ramelot, L. K. Nicholson, *J. Mol. Biol.* **2001**, *307*, 871.

131. K. Ando, M. Oishi, S. Takeda, K. Iijima, T. Isohara, A. C. Nairn, Y. Kirino, P. Greengard, T. Suzuki, *J. Neurosci.* **1999**, *19*, 4421.

132. M. S. Lee, S. C. Kao, C. A. Lemere, W. M. Xia, H. C. Tseng, Y. Zhou, R. Neve, M. K. Ahlijanian, L. H. Tsai, *J. Cell. Biol.* **2003**, *163*, 83.

133. N. G. Milton, *Neuroreport* **2001**, *12*, 3839.

134. C. J. Phiel, C. A. Wilson, V. M. Lee, P. S. Klein, *Nature* **2003**, *423*, 435.

135. J. C. Cruz, D. Kim, L. Y. Moy, M. M. Dobbin, X. Sun, R. T. Bronson, L. H. Tsai, *J. Neurosci.* **2006**, *26*, 10536.

136. M. Flajolet, G. He, M. Heiman, A. Lin, A. C. Nairn, P. Greengard, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4159.

137. A. Ferreira, Q. Lu, L. Orecchio, K. S. Kosik, *Mol. Cell. Neurosci.* **1997**, *9*, 220.

138. A. Alvarez, R. Toro, A. Caceres, R. B. Maccioni, *FEBS Lett.* **1999**, *459*, 421.

139. T. Prapong, J. Buss, W. H. Hsu, P. Heine, H. West Greenlee, E. Uemura, *Exp. Neurol.* **2002**, *174*, 253.

140. M. Rapoport, H. N. Dawson, L. I. Binder, M. P. Vitek, A. Ferreira, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6364.

141. P. Y. Chou, G. D. Fasman, *Biochemistry* **1974**, *13*, 211.

142. A. Chakrabartty, T. Kortemme, R. L. Baldwin, *Protein Sci*. **1994**, *3*, 843.

143. C. K. Smith, J. M. Withka, L. Regan, *Biochemistry* **1994**, *33*, 5510.

144. M. T. Pastor, A. Esteras-Chopo, M. Lopez de la Paz, *Curr. Opin. Struct. Biol.* **2005**, *15*, 57.

145. K. Pagel, B. Koksch, *Curr. Opin. Chem. Biol.* **2008**, *12*, 730.

146. D. N. Woolfson, *Adv. Protein Chem.* **2005**, *70*, 79.

147. H. Dong, J. D. Hartgerink, *Biomacromolecules* **2007**, *8*, 617.

148. F. H. Crick, *Nature* **1952**, *170*, 882.

149. E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **1991**, *254*, 539.

150. E. Wolf, P. S. Kim, B. Berger, *Protein Sci.* **1997**, *6*, 1179.

151. P. Burkhard, J. Stetefeld, S. V. Strelkov, *Trends Cel.l Biol.* **2001**, *11*, 82.

152. P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, *262*, 1401.

153. P. B. Harbury, P. S. Kim, T. Alber, *Nature* **1994**, *371*, 80.

154. L. Gonzalez, Jr., D. N. Woolfson, T. Alber, *Nat. Struct. Biol.* **1996**, *3*, 1011.

155. D. L. McClain, H. L. Woods, M. G. Oakley, *J. Am. Chem. Soc.* **2001**, *123*, 3151.

156. O. D. Monera, C. M. Kay, R. S. Hodges, *Biochemistry* **1994**, *33*, 3862.

157. D. L. McClain, J. P. Binfet, M. G. Oakley, *J. Mol. Biol.* **2001**, *313*, 371.

158. T. J. Graddis, D. G. Myszka, I. M. Chaiken, *Biochemistry* **1993**, *32*, 12664.

159. P. Burkhard, S. Ivaninskii, A. Lustig, *J. Mol. Biol.* **2002**, *318*, 901.

160. N. E. Zhou, C. M. Kay, R. S. Hodges, *Protein Engineering* **1994**, *7*, 1365.

161. W. D. Kohn, C. M. Kay, R. S. Hodges, *Protein Sci.* **1995**, *4*, 237.

162. D. N. Woolfson, T. Alber, *Protein Sci.* **1995**, *4*, 1596.

163. M. Mutter, R. Gassmann, U. Buttkus, K. H. Altmann, *Angew. Chem. Int. Ed.* **1991**, *30*, 1514.

164. Y. Takahashi, A. Ueno, H. Mihara, *Chem. Eur. J.***1998**, *4*, 2475.

165. B. Ciani, E. G. Hutchinson, R. B. Sessions, D. N. Woolfson, *J. Biol. Chem.* **2002**, *277*, 10150.

166. K. Pagel, S. C. Wagner, R. Rezaei Araghi, H. von Berlepsch, C. Bottcher, B. Koksch, *Chem. Eur. J.* **2008**, *14*, 11442.

167. T. Kimmerlin, D. Seebach, *J. Pept. Res.* **2005**, *65*, 229.

168. R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

169. R. M. Hofmann, T. W. Muir, *Curr. Opin. Biotech.* **2002**, *13*, 297.

170. J. McMurray, D. R. t. Coleman, W. Wang, M. L. Campbell, *Biopolymers* **2001**, *60*, 3.

171. N. L. Daly, R. Hoffmann, L. Otvos, Jr., D. J. Craik, *Biochemistry* **2000**, *39*, 9039.

172. M. Inoue, A. Hirata, K. Tainaka, T. Morii, T. Konno, *Biochemistry* **2008**, *47*, 11847.

173. L. X. Zhou, Z. Y. Zeng, J. T. Du, Y. F. Zhao, Y. M. Li, *Biochem. Biophys. Res. Commun.* **2006**, *348*, 637.

174. W. C. Chan, P. D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, **2000**.

175. T. Wakamiya, R. Togashi, T. Nishida, K. Saruta, J. Yasuoka, S. Kusumoto, S. Aimoto, K. Y. Kumagaye, K. Nakajima, K. Nagata, *Bioorg. Med. Chem.* **1997**, *5*, 135.

176. E. A. Ottinger, L. L. Shekels, D. A. Bernlohr, G. Barany, *Biochemistry* **1993**, *32*, 4354.

177. T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, W. Schafer, *Justus Liebigs Annalen der Chemie* **1953**, *583*, 129.

178. D. S. Kemp, S. L. Leung, D. J. Kerkman, *Tetrahedron Letters* **1981**, *22*, 181.
179. D. S. Kemp, D. J. Kerkman, *Tetrahedron Lett.* **1981**, *22*, 185.

180. P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776.

181. T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068.

182. P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325.

183. Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc.* **1999**, *121*, 11684.

184. M. Huse, M. N. Holford, J. Kuriyan, T. W. Muir, *J. Am. Chem. Soc.* **2000**, *122*, 8337.

185. F. Mende, O. Seitz, *Angew. Chem. Int. Ed.* **2007**, *46*, 4577.

186. R. von Eggelkraut-Gottanka, A. Klose, A. G. Beck-Sickinger, M. Beyermann, *Tetrahedron Lett.* **2003**, *44*, 3551.

187. J. Offer, C. N. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642.

188. C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorg. Med. Chem.* **2004**, *12*, 2749.

189. L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526.

190. Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248.

191. C. Haase, H. Rohde, O. Seitz, *Angew. Chem. Int. Ed.* **2008**, *47*, 6807.

192. D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064.

193. T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705.

194. H. Paulus, *Annu. Rev.Biochem.* **2000**, *69*, 447.

195. T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.

196. G. J. Cotton, B. Ayers, R. Xu, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 1100.

197. G. J. Cotton, T. W. Muir, *Chem. Biol.* **2000**, *7*, 253.

198. D. Schwarzer, P. A. Cole, *Curr. Opin. Chem. Biol.* **2005**, *9*, 561.

199. K. P. Chiang, M. S. Jensen, R. K. McGinty, T. W. Muir, *ChemBioChem* **2009**, *10*, 2182.

200. C. P. R. Hackenberger, C. T. Friel, S. E. Radford, B. Imperiali, *J. Am. Chem. Soc.* **2005**, *127*, 12882.

201. A. Rak, O. Pylypenko, T. Durek, A. Watzke, S. Kushnir, L. Brunsveld, H. Waldmann, R. S. Goody, K. Alexandrov, *Science* **2003**, *302*, 646.

202. M. R. Seyedsayamdost, C. S. Yee, J. Stubbe, *Nat. Protoc.* **2007**, *2*, 1225.

203. M. C. Chang, C. S. Yee, D. G. Nocera, J. Stubbe, *J. Am. Chem. Soc.* **2004**, *126*, 16702.

204. J. Mukhopadhyay, A. N. Kapanidis, V. Mekler, E. Kortkhonjia, Y. W. Ebright, R. H. Ebright, *Cell* **2001**, *106*, 453.

205. J. W. Wu, M. Hu, J. Chai, J. Seoane, M. Huse, C. Li, D. J. Rigotti, S. Kyin, T. W. Muir, R. Fairman, J. Massague, Y. Shi, *Mol. Cell.* **2001**, *8*, 1277.

206. B. M. Chacko, B. Y. Qin, A. Tiwari, G. Shi, S. Lam, L. J. Hayward, M. De Caestecker, K. Lin, *Mol. Cell.* **2004**, *15*, 813.

207. M. E. Hahn, T. W. Muir, *Angew. Chem. Int. Ed.* **2004**, *43*, 5800.

208. W. Zheng, Z. Zhang, S. Ganguly, J. L. Weller, D. C. Klein, P. A. Cole, *Nat. Struct. Biol.* **2003**, *10*, 1054.

209. W. Lu, K. Shen, P. A. Cole, *Biochemistry* **2003**, *42*, 5461.

210. I. Giriat, T. W. Muir, *J. Am. Chem. Soc.* **2003**, *125*, 7180.

211. H. Staudinger, J. Meyer, *Helvetica Chimica Acta* **1919**, *2* 635.

212. B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939.

213. E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141.

214. R. Kleineweischede, C. P. Hackenberger, *Angew. Chem. Int. Ed.* **2008**, *47*, 5984.

215. A. Tam, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2007**, *129*, 11421.

216. J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem. Int. Ed.* **2006**, *45*, 1248.

217. G. D. Fasman, *Circular dichroism and the conformational analysis of biomolecules.*, Plenum Press: New York, **1996**.

218. K. Pagel, Doctoral Thesis, Freie Universität Berlin, **2007**.

219. S. M. Kelly, T. J. Jess, N. C. Price, *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2005**, *1751*, 119.

220. N. Greenfield, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108.

221. W. C. Johnson, *Proteins-Structure Function and Genetics* **1990**, *7*, 205.

222. J. Reed, T. A. Reed, *Anal. Biochem.* **1997**, *254*, 36.

223. N. J. Greenfield, *Anal. Biochem.* **1996**, *235*, 1.

224. H. Levine, *Protein Sci.* **1993**, *2*, 404.

225. R. Khurana, C. Coleman, C. Ionescu-Zanetti, S. A. Carter, V. Krishna, R. K. Grover, R. Roy, S. Singh, *J. Struct. Biol.* **2005**, *151*, 229.

226. M. R. H. Krebs, E. H. C. Bromley, A. M. Donald, *J. Struct. Biol.* **2005**, *149*, 30.

227. S. Jeganathan, A. Hascher, S. Chinnathambi, J. Biernat, E. M. Mandelkow, E. Mandelkow, *J. Biol. Chem.* **2008**, *283*, 32066.

228. H. Ding, T. A. Matthews, G. V. W. Johnson, *J. Biol. Chem.* **2006**, *281*, 19107.

229. I. Landrieu, L. Lacosse, A. Leroy, J. M. Wieruszeski, X. Trivelli, A. Sillen, N. Sibille, H. Schwalbe, K. Saxena, T. Langer, G. Lippens, *J. Am. Chem. Soc.* **2006**, *128*, 3575.

230. S. G. Greenberg, P. Davies, J. D. Schein, L. I. Binder, *J. Biol. Chem.* **1992**, *267*, 564.

231. A. Abraha, N. Ghoshal, T. C. Gamblin, V. Cryns, R. W. Berry, J. Kuret, L. I. Binder, *J. Cell Sci.* **2000**, *113*, 3737.

232. T. Li, H. K. Paudel, *Biochemistry* **2006**, *45*, 3125.

233. T. Li, C. Hawkes, H. Y. Qureshi, S. Kar, H. K. Paudel, *Biochemistry* **2006**, *45*, 3134.

234. M. Goedert, R. Jakes, *EMBO J.* **1990**, *9* 4225.