Synthesis of Functional Proteins via Multiple Bioorthogonal Ligations

by

Monica Gene Pirigyi

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Degree in Biochemistry with Distinction

Spring 2013

Synthesis of Functional Proteins via Multiple Bioorthogonal Ligations

by

Monica Gene Pirigyi

Approved: _____
Neal J. Zondlo, Ph. D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Joseph M. Fox, Ph. D.
Committee member from the Department of Chemistry and Biochemistry

Approved: _____
Salil Lachke, Ph. D.
Committee member from the Board of Senior Thesis Readers

Approved: _____
Michael Arnold, Ph.D.
Director, University Honors Program

ACKNOWLEDGMENTS

I am lucky to have had a tremendous amount of support and mentorship throughout my career in undergraduate research. I would especially like to thank Neal Zondlo for giving me the opportunity to work and grow in his lab. Neal is a passionate and devoted scientist who has motivated me to be the same way. I am very grateful for his constant encouragement, support, and confidence in me.

I am grateful for the wonderful group of graduate and undergraduate students in the lab that I have had the pleasure of working with. Everyone was always willing to teach me and answer any questions that I had. I would like to give a special thanks to Anil Pandey, my graduate mentor, for teaching me everything I know. I truly appreciate your patience, encouragement, and the help you gave me in deciphering data and planning experiments.

I would like to thank Dr. Joseph Fox and Dr. Salil Lachke, my second and third readers respectively, for aiding me through the process of formulating my thesis. I would also like to thank Dr. Harold White for his work in undergraduate education as well as the University of Delaware's Howard Hughes Medical Institute (HHMI) program. The HHMI program funded two consecutive summers of research for me.

And most importantly, I would like to thank my family, boyfriend, and friends for continually being there to support me along the way. Your encouragement has helped me surpass difficult failures and setbacks, and also to reach new heights after successes. Without your support this process would have been much more difficult.

TABLE OF CONTENTS

iv

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Large protein synthesis is a present day biochemical challenge. Proteins can efficiently be expressed from various sources; however, post-translational modifications to introduce specific phosphorylation and spectroscopic labels can be challenging. One promising mechanism in creating these large, functional proteins is bioconjugation. This work is directed toward the chemical synthesis of microtubule-binding proteins and microtubule binding studies to elucidate the mechanism of tau aggregation in Alzheimer's disease. Peptides from the tubulin-binding domain of Tau were synthesized on an automated peptide synthesizer and modifications were conducted, on both solid and solution phase, to allow incorporation of specific phosphorylation events and multiple bioorthogonal functionalities. Peptides were modified and peptide conjugation reactions were conducted. To elucidate the efficiency of these bioconjugation reactions we have used model peptides and demonstrated clean, efficient conversions in aqueous solutions. Further, we are utilizing these tools to synthesize proteins with multiple tubulin-binding domain repeats and to study their ability to bind, polymerize, and depolymerize microtubules in both their nonphosphorylated and phosphorylated states.

Chapter 1

INTRODUCTION

1.1   Functional Protein Synthesis

Synthesis of a fully functional protein is a present day biochemical challenge. Synthetic proteins can be selectively modified at individual residues with post-translational modifications, isotopic labels, unnatural amino acids, sugar groups, polymers, fluorophores, and other chemical markers that have a plethora of applications (22). Peptides can be synthesized on solid phase in automated fashion, in which the peptide is built on an insoluble, porous resin bead using fluorenylmethyloxycarbonyl (Fmoc) or *tert*-butyloxycarbonyl (Boc) protected amino acids. However modern synthesizers are limited to about 50 amino acids and have compromised yields (1). These peptides could be modified and studied, however the results are not necessarily indicative of how the full length protein would interact and behave (2).

A number of methods have been developed to study modified fully functional proteins. Proteins can be pseudo-modified by mutating the amino acid sequence. For instance, pseudo-phosphorylation is done by replacing a serine or threonine residue with an aspartic or glutamic acid. However, the carboxylic acid groups on the acidic residues are not a perfect representation of a phosphorylated protein (2). Proteins can

also be enzymatically modified in vitro, but these modifications do not yield a homogeneous product because the proteins are modified in different positions and to varying extents (2).

Bioorthogonal ligations and native chemical ligation (NCL) can allow for the synthesis of a functional protein that can be site-specifically modified with phosphates, fluorescent probes, and other moieties. NCL links an N-terminal thiolate to a C-terminal thioester, thereby extending the length of peptides that can be synthesized on solid phase (22). The Hackenberger group from FU Berlin used this chemistry to study the ability of the protein tau to stabilize microtubules in its nonphosphorylated and phosphorylated state. They demonstrated the ability to synthesize a semisynthetic protein using expressed protein ligation (EPL) via NCL to synthesize the tau protein. The group ligated together a portion of recombinant tau, comprised of 441 amino acids, and a fifty residue peptide synthesized on solid phase. By incorporating on a synthetic peptide they had the ability to selectively phosphorylate certain residues and incorporate a biotin tag. Tubulin polymerization assays showed that there were no functional differences between the wild type tau and the semisynthetic tau, so modification was achieved and the behavior of the full protein was still retained (2).

Semisynthetic proteins still pose a challenge, however, if the protein is difficult to purify to homogeneity. Intrinsically unstructured proteins, like tau, are difficult to purify through standard methods and the procedure is time-consuming (2). Complete synthesis of a functional protein via high yielding bioorthogonal ligations of solid

phase synthesized peptides could significantly reduce the difficulties in protein purification. Importantly, the protein would be available for selective modification of certain residues with different moieties and probes. The consequence on the behavior and structure of those modifications could be studied using protein assays.

1.2 Bioorthogonal Ligations

Bioorthogonal reactions are a unique type of chemistry that function under normal physiological conditions and do not interfere with any native biological functional groups, with the exception of metal catalysts. These reactions are rapid and selective (3). As shown in Figure 1, each bioorthogonal functionality reacts only with its counterpart, inert to the variety of functionalities typically found in vivo.

$$A + B \longrightarrow A-B$$
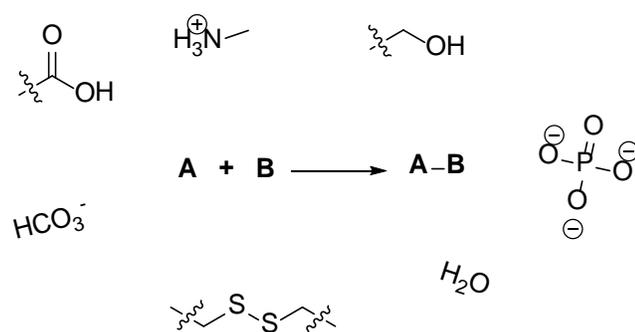
Figure 1:   A bioorthogonal reaction. The bioorthogonal functionalities A and B react in the presence of biologically relevant functionalities without crossreactivity (3).

Bioorthogonal chemistry has extensively been used in protein bioconjugation, in which two successive steps are required. First, a bioorthogonal functionality must be incorporated onto the biomolecule, in this case a protein. The functionality can be

incorporated synthetically or engineered into the biosynthetic pathway (4). Functionalities are typically incorporated onto the side chains of cysteine and lysine, and less commonly tyrosine, glutamate, aspartate, and histidine (3). The N-terminus can also be functionalized with different bioorthogonal moieties, but there may be competition from other basic residues in the protein. After the protein has been functionalized a bioorthogonal reaction must take place between it and a chemical probe or other biomolecule (4).

These bioorthogonal reactions must take place rapidly under physiological conditions, have high yields, and have no cross-reactivity with any other molecules found in a biological system. In addition, if these reactions are conducted in cells, the reactants and products must be stable under physiological conditions and must be non-toxic to living organisms (4).

The bioorthogonal reactions utilized in synthesizing a functional peptide are described below.

1.2.1 Hüisgen Azide-Alkyne Chemistry

The Huisgen azide-alkyne reaction, commonly known as "click chemistry", is a selective and efficient bioorthogonal reaction. The 1,3-dipolar cycloaddition is catalyzed by Cu^{I} and cyclizes an azide and an alkyne into a triazole, as shown in Figure 2 (20). Azides are not biologically present and are usually small enough to not disturb the substrate (3). In addition, azides and alkynes have weak acid-base

properties, making them essentially inert to many biological compounds in living cells (20).

Click chemistry using a copper catalyst is not applicable in living biological systems because of its cytotoxicity (4). The Bertozzi group from UC Berkley has developed a copper-independent reaction that utilizes the ring strain of cyclooctyne to cyclize with azide, as shown in figure 2 (21). The Fox group from the University of Delaware has also developed a copper-independent Diels-Alder cycloaddition between *trans*-cyclooctene and tetrazine, as shown in figure 2 (5). The Cu-free click reactions developed by the Bertozzi and Fox groups have good reaction rates. However, for the purpose of synthesizing a functional protein in vitro, the cytotoxicity of the copper catalyzed click reaction is insignificant, as the protein will be purified from reagents.
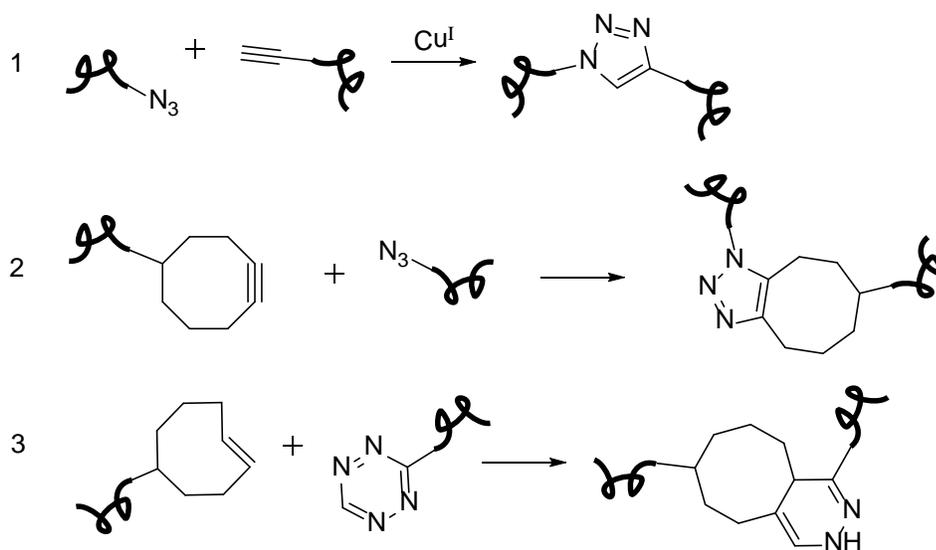
Figure 2:   1. Copper catalyzed click reaction using an azide and alkyne; 2. Bertozzi: Copper-free click chemistry using cyclooctyne and an azide (21); 3. Fox: Copper-free click chemistry using *trans*-cyclooctene and tetrazine (5).

5

1.2.2 Oxime Chemistry

An oxyamine and aldehyde react to form a stable oxime, shown in Figure 3. Aldehydes are a useful bioorthgoonal group because they are small and inert towards many other functional groups present in biological systems. The Zondlo group demonstrated that the oxime reaction proceeds rapidly and without a catalyst using aldehydes and 4S-aminoxyproline, among others (18).

Figure 3:   An oxyamine and aldehyde react to form an oxime.

1.2.3 Suzuki-Miyaura Chemistry

The Suzuki-Miyaura cross-coupling reaction forms a stable carbon-carbon bond between a boronic acid and aryl halide. The Suzuki group first published this reaction in 1981and reported good yields. They ran their reactions in benzene at reflux temperature using sodium carbonate and 3 mole% of the catalyst tetrakis(triphenylphosphine)Pd^0 (Pd(PPh$_3$)$_4$) for 6 hours, or longer for sterically hindered bromoarenes (Figure 4) (16). The phosphine catalyst is subjective to oxidation, so degassed solvents were used.
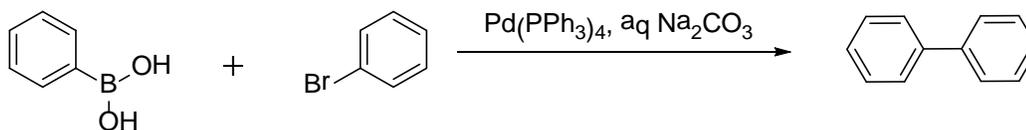
Figure 4:   Suzuki-Miyaura Reaction to form a stable carbon-carbon bond using the
             catalyst Pd(PPh₃)₄.

In order for the Suzuki-Miyaura reaction to be biologically friendly and applicable to proteins, it had to react under benign conditions, in aqueous solution, and at ambient temperature. The Davis group found that the sodium salt of the ligand 2-amino-4,6-dihydroxypyrimidine formed a soluble complex with Pd(OAc)₂ in water. The ligand catalyzed the Suzuki-Miyaura reaction in model peptides with good yields at 37°C, and did not require oxygen exclusion or organic solvents (17). The Zondlo group demonstrated that the reaction between boronic acid and aryl bromide proceeds rapidly and cleanly using the Davis conditions and generates a high conversion under neutral, aqueous conditions (Figure 5) (18).

It is unclear if the palladium catalyst proposes a toxicity problem to living organisms. Boronic acids also have a modest affinity for polysaccharides, so this reaction may not be ideal for biological systems (4). However, for the purpose of synthesizing a functional protein this reaction will create a stable ligation.
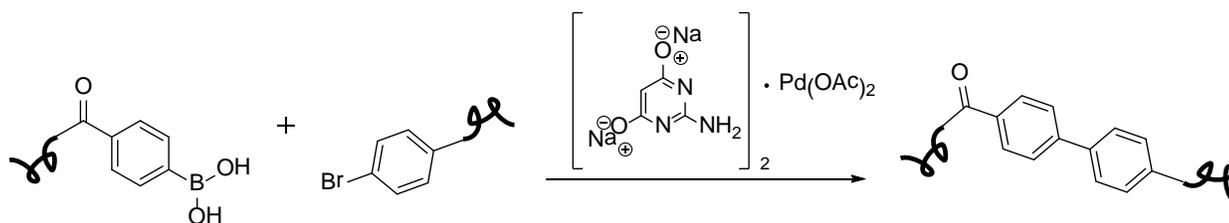
Figure 5:   Suzuki-Miyaura cross coupling reaction catalyzed by the Davis Ligand
             (17).

1.3   Alzheimer's Disease

Alzheimer's disease (AD) is the sixth leading cause of death in the United States, affecting one out of eight older Americans (6). The neurodegenerative disease results in a decline in cognitive function and loss of memory. Patients with AD develop two types of brain lesions, senile plaques, made up of Aβ amyloid fibrils, and neurofibrillary tangles (NFT).

NFTs are aggregates of paired helical filaments (PHFs), made up of the neuronal protein tau. They are observed largely in the pyramidal cells of the entorhinal cortex and hippocampus, as well as the infragranular layers of the association cortical areas (7).

AD was first discovered in 1906 by Alois Alzheimer. Back then the disease was rare because few people lived long enough to be at risk. Decades later with the life expectancy being some thirty years higher, AD is a serious problem. 50% of people over the age of 85 exhibit symptoms of Alzheimer's disease. Of the top ten leading causes of death in America, AD is the only disease without a method of prevention or cure (6). Aside from its health risks, AD is draining the American economy. In 2012, $200 billion dollars were spent to care for those with AD, and without action the bill will total $1.1 trillion dollars in 2050 (6). Research on the cause, early detection, and treatment of AD is essential to the future of our country.

1.4 Microtubule-Associated Protein Tau

Tau is an intrinsically unstructured protein that belongs to the microtubule-associated protein family. Tau is typically found in the neurons of the peripheral and central nervous system, though it is also present in non-neuronal cells in small quantities (7). Tau's main function is to bind to and stabilize the assembly of tubulin subunits into microtubules. Tau also supports the outgrowth of axons and regulates axonal transport (8).

Tau is present in six different isoforms in the brain that result from alternative splicing, and they are not equally expressed in neurons (9, 7). The longest isoform of tau is 441 amino acids long, and is shown schematically in Figure 6. The protein is composed of two hydrophobic domains, A and B, the proline rich domain (PRD), and the several tubulin-binding domain (TBD) repeats near the carboxy terminus. The different isoforms of tau are distinguished by the number of TBD repeats present, either three or four, and the absence or presence of one or two inserts at the amino-terminal end (7).

Figure 6:   Diagram of the largest tau isoform in the human brain. The protein consists of the proline rich domain (PRD), shown in green, two hydrophobic regions, A and B shown in red, and either three our four tubulin binding domains (TBD) shown in blue.

1.5   Post-translational Phosphorylation of Tau

Tau undergoes a variety of post-translational modifications that provide the protein with a wide variety of functional and regulatory behaviors. Phosphorylation is the most important post-translational modification that affects tau's ability to bind microtubules (10). Tau is phosphorylated by a variety of protein kinases. In vitro, tau is mostly phosphorylated by mitogen activated protein kinase (MAP), tau-tubulin kinase, glycogen synthase kinase 3 (GSK3), and cyclin-dependent protein kinase 5 (cdk5) (7). Tau can also be phosphorylated by casein kinase II, calcium/calmodulin-dependent kinase II (CaMKII), p34 kinase, and cyclic AMP-dependent protein kinase (pKA) (10).   Microtubule affinity-regulating kinase (MARK) phosphorylates the KXGS motifs in the TBD of tau (23). In the longest isoform of tau, MARK has been found to phosphorylate Ser_{262}, Ser_{293}, Ser_{324}, and Ser_{356} with a resulting detachment of tau from the microtubules (24). Overexpression of MARK leads to the hyperphoshorylation of microtubule-associated proteins, like tau, and disrupts the microtubule array (26). The kinases that phosphorylate tau in vivo have been extensively investigated, but haven't been fully identified yet.

There are presumed to be seventy-nine possible phosphorylation sites on the longest isoform of tau, all of which are either serine or threonine residues (7). Several of the phosphorylation sites are located within the TBD. Ser_{262} is located in the first repeat, Ser_{285} between repeat one and repeat 2, Ser_{305} between repeat two and three, Ser_{324} in repeat two, and Ser_{352} and Ser_{356} in the fourth repeat (7). Ser_{262} and Ser_{356} are among the phosphorylation sites used as markers for Alzheimer's disease.

Phosphorylation of Ser_{262} elicits an early phosphorylation-related change in tau and is an epitope to the antibody S262. The antibody 12E8 reacts with Ser_{356} as well as Ser_{262}. Phosphorylation of Ser_{356} and Ser_{262} can be detected by 12E8 in pre-NFTs (phospho-tau, normal cell morphology), intra-neuronal NFTs (phospho-tau, filament aggregation, deteriorated dendrites), and extra-neuronal NFTs (phospho-tau filaments, collapsed dendrites) (28).

The amount of phosphorylation on the tau protein is a delicate balance and can affect microtubule binding (11). Tau is normally weakly phosphorylated and does not aggregate to form PHFs (7). The PHFs are made up of hyperphosphorylated tau (7). In PHF-tau approximately thirty serine or threonine residues are phosphorylated (12). The hyperphosphorylation may be either caused by a decrease in phosphatase activity or an increase in kinase activity (7).

1.6   Microtubule-binding Domain (TBD)

The TBD is made up of three or four repeats consisting of approximately thirty one amino acids, separated by one, two, or zero inter-repeat residues (13). The repeat region is thought to act like a flexible linker that binds microtubules; however its main function is to stabilize microtubules against disassembly, not initiating microtubule polymerization, though in vitro it effectively promotes microtubule assembly and polymerization (14). The TBD has a positive charge that binds the negatively-charged C-terminal end of the microtubules electrostatically (27). A single repeat domain itself

weakly binds microtubules, but in combination with the flanking domains the binding is strong (13).

The TBD has been identified as the core structure of the PHFs that make up NFTs, and has been identified as the most important region of the protein involved in tau aggregation (8, 27).The nonphosphorylated TBD repeats have been shown to aggregate into filaments similar to the PHFs found in Alzheimer's disease (25). Both the presence of a tau protein with an abnormal conformation and post-translational modifications have been shown to be involved in tau aggregation and AD development (27). The Mandelkow group developed a generalized mechanism for tau aggregation into PHFs. Tau is normally unstructured, but can conformationally change to form β-sheet interactions with other molecules. In this manner, two tau molecules can form a dimer in the nucleation step that then allows for other tau molecules to associate into the full PHF. The dimer is thought to be formed by β-sheet development around the hexapeptides in repeats 2 and 3 of the TBD. When only the TBD repeats are used the PHFs form more rapidly (28).

Mandelkow described the "paperclip model" which shows that though tau is unstructured, it has a superstructure that is modulated by post-translational modifications. In the "paperclip model", both the C and N terminus of tau fold inward toward the TBD, as shown in Figure 7 (31). Phosphorylation of the Ser_{199}, Ser_{202}, and Thr_{205} in the PRD, shown by a reaction with the AT8 antibody produce, moves the N-terminal domain away from the C-terminal domain (28, 32). Phosphorylation of Ser_{396} and Ser_{404} towards the C-terminus of tau, shown by a reaction with the PHF1

antibody, moves the C-terminal domain towards the N-terminal domain (28, 32). However phosphorylation of residues in both the PRD and C-terminus of tau, shown by the reaction with both AT8, AT100 (Thr$_{212}$ and Ser$_{214}$), and PHF1 results in a compaction of the paperclip and triggers the reactivity of antibodies Alz50 and MC1 that signify early stage Alzheimer's disease (28, 32). Compaction of the paperclip allows tau to form aggregates more rapidly (32).
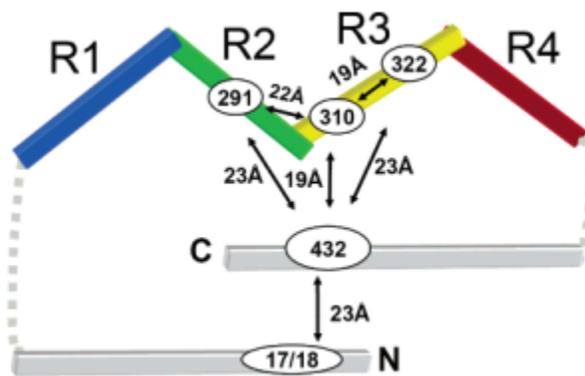
Figure 7:   Mandelkow's "Paperclip Model" of tau. The TBD repeats, shown in blue, green, yellow, and red, associate with the C-terminal and N-terminal ends of the tau protein (31).

Aside from aggregation, hyperhosphorylation has a profound impact on the ability of the TBD to bind microtubules and results in microtubule depolymerization (27). However, there are so many different kinases that can phosphorylate tau that it is difficult to characterize the modified residues (13). Ser$_{262}$, located in the first TBD repeat, has been shown to be phosphorylated in PHF-tau (10). Phosphorylation of Ser$_{262}$ strongly inhibits the ability of tau to bind microtubules, but also inhibits PHF formation (10, 29).

Ser$_{356}$, located in the fourth TBD repeat, is another known abnormal phosphorylation site (8). Ser$_{356}$ is located in a KXGS motif, and is specifically phosphorylated by MARK (24). Researchers at Tsinghua University in Beijing studied the effects of phosphorylation of Ser$_{356}$ in the fourth repeat of the TBD using turbidity, ThT fluorescence, and electron microscopy. Results showed that phosphorylation of Ser$_{356}$ in the fourth TBD repeat causes a modulated effect on aggregation (8). Studying the effect of phosphorylation of Ser$_{356}$ on microtubule binding and stabilization could provide further insight on the effect of post-translational modifications on tau that ultimately lead to microtubule destabilization and polymerization, PHF formation, and the general mechanism of Alzheimer's disease.

1.7 Chemical Synthesis of a Functional Protein via Multiple Bioorthogonal Ligations

Chemical synthesis of tau is a useful alternative to pseudo-phosphorylation, enzymatic modification, and the use of semisynthetic proteins because it allows for the selective modification of certain residues without the difficulty of multiple protein purifications. Multiple bioorthogonal reactions could be used to synthesize the problematic TBD. The TBD is highly prone to aggregation and the full repeat domain is too long to synthesize by solid phase peptide synthesis alone. For ease of synthesis, one repeat of the TBD can be synthesized, and then ligated to other repeats to form the entire repeat domain.

Of the four repeats, the fourth domain is the optimal choice for synthesis. The fourth repeat domain contains Ser$_{356}$ allowing for the analysis of the effect of phosphorylation on microtubule-binding. It does not contain a cysteine residue in its

sequence, like repeats two and three (Figure 8). This allows for the incorporation of a C-terminal cysteine for selective solution phase modifications on the thiol. Lastly, the fourth repeat is more soluble than repeat one, and will be easier to work with.

R1 240-270   K**S**RLQTAPVPMPDLKNVK*SKIGS*TENLKHQP
R1 271-301   GGGKVQIINKKLDL*S*NVQ*SKCGS*KDNIKHVP
R3 302-333   GGG*S*VQIVYKPVDL*S*KVT*SKCGS*LGNIHHKPG
R4 334-365   GGQVEVK**S**EKLDFKDRVQ*SKIGS*LDNITHVPG**C**

Figure 8:   Sequences for the four repeats of the TBD in the longest isoform of tau. The serine residues are highlighted in green and the cysteine residues are highlighted in blue, with the exception of Cys_{366} in red because it is not native to the sequence. The KXGS motifs are italicized to highlight possible sites for MARK phosphorylation. The underlined residues are known phosphorylation sites.

The fourth repeat of the microtubule-binding domain was synthesized, $tau_{334-365}$, with the addition of a C-terminal cysteine residue to incorporate groups via solution phase thiol chemistry (Figure 9). During synthesis, the Ser_{356} was protected with a trityl protecting group as opposed to the typical *tert*-butyl protecting group. The trityl protecting group allows the alcohol to be modified with a phosphate group. The $tau_{334-365}$ peptide was modified, both N-terminally and C-terminally, to introduce bioorthogonal functionalities.

The N-terminal $tau_{334-365}$ peptide was acetylated at the N-terminus for stability and the C-terminal $tau_{334-365}$ peptide was labeled with fluorescein at its C-terminal cysteine. The peptides were modified to incorporate a boronic acid, azide, and hydroxylamine to correspondingly ligate with an aryl halide, alkyne, and aldehyde (Figure 10). Using bioorthogonal chemistry we hypothesize that we can recreate the

microtubule-binding domain with novel linkers and the ability to bind and stabilize microtubules in vitro. The ability of the unphosphorylated, synthetic microtubule-binding domain to bind microtubules was tested using a microtubule-binding assay. The same tau$_{334-365}$ peptides were synthesized with the addition of a phosphate group to the hydroxyl of Ser$_{356}$. We hypothesize that we can observe the effect phosphorylating Ser$_{356}$ has on microtubule-binding using microtubule-binding assays. This method improves upon past efforts to study the effect of phosphorylation on microtubule stabilization because it uses actual phosphate groups, as opposed to pseudo-phosphates, and allows for the study of the entire TBD as opposed to a small peptide, giving a more realistic picture. In addition, it eliminates the process of purifying portions of expressed tau protein. Using bioorthogonal ligations to synthesize fully functional proteins can allow for selective residue modifications and the ability to see the effects of those modifications on a more macroscopic level.

tau$_{334-365}$ **G$_{334}$GQVEVKSEKLDFKDRVQSKIGS̲LDNITHVPG$_{365}$C**

Figure 9:   The sequence of the fourth repeat domain of the longest isoform of tau, with the first thirty amino acids corresponding to residues 334-365. A C-terminal cysteine was added to the sequence for modification purposes. The underlined Ser$_{356}$ was protected with a trityl group for selective phosphorylation.
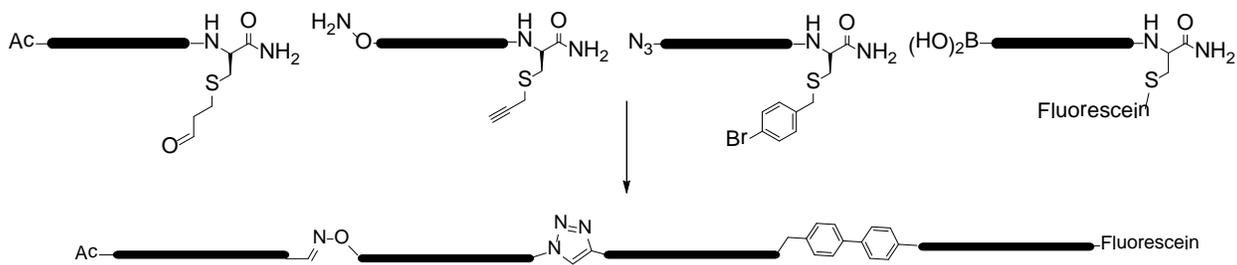
Figure 10: The schematic for synthesis of the four repeat microtubule-binding domain. The sequence of each peptide is shown in Figure 6. The peptides will be modified N-terminally and C-terminally to allow for the following reactions: oxime reaction, Hüisgen azide-alkyne reaction, and the Suzuki-Miyaura reaction, sequentially.

Chapter 2

MATERIALS AND METHODS

2.1   Materials

Fmoc-L-amino acids were purchased from Nova Biochem (San Diego, CA), AnaSpec, and Chem-Impex. Rink Amide MBHA resin, diisopropylethylamine (DIPEA), fluorenylmethoxycarbonyloxysuccinimate (Fmoc), t-butylcarbonylanhydride, and O-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Chem-Impex (Wood Dale, IL). Pyridine, dimethyl sulfoxide (DMSO), acetonitrile, dichloromethane (DCM), acetone, potassium carbonate, bis-acrylamide, sodium hydroxide, sodium bicarbonate, ammonium persulfate (APS), and N, N, N', N-tetramethylethylenediamine (TEMED), and acetic anhydride were purchased from Fischer Scientific (Fair Lawn, NJ). Ethanedithiol was purchased from TCI America (Portland, OR). Phenol, thioanisole, triethylsilane (TES), ether, propargyl bromide, dimethylsulfate (DMS), trans-4-hydroxy-L-proline, sodium azide, N-hydroxy phthalimide, 4-nitrobenzoic acid, hydrazine hydrate, and dimethylformamide (DMF) were purchased from Acrôs. Piperidine, triphenylphosphine, and diisopropyl azodicarboxylate (DIAD) were purchased from Sigma Aldrich. 1-Hydroxy-7-azabenzotriazole (HoAT) was purchased from AnaSpec and N-hydroxybenzotriazole (HOBT) was purchased from Fluka Chemika. 4-Carboxyphenylboronic acid was purchased from Combi-Blocks, Inc.

Deionized water was purified by a Millipore Synergy 185 water purification system with Simpak2 cartridge. Post-synthetic solid-phase reactions were performed in capped disposable fritted columns (Image Molding).

2.2   Peptide Synthesis

Peptides (0.25 mmol) were synthesized on Rink amide resin via standard solid phase peptide synthesis using a Rainin PS3 peptide synthesizer. Prior to amino acid coupling, the resin was swelled in DMF (4   10 minutes). Amino acids were coupled for either 1 or 2 hours or double-coupled, using 4 equivalents of HBTU and 4 equivalents of the Fmoc protected amino acid. Each coupling of an amino acid included the following four steps: (1) Removal of the Fmoc protecting group using 20% piperidine in DMF (3   5 minutes); (2) Resin wash in DMF (5   1 minute); (3) amino acid coupling using HBTU and 8% DIPEA in DMF (1 hour, 2 hours, or double coupling); (4) resin wash in DMF (3   1 minute). The tau$_{334\text{-}365}$ peptide was made in increments, coupling amino acids 1 through 10, then adding 11-22, then 23 through 29, and finally 30 through 33. The resin was washed with DMF (3   ), DCM (3   ), and ether (3   ) to dry.

Tau$_{334\text{-}365}$ peptides were cleaved for 3 hours in 15 mL of TFA and 80 μL each of water, thioanisole, phenol, and EDT. The TFA was partially evaporated to about 4 mL with nitrogen, and then the solution was filtered through a fritted tube to remove the resin beads. The TFA/peptide mixture was then mixed with cold ether in a 1:10 ratio to precipitate the peptide. The solution was centrifuged for 5 minutes at 3.5 K, and

then the ether was removed via pipette. The dried peptides were dissolved in 10 microliters of DMF and 2 milliliters of pH 2.9 phosphate buffer.

The KKHMCX peptides were cleaved for 3 hours using a solution of 90% TFA, 5% TIS, and 5% H_2O. The TFA was removed by evaporation, the peptides were precipitated with cold ether, and the precipitate was dissolved in water. Peptides were filtered and purified by reverse phase HPLC on Vydac semi-preparative (C18) column using linear gradients of buffer B (20% H_2O, 80% MeCN, and 0.06% TFA) in buffer A (98% H_2O, 2% MeCN, and 0.06% TFA). The peptides were analyzed for purity by reinjection on a Varian HPLC analytical column (Microsorb-MV 100-5 C18 250   4.6 mm), and the peaks were characterized using mass spectrometry on either the Ultima Q Tof (Quadrupole-Time of Flight tandem mass spectrometer) or the Shumadzu LC/MS 2020 instrument.

$Tau_{334-365}$ peptide concentrations were determined by UV-Vis spectroscopy based on the equation: $(\mu g/mL) = (A_{215} - A_{225}) \times 144$ (15). KKHMCX peptide concentrations were determined by UV-VIS spectroscopy based on 4-iodophenylalanine absorbance ($\varepsilon_{280} = 280$ M^{-1} cm^{-1} in water).

2.3   N-Terminal Modification of TBD Peptides

To accomplish modification of the $tau_{334-365}$ peptides, the resin was swelled with DMF (10 minutes) and then the peptides were Fmoc deprotected using 20% Piperidine in DMF (3   5 minutes).

2.3.1   Tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$)

The tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) peptide was synthesized using 4-carboxyphenylboronic acid (0.5 M), 1-hydroxy-7-azabenzotriazole (HoAt) (0.5 M), HATU (1.5 eqv), and DIPEA (0.5 M) in DMF (Figure 11, 12). The reaction ran for 24 hours with mixing (Figure 6). Afterward the resin was washed with DMF (3), DCM (3), and ether (3).
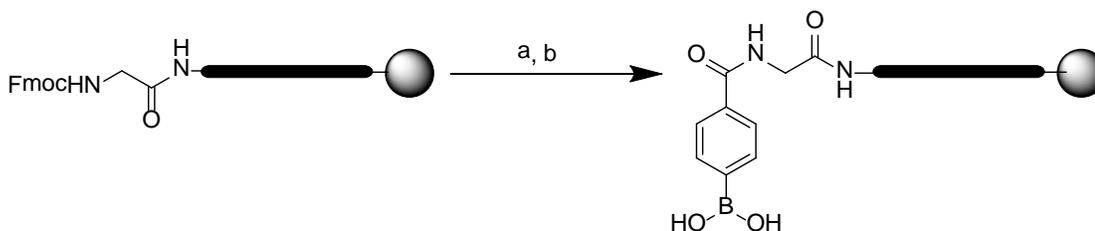
Figure 11:   Scheme for modification of tau$_{334\text{-}365}$ with boronic acid on resin. (a) 20% Piperidine/DMF (3   5 min); (b) 4-carboxyphenyl boronic acid, HATU, HOAt, DMF (24 hrs).
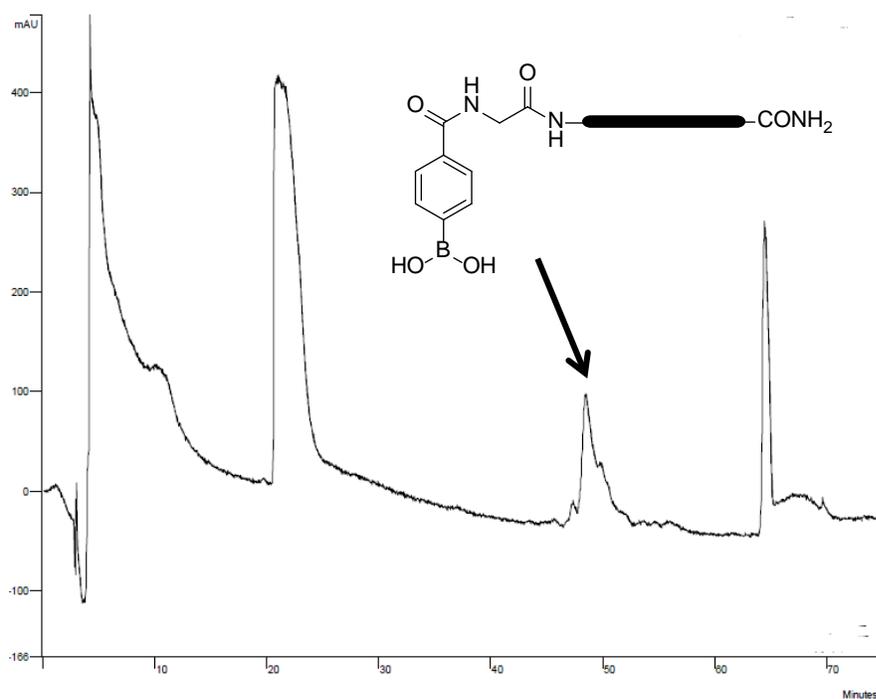
Figure 12: Crude HPLC chromatogram of tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) peptide. The peak with a t_R= 48.5 min. corresponds to the peptide with the boronic acid modification.

2.3.2 Tau$_{334\text{-}365}$ N-(N$_3$)

The azido modification of the peptide was synthesized in two steps. First, the N-terminus of the peptide was modified using bromoacetic acid (0.5 M), DMAP (0.05 M), and DIC (0.5 M) in anhydrous THF to achieve the bromoacetyl modification. The reaction was conducted for 24 hours with gentle mixing. The azido peptide was synthesized by reacting the bromoacetyl modified peptide with sodium azide (1 M) in anhydrous DMF for 24 hours at 40 °C with gentle stirring (Figure 13, 14). Afterward the resin was washed with DMF (3×), DCM (3×), and ether (3×).
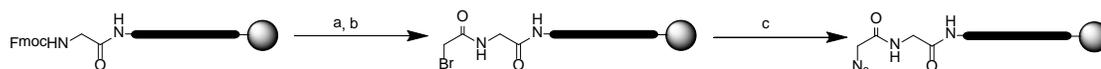
Figure 13: Scheme for modification of tau$_{334-365}$ with azide on resin. (a) 20% Piperidine/DMF (3   5 min); (b) bromoacetic acid, DIC, DMAP, Anhydrous THF; (c) NaN$_3$, DMF, 40 °C.
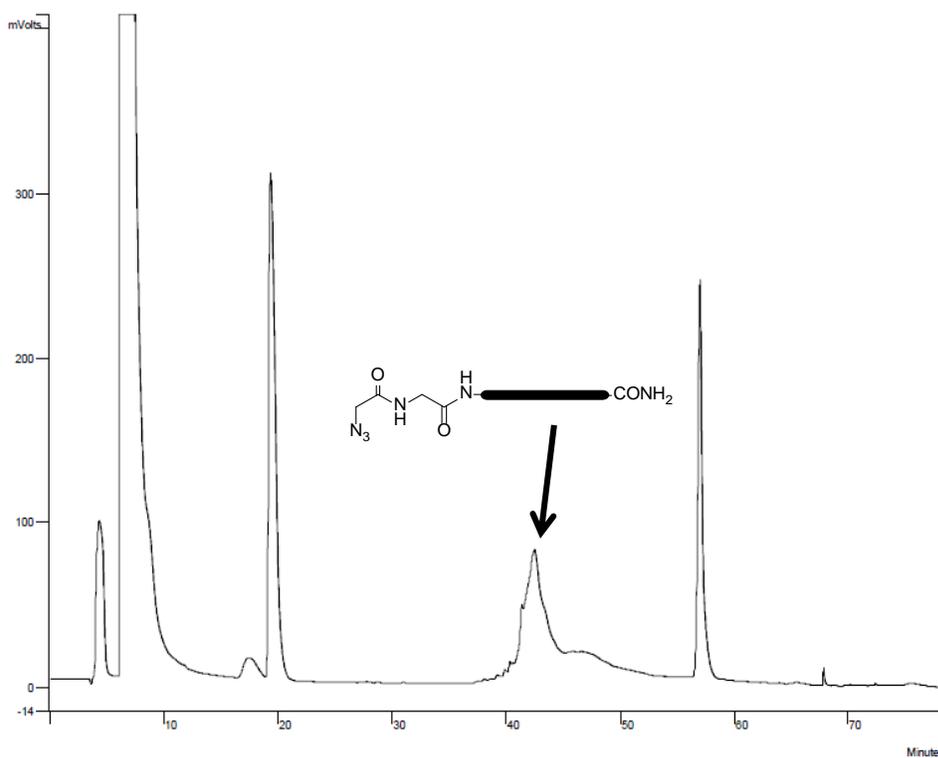
Figure 14: Crude HPLC chromatogram after the tau$_{334-365}$ peptide was modified with an azide. The peak with a t_R= 42.4 min. represents tau$_{334-365}$ N-(N$_3$).

2.3.3   Tau$_{334-365}$ N-(Ac)

The N-terminus acetylation was achieved using 5% acetic anhydride in pyridine (3 × 5 minutes) (Figure 15). Afterward the resin was washed with DMF (3×),

DCM (3×), and ether (3×). The reaction proceeded with excellent conversion (Figure 16).
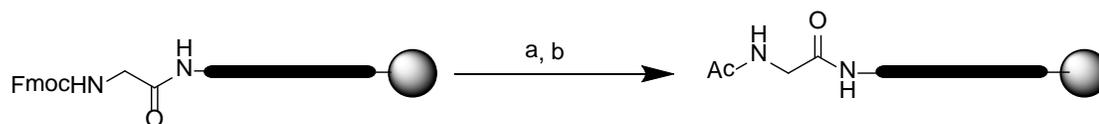
Figure 15:   Scheme for modification of $tau_{334-365}$ via acetylation. (a) 20% Piperidine/DMF (3 × 5 min); (b) 5% acetic anhydride/pyridine (3 × 5 min).
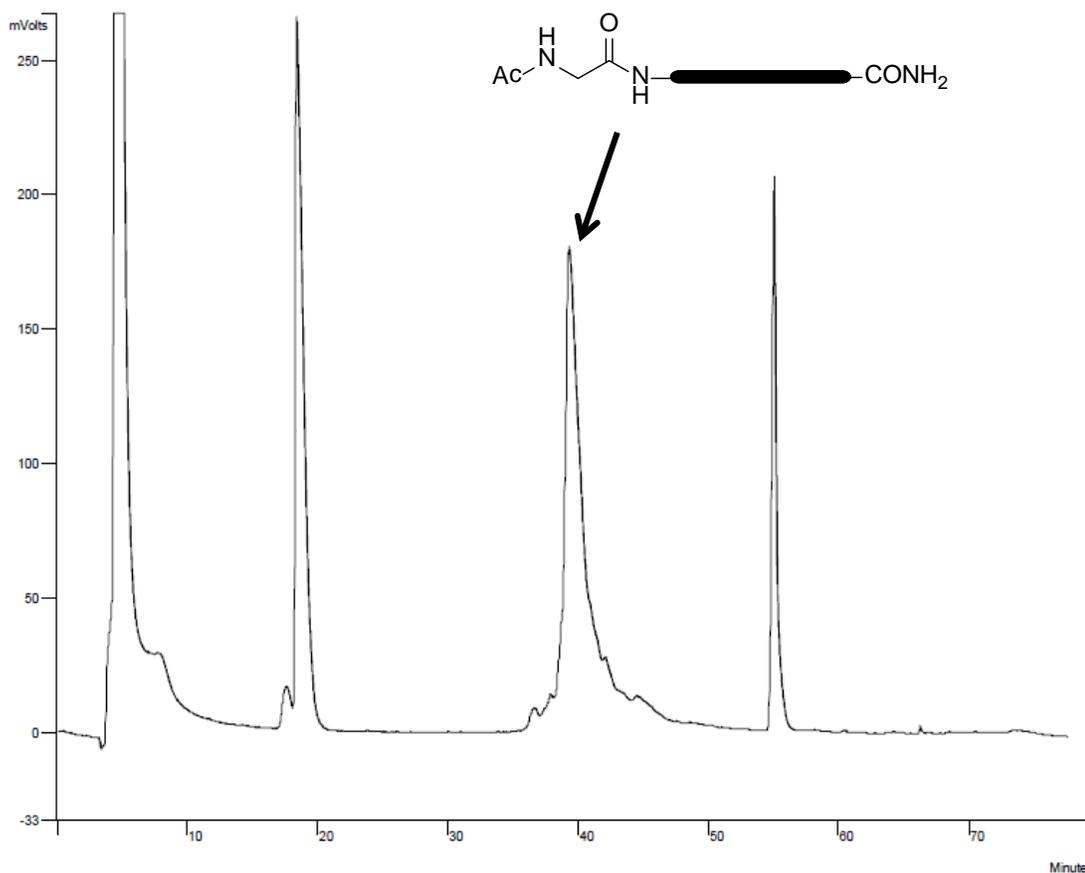
Figure 16:   Crude HPLC chromatogram of $tau_{334-365}$ N-(Ac) peptide. The large peak at t_R= 39.4 min. represents the acetylated peptide.

2.3.4   Tau$_{334-365}$ N-(O-phthalimide)

The oxyamine peptide was synthesized via coupling Fmoc-$_L$-Hyp-(4R-O-phthalimide) to the N-terminus. Fmoc-$_L$-Hyp-(4R-O-phthalimide) (100 mg) was coupled to the tau$_{334-365}$ peptide using the coupling reagents HATU (1.5 eqv) and HOBT (0.25 eqv) in an 8% solution of DIPEA in anhydrous DMF. The reaction ran for approximately 3 hours. The Hyp-(4R-O-phthalimide) was subsequently Fmoc detprotected using 8% piperidine in DMF (3   5 min) and acetylated using 5% acetic anhydride in pyridine (3   5 min) (Figure 17, 18). The deprotection of the phthalimide can be accomplished via reaction with hydrazine.
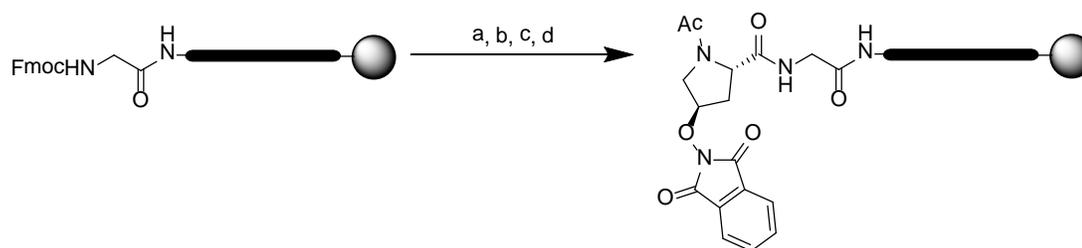
Figure 17:   Scheme for the coupling of tau$_{334-365}$ with Fmoc-$_L$-Hyp-(4R-O-phthalimide). (a) 20% Piperidine/DMF (3   5 min); (b) Fmoc-$_L$-Hyp-(4R-O-phthalimide), HATU, HOBT, DIPEA, DMF (3 hr); (c) 20% Piperidine/DMF (3   5 min); (d) 5% acetic anhydride/pyridine (3   5 min).
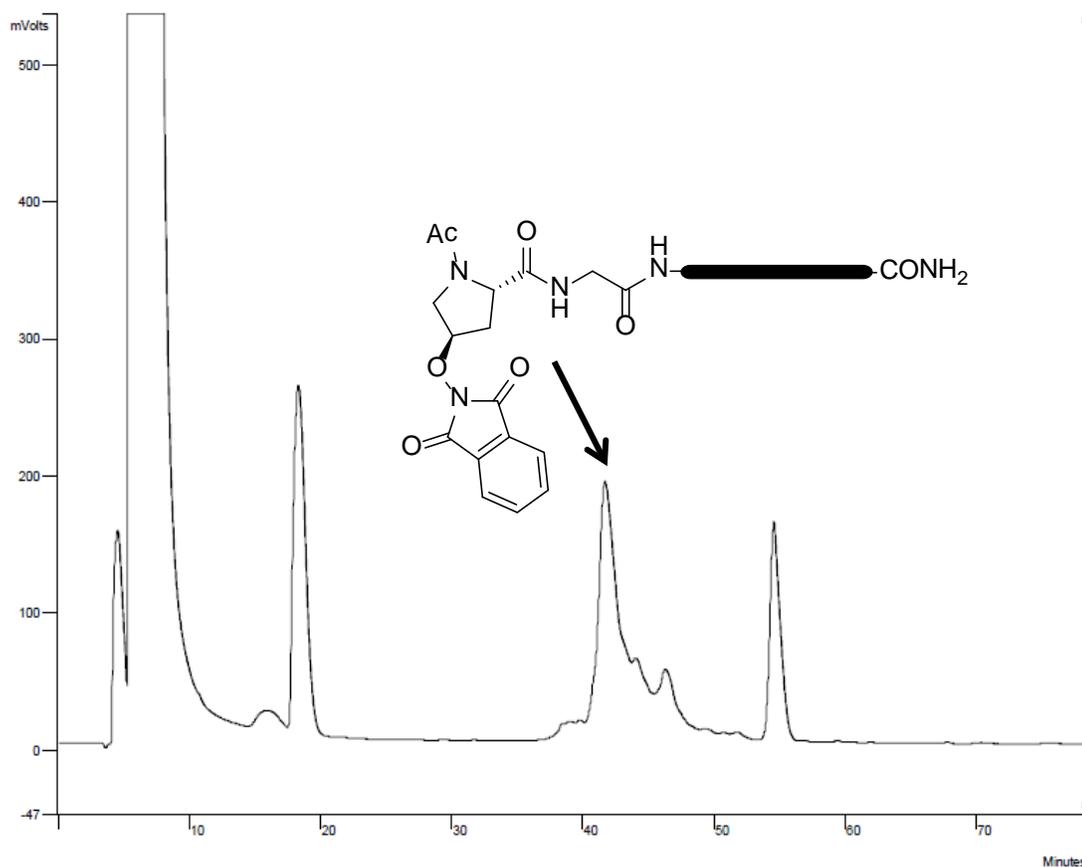
Figure 18:   Crude HPLC chromatogram of the acetylated tau$_{334-365}$ N-(O-phthalimide) peptide. The peak with a t_R= 41.7 min. represents the phthalimide modified peptide.

2.3.4.1   Synthesis of Fmoc-$_L$-Hyp-(4R-O-phthalimide)

Fmoc-$_L$-Hyp-(4R-O-phthalimide) was synthesized from the inexpensive starting material *trans*-4-$_L$-hydroxyproline (Hyp). The amino and carboxylic acid groups were protected with Boc and Methyl protecting groups, respectively. The inversion of the stereocenter at C-γ was accomplished via Mitsonobu reaction to yield the compound (2S, 4R)Hyp-(O-phthalimide). After acidic Boc deprotection and

hydrolysis of the ester, the amino group was Fmoc-protected for application in solid-phase peptide synthesis. The 8-step synthesis scheme is explained below:

To a solution of compound *trans*-4-hydroxy-L-proline (10.0 g, 76.3 mmol) in H_2O (50 mL) at 0 ºC was added $NaHCO_3$ (19.2 g, 229 mmol) slowly with stirring. A solution of Boc_2O (33.3 g, 152 mmol) in 1,4-dioxane (50 mL) was slowly added to the reaction mixture. The resulting reaction mixture was stirred at 0 ºC for 1 h and then stirred at room temperature overnight. The crude reaction mixture was acidified to pH 2 and extracted with ethyl acetate (3×). The organic layer was dried with Na_2SO_4 and solvent was removed in *vacuo*. The compound was purified by silica gel column chromatography (0-3% CH_3OH/CH_2Cl_2 v/v) to obtain N-Boc-2S,4R-hyp (14.80 g) as a viscous oil (84% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 4.27 (m, 1H), 3.56-3.52 (m, 1H), 3.27-3.25 (m, 1H), 2.40-2.30 (m, 1H), 1.99-1.93 (m,2H), and 1.40 (s, 9H). The NMR spectrum corresponded to the literature values (18, 34).
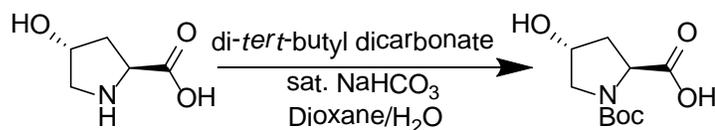
Figure 19:   Synthesis of N-Boc-2S,4R-hyp from *trans*-4-hydroxy-L-proline.

To a solution of N-Boc-2S,4R-hyp (6.0 g, 24.5 mmol) in dry acetone (100 mL) was added K_2CO_3 (10.1g, 73.5 mmol), followed by a slow addition of dimethyl sulfate (DMS) (3.9 mL, 41.7 mmol). The reaction mixture was refluxed with stirring for 2 h. The reaction mixture was allowed to cool to room temperature and subsequently filtered on a water aspirator. The solvent was removed in *vacuo*. The compound was purified by silica gel column chromatography (0-3% CH_3OH/CH_2Cl_2v/v) to obtain N-

Boc-2S,4R-hyp-OCH$_3$ (5.50 g) as a white solid (92% yield). ^{1}H NMR (400 MHz, CD$_3$OD) δ 4.27 (m, 1H), 3.58 (s, 3H), 3.56-3.52 (m, 1H), 3.27-3.25 (m, 1H), 2.40-2.30 (m, 1H), 2.01-1.95 (m, 2H), and 1.40 (s, 9H). The NMR spectrum corresponded to the literature values (18, 34).

Figure 20:   Synthesis of N-Boc-2S,4R-hyp-OCH$_3$ from N-Boc-2S,4R-hyp.

To N-Boc-2S,4R-hyp-OCH$_3$ (5.0 g, 20.4 mmol) dissolved in anhydrous THF (200 mL) at 0 ºC was added Ph$_3$P (8.0 g, 30.6 mmol) and 4-nitrobenzoic acid (4.1 g, 24.5 mmol). DIAD (6.0 mL, 30.6 mmol) was added slowly to the reaction mixture and was stirred vigorously for 1 h at 0 ºC. The ice bath was removed and stirring was continued for 6 h at room temperature. The white precipitate was filtered and the solvent was removed in *vacuo*. The resulting yellow precipitate was purified by silica gel column chromatography (0-2% CH$_3$OH/CH$_2$Cl$_2$ v/v) to yield compound N-Boc-2S,4S-Hyp(OBzNO$_2$)-OCH$_3$ (5.10 g) as a colorless oil (66% yield). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.56-7.45 (m, 4H), 4.25-4.23 (m, 2H), 3.63 (s, 3H), 3.51-3.48 (m, 1H), 3.25-3.24 (m, 1H), 2.32-2.27 (m, 1H), 1.96-1.93 (m, 1H), and 1.40 (s, 9H). The NMR data corresponded to the literature values (18, 34).
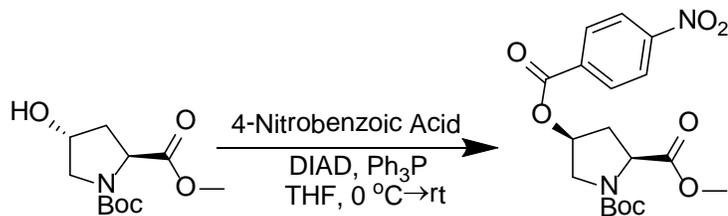
Figure 21: Synthesis of N-Boc-2S,4S-Hyp(OBzNO$_2$)-OCH$_3$ from N-Boc-2S,4R-hyp-OCH$_3$.

N-Boc-2S,4S-Hyp(OBzNO$_2$)-OCH$_3$ (3.0 g, 7.9 mmol) was hydrolyzed by reacting with NaN$_3$ (2.1 g, 31.6 mmol) in methanol (50 mL) at 45 ºC for 24 h. The solvent was removed *in vacuo* and the residue was dissolved in CH$_2$Cl$_2$ and washed with NaHCO$_3$ and brine. The organic phase was dried with Na$_2$SO$_4$ and concentrated in *vacuo*. The resulting viscous oil was purified by silica gel column chromatography (0-3% CH$_3$OH/CHCl$_3$v/v) to obtain N-Boc-2S,4S-Hyp-OCH$_3$ (1.50 g) as a colorless oil (79% yield). ^{1}H NMR (400 MHz, CD$_3$OD) δ 4.27 (m, 1H), 3.58 (s, 3H), 3.56-3.52 (m, 1H), 3.27-3.25 (m, 1H), 2.40-2.30 (m, 1H), 2.01-1.95 (m, 2H), and 1.40 (s, 9H). The NMR data corresponded to the literature values (18, 34).
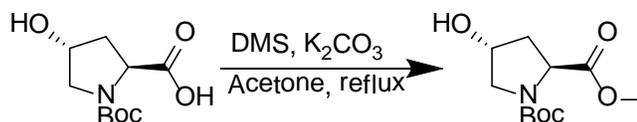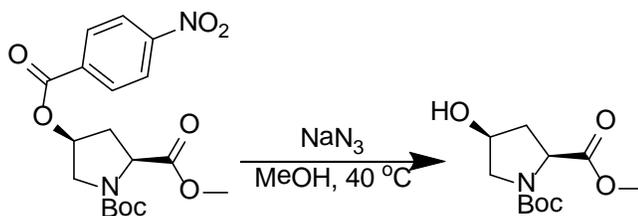

Figure 22: Synthesis of N-Boc-2S,4S-Hyp-OCH$_3$ from N-Boc-2S,4S-Hyp(OBzNO$_2$)-OCH$_3$.

N-Boc-2S,4S-hyp-OCH$_3$ (5.0 g, 20.4 mmol) dissolved in anhydrous THF (200 mL) at 0 ºC was added Ph$_3$P (8.0 g, 30.6 mmol) and N-hydroxy phthalimide (4.1 g, 24.5 mmol). DIAD (6.0 mL, 30.6 mmol) was added slowly to the reaction mixture and was stirred vigorously for 1 h at 0 ºC. The ice bath was removed and stirring was continued for 6 h at room temperature. The white precipitate was filtered and the solvent was removed in *vacuo*. The resulting yellow precipitate was purified by silica gel column chromatography (0-2% CH$_3$OH/CH$_2$Cl$_2$v/v) to yield compound N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$ (5.10 g) as a colorless oil (66% yield). ^{1}H NMR (400 MHz, CD$_3$OD) δ 7.80-7.72 (m, 4H), 4.18 (m, 1H), 3.63 (s, 3H), 3.51-3.48 (m, 1H), 2.32-2.27 (m, 1H), 2.20-1.96 (m, 1H), and 1.40 (s, 9H). The NMR data corresponded to the literature values. ^{13}C NMR (100 MHz, CD$_3$OD) δ 134.78, 128.76, 123.75, 85.71, 84.69, 80.48, 69.82, 57.71, 57.38, 52.32, 52.11, 51.19, 50.75, 35.59, 34.81, 28.33, 28.22, 21.93 (18, 34). The ^{1}H NMR and ^{13}C NMR spectra are shown in Figures 24 and 25, respectively.
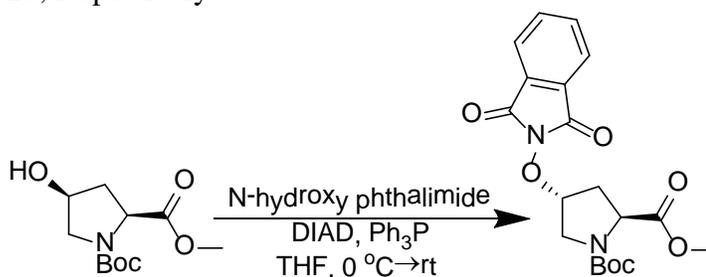
Figure 23:   Synthesis of N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$ from N-Boc-2S,4S-hyp-OCH$_3$.
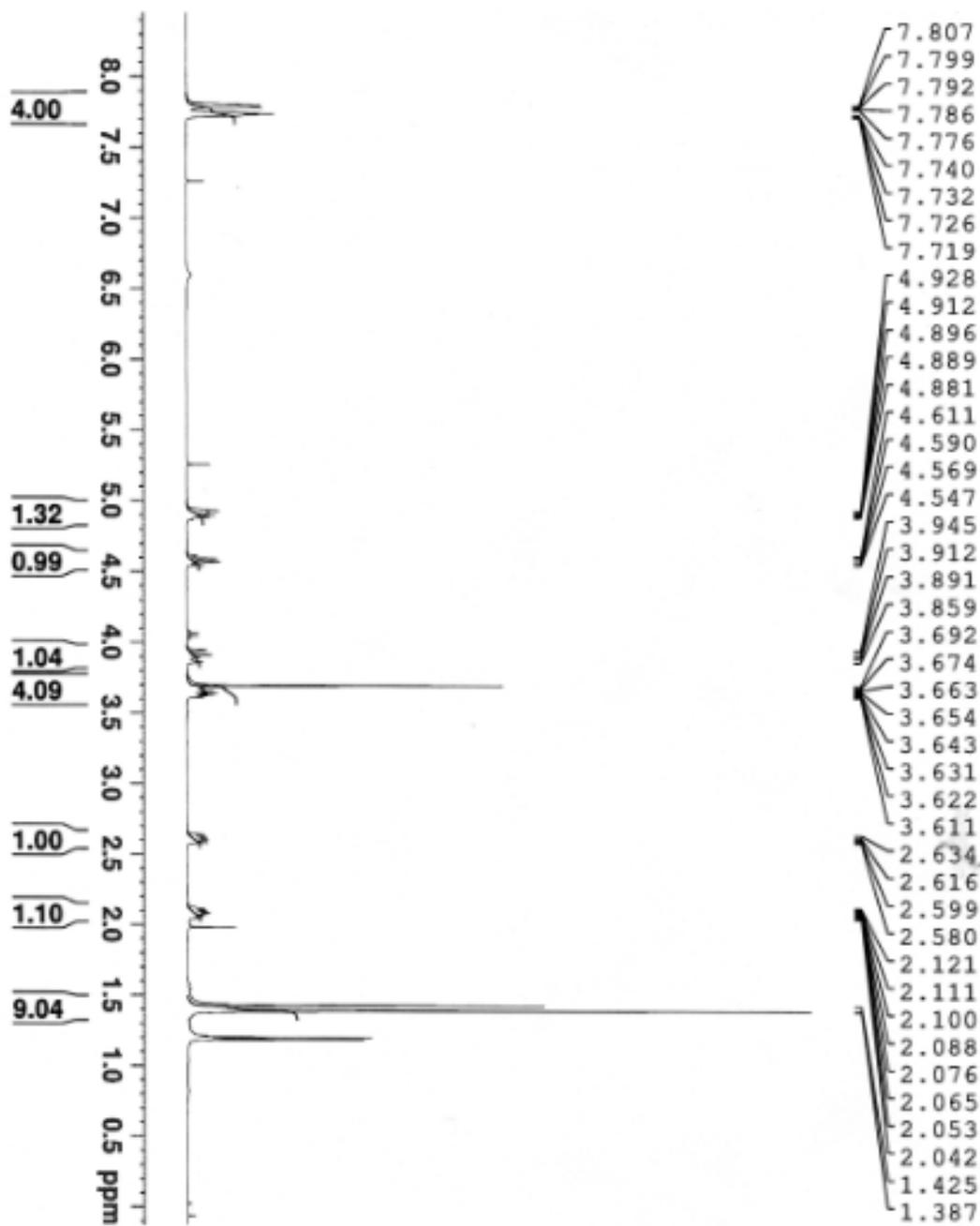
Figure 24: ^{1}H NMR of N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$.
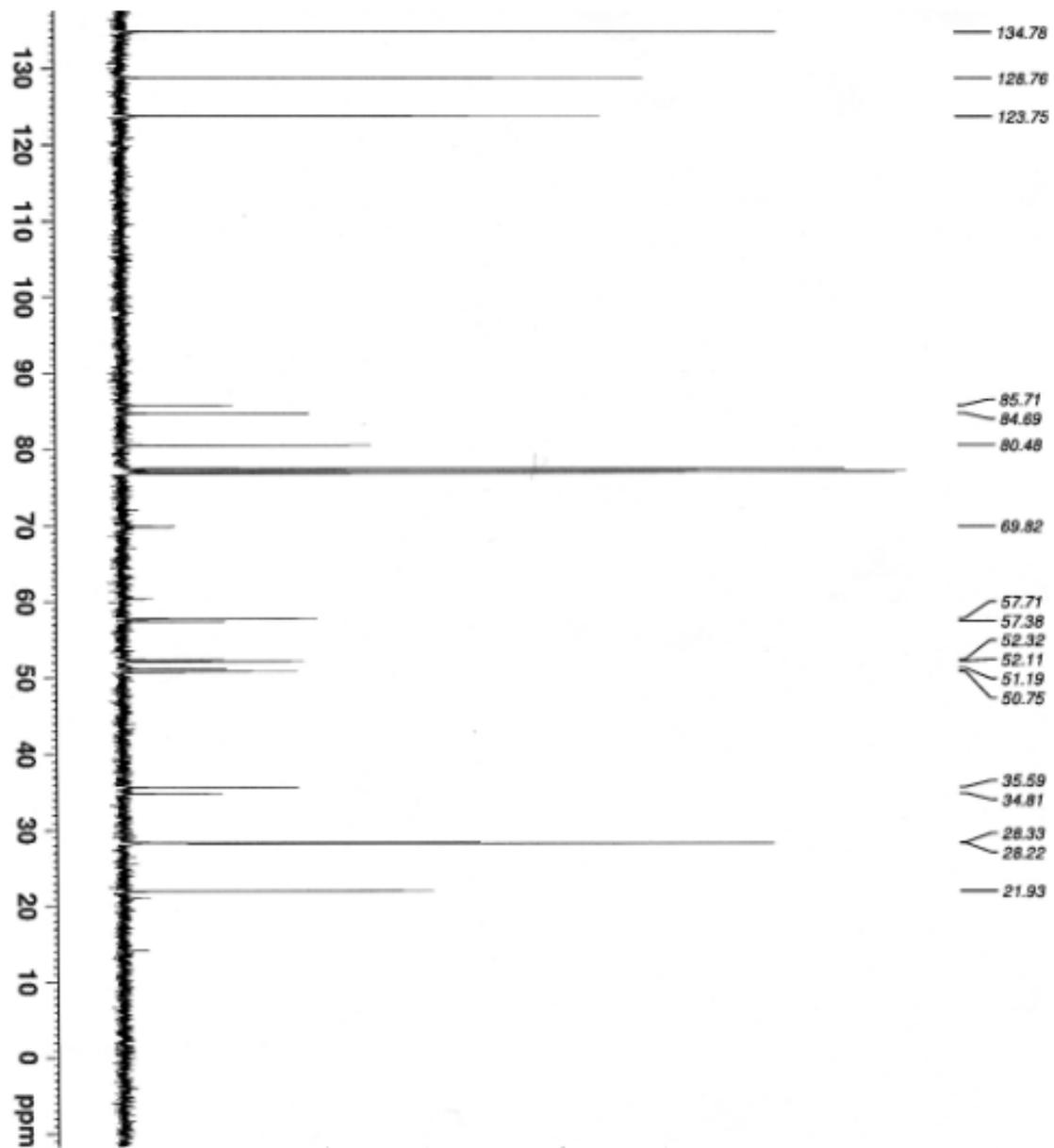
Figure 25: ^{13}C NMR of N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$.

To N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$ dissolved in 1,4-dioxane (10 mL) was added 4N HCl (10 mL). The reaction mixture was stirred at room temperature

overnight. The solvent was removed in *vacuo* to obtain N-Boc-2S,4R-Hyp(O-phthalimide) (18, 34).
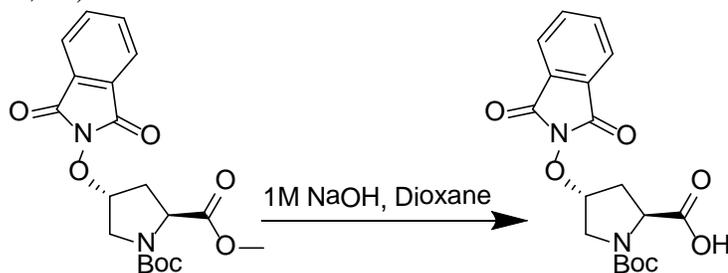
Figure 26:   Synthesis of N-Boc-2S,4R-Hyp(O-phthalimide) from N-Boc-2S,4R-Hyp(O-phthalimide)-OCH$_3$.

To N-Boc-2S,4R-Hyp(O-phthalimide) dissolved in H$_2$O was added LiOH. The reaction was mixed vigorously at room temperature overnight. The organic layer was dried with Na$_2$SO$_4$ and solvent was removed in *vacuo*. (18, 34).
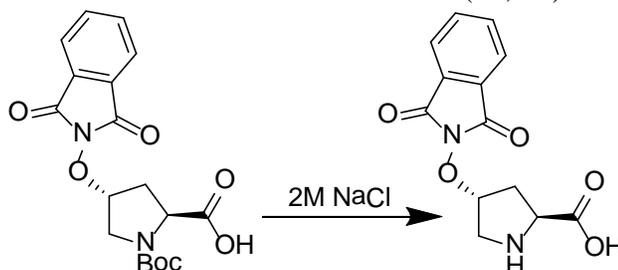
Figure 27:   Synthesis of N-2S,4R-Hyp(O-phthalimide) from N-Boc-2S,4R-Hyp(O-phthalimide).

To N-2S-4R-hyp-(O-phthalimide) dissolved in H$_2$O (10 mL) was added NaHCO$_3$ (1.9 g, 22.9 mmol). The solution was cooled to 0 ℃ and a solution of Fmoc-OSu (5.8 g, 17.2 mmol) in 1,4-dioxane (10 mL) was added. The reaction mixture was stirred vigorously at 0 ℃ for 1 h, after which the ice bath was removed and the

solution stirred at room temperature for overnight. Water (10 mL) was then added to the reaction mixture and the aqueous layer was extracted with ethyl acetate (2×). The organic layers were back-extracted twice with a saturated sodium bicarbonate solution. The combined aqueous layer was acidified to a pH of 1 with 2N HCl, then extracted with ethyl acetate (2×). The combined organic layers were dried with sodium sulfate and concentrated in *vacuo*. The resulting white residue was purified by silica gel column chromatography (0-5% CH_3OH/CH_2Cl_2 v/v) to obtain compound N-Fmoc-2S,4R-Hyp-(O-phthalimide) (1.51g) as a white solid. ^{1}H NMR (400 MHz, MeOD) δ 7.88-7.85 (m, 4H), 7.72-7.20 (m, 8H), 4.70 (m, 2H), 4.43-4.40 (t, 1H), 4.22 (t, 1H), 3.55-3.44 (m, 1H), and 2.08-1.99 (m, 2H). The NMR data corresponded to literature values (Figure 29) (18, 34).
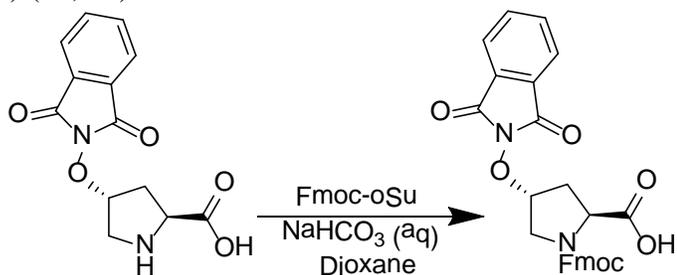
Figure 28: Synthesis of N-Fmoc-2S,4R-Hyp-(O-phthalimide) from N-2S-4R-hyp-(O-phthalimide).

34
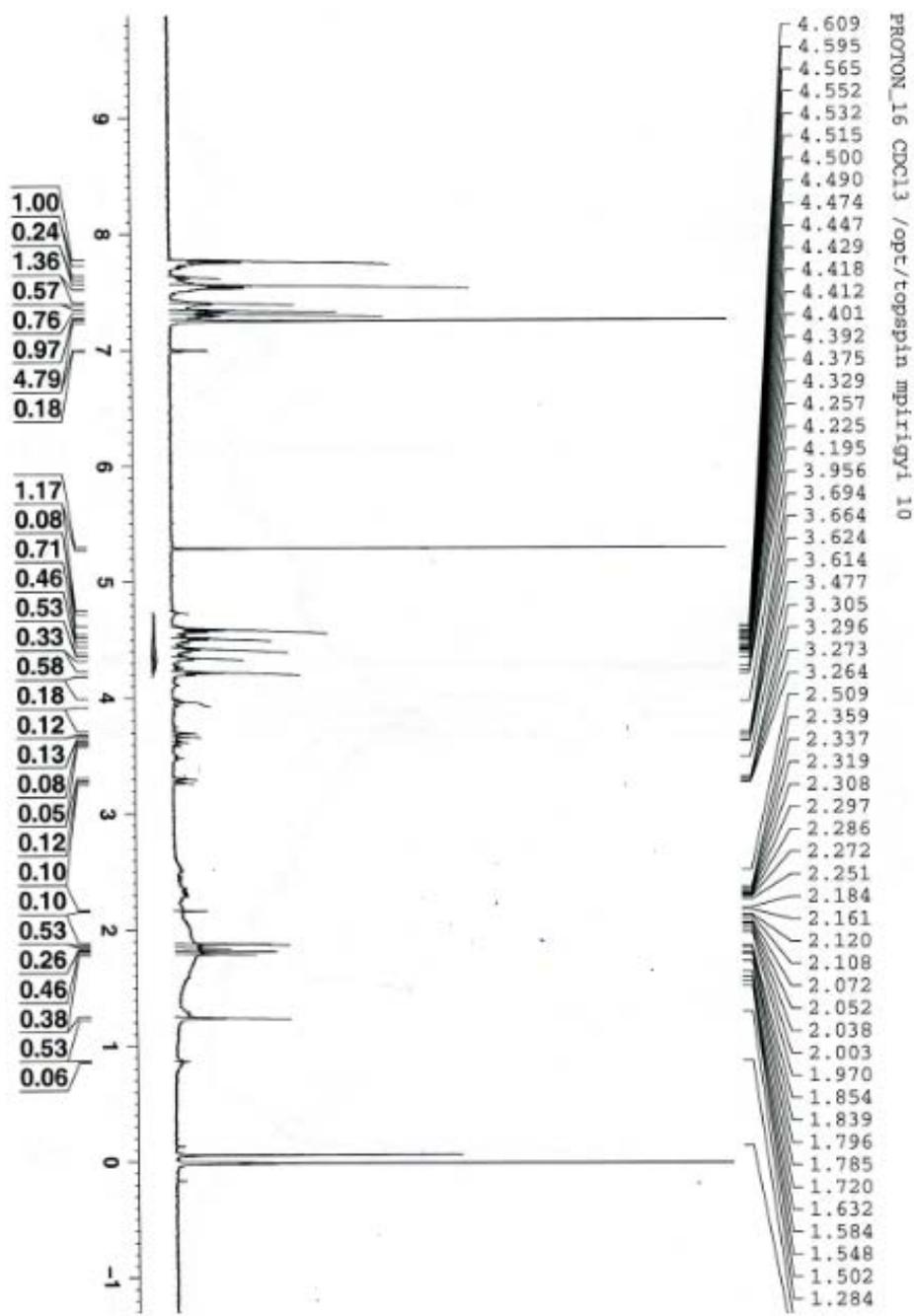
Figure 29: ¹H NMR of Fmoc-protected Hyp-(4R-O-phthalimide).

2.4   C-terminal Modifications of Model Peptide

In order to optimize the reaction conditions for solution phase reactions on cysteine and check for cross reactivity, the modification reactions were initially conducted on the model peptide Ac-KKHMCX-NH$_2$, where X is 4-iodophenylalanine (19).

2.4.1   Fluorescein Labeling of Ac-KKHMCX-NH$_2$

The model peptide was labeled with fluorescein at the C-terminal cysteine residue. 5-iodoacetamidofluorescein (5-IAF) (0.5 mg/mL) was allowed to react with the peptide in pH 6 phosphate buffer for approximately 24 hours (Figure 30, 31).
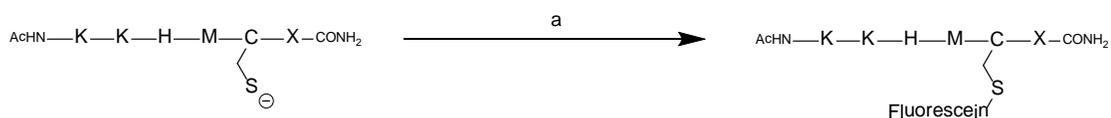
Figure 30:   Scheme for the C-terminal cysteine modification of the model peptide with 5-iodoacetamidofluorescein. (a) 5-IAF, pH 6 phosphate buffer.
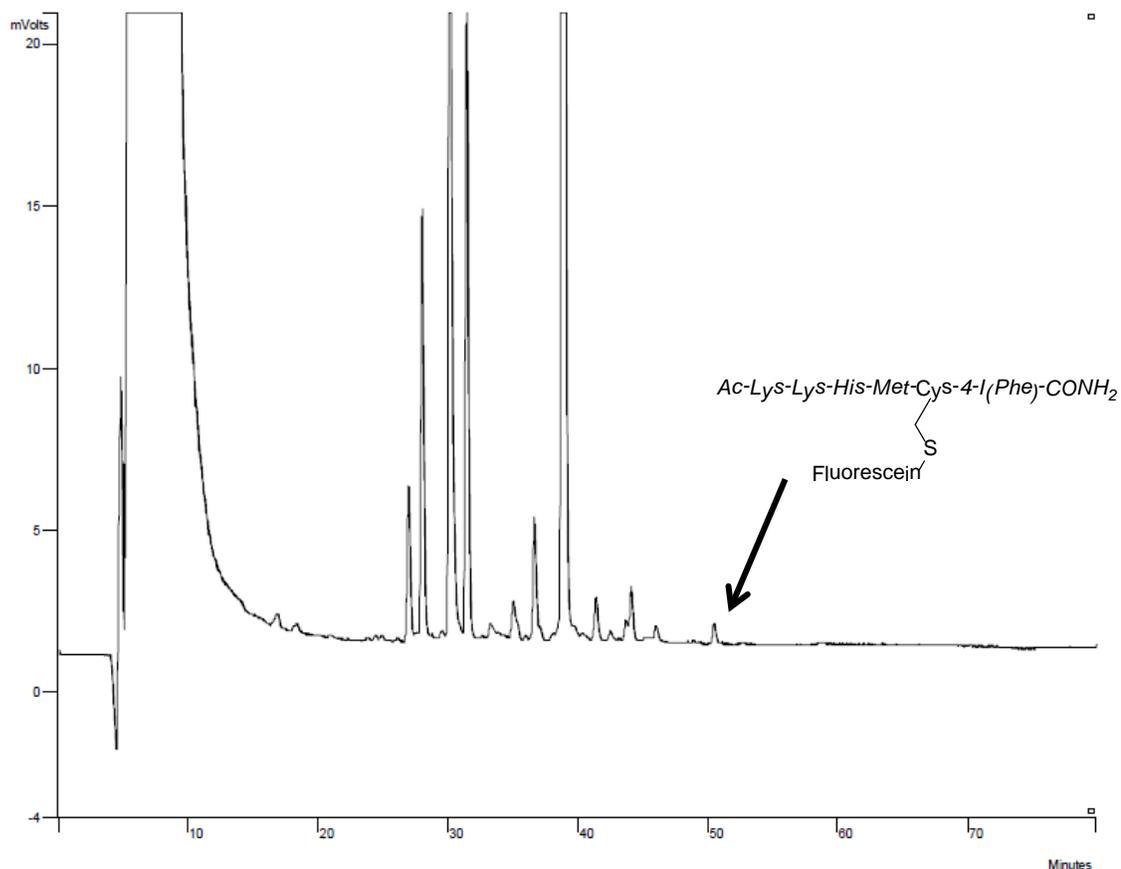
Figure 31: Crude HPLC chromatogram for fluorescein labeling of AC-KKHMCX-NH$_2$. The peak at t_R= 52 corresponds to the product. The other peaks on the chromatogram correspond to 5-IAF and impurities present in the commercial material.

2.4.2   Modification of Ac-KKHMCX-NH$_2$ with an aryl halide on Cysteine

The 4-bromobenzyl modified model peptide was synthesized using 4-bromo benzyl bromide (250 µM) and the model peptide (50 µM) in pH 7.12 phosphate buffer (50 mM) containing 250 mM KF at room temperature. The reaction was conducted in solution phase and was allowed to react for one hour with stirring (Figure 32, 33).

Figure 32: Scheme for C-terminal cysteine modification of the model peptide with 4-bromo benzyl bromide. (a) 4-bromobenzyl bromide, pH 7.1 phosphate buffer.
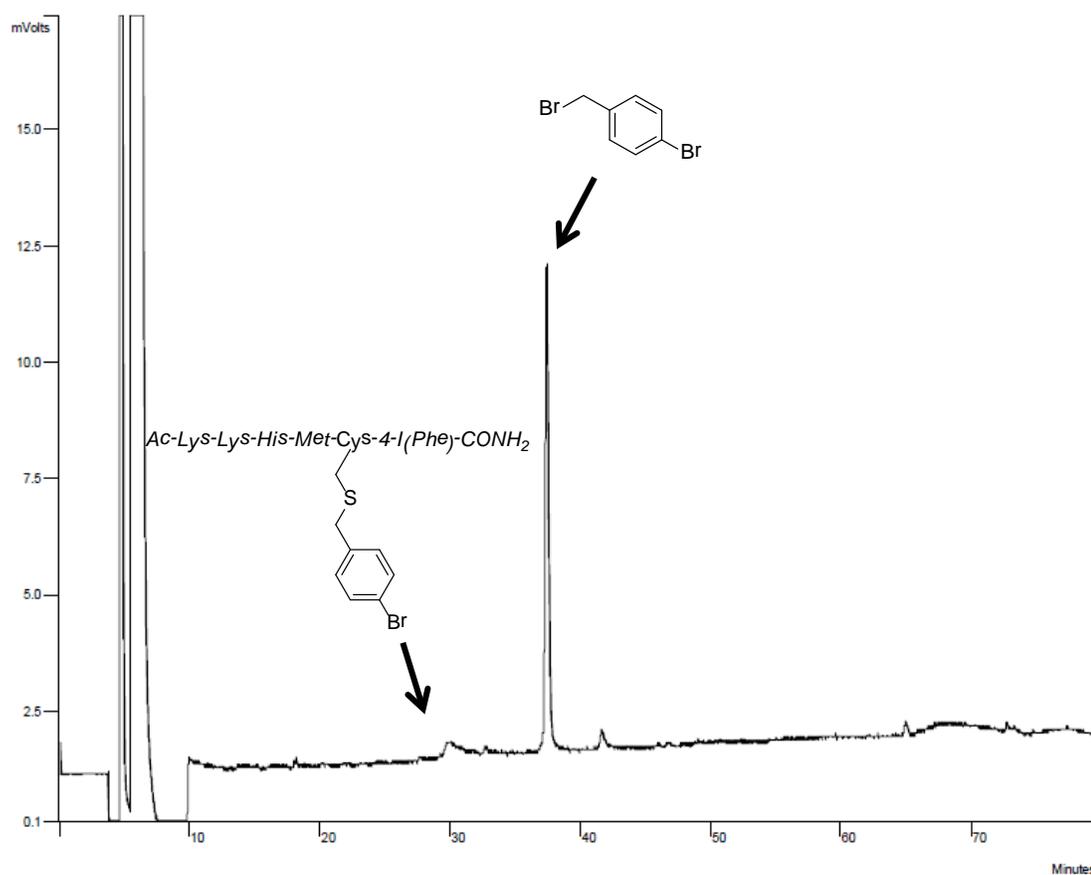
Figure 33: Crude HPLC chromatogram after the model peptide was modified with 4-bromobenzyl bromide. The larger peak at t_R= 34 min represents the 4-bromo benzyl bromide reagent, and the small peak represents the product with a t_R= 29.3 min.

2.4.3 Modification of Ac-KKHMCX-NH$_2$ with an aldehyde on Cysteine

The model peptide with the aldehyde modification was synthesized using acrolein (4 mM), model peptide, and pH 7.1 phosphate buffer (50 mM) containing 250 mM KF at room temperature. The reaction was conducted in solution phase and was allowed to react for one hour with stirring (Figure 34). Excellent conversion was observed as all of the starting material was consumed as demonstrated from the crude HPLC chromatogram (Figure 35).
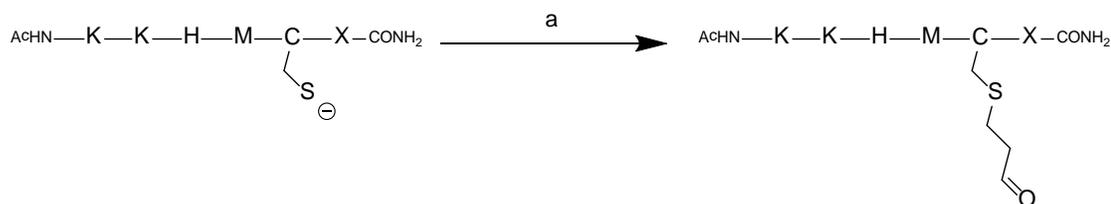
Figure 34: Scheme for C-terminal cysteine modification of the model peptide to incorporate an aldehyde group. (a) acrolein, pH 7.6 phosphate buffer.
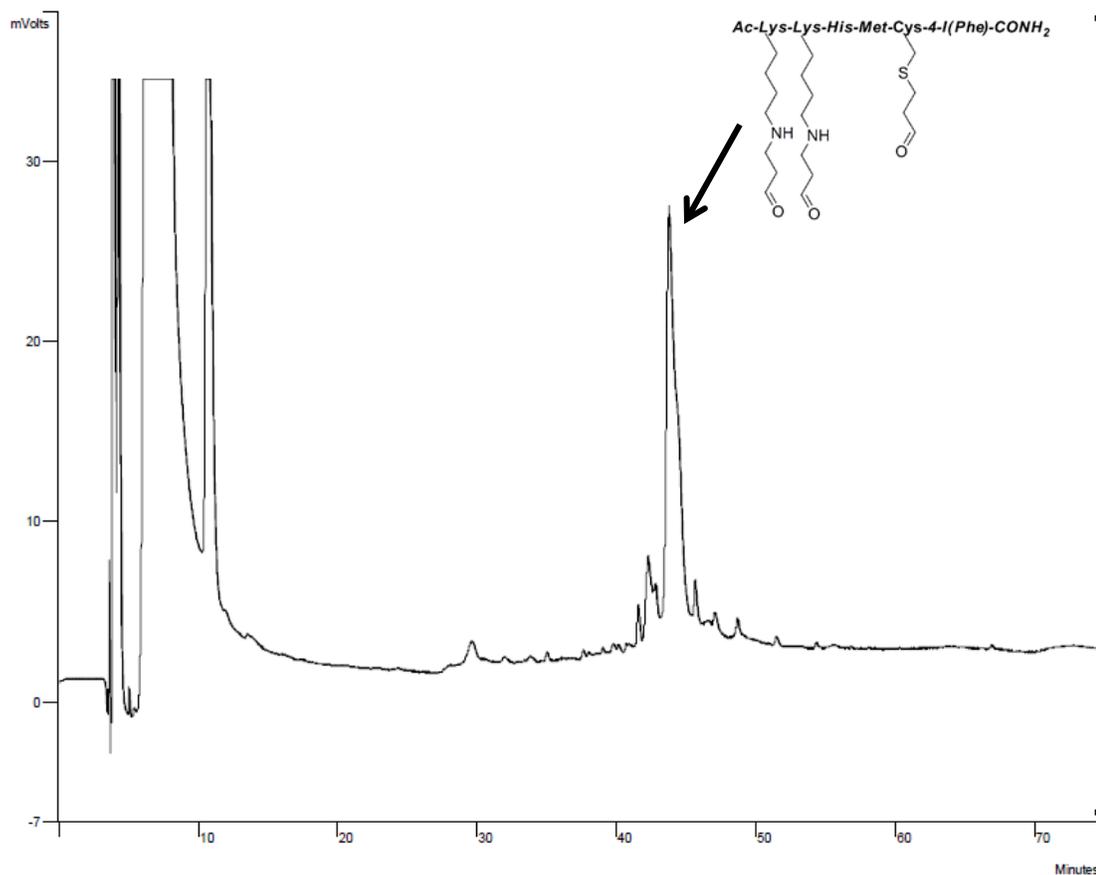
Figure 35: Crude HPLC chromatogram after the model peptide was modified with acrolein. The peak at t_R= 43 min represents the product.

2.4.4 Modification of Ac-KKHMCX-NH$_2$ with an alkyne on Cysteine

The alkyne group was incorporated via reaction with propargyl bromide (2 mM) and the model peptide in pH 7.1 phosphate buffer (50 mM) containing 250 mM KF at room temperature. The reaction was done in solution phase and was allowed to react for one hour with mixing (Figure 36). Excellent conversion to the product was observed (Figure 37).
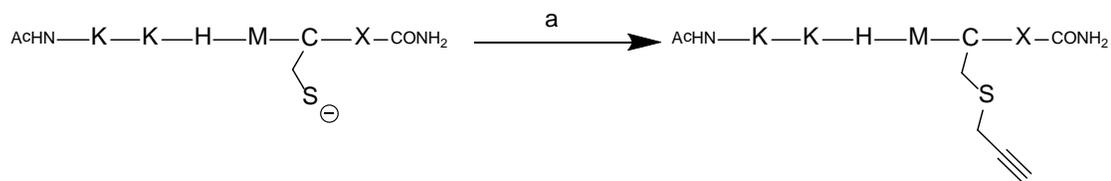
Figure 36: Scheme for C-terminal cysteine modification of model peptide with propargyl bromide. (a) Propargyl bromide, pH 7.1 phosphate buffer.
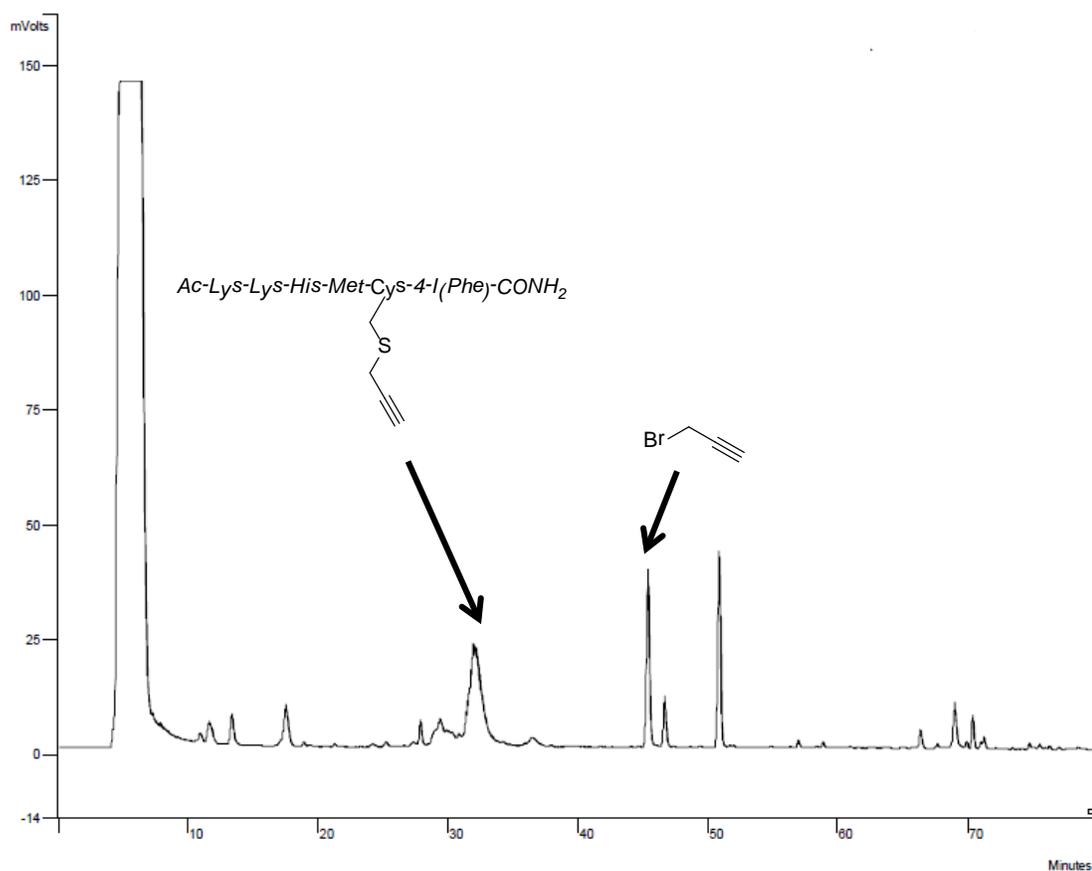
Figure 37: Crude HPLC chromatogram after the model peptide was modified with propargyl bromide. The peak at t_R=32 min represents the product and the large peak at t_R=45 min represents the propargyl bromide reagent.

2.5   C-terminal Modifications of TBD Peptides

2.5.1   tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) Cys-(fluorescein)

The cysteine of tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) peptide was modified with 5-IAF (Figure 38).

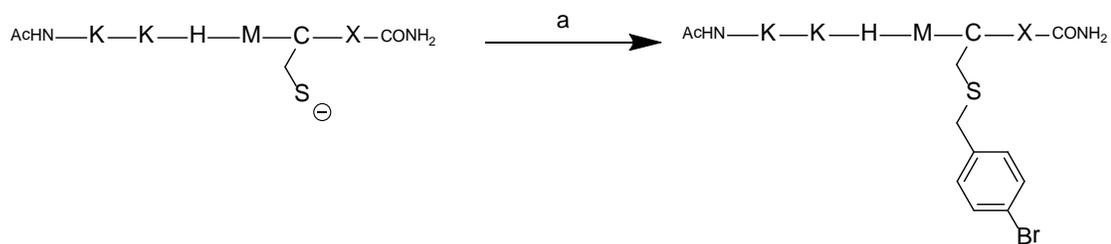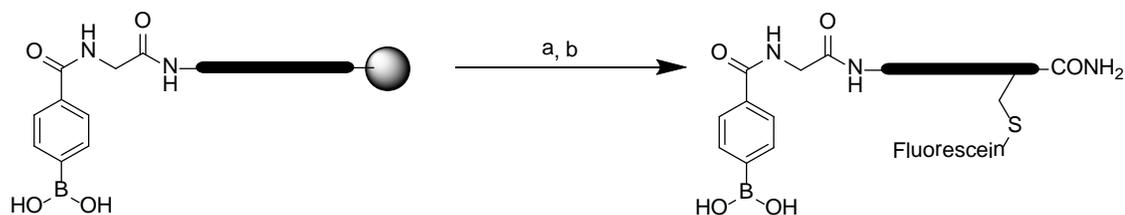Figure 38:   Scheme for the C-terminal cysteine modification of tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) peptide with 5-iodoacetamidofluorescein. (a) Cleavage with Reagent K in TFA; (b) 5-IAF, pH 6 phosphate buffer.

2.5.2   tau$_{334\text{-}365}$ N-(N$_3$) Cys-(aryl halide)

The cysteine of the tau$_{334\text{-}365}$ N-(N$_3$) peptide was modified with 4-bromo benzyl bromide. The doubly modified peptide was synthesized by reacting the tau$_{334\text{-}365}$ N-(N$_3$) peptide with 4-bromo benzyl bromide (250 μM) and DTT (100 μM) in pH 7.6 phosphate buffer (100 mM phosphate) containing 200 mM NaCl. The reaction was conducted for 2 hours with mixing (Figure 39). The reaction proceeded with excellent conversion (Figure 40).
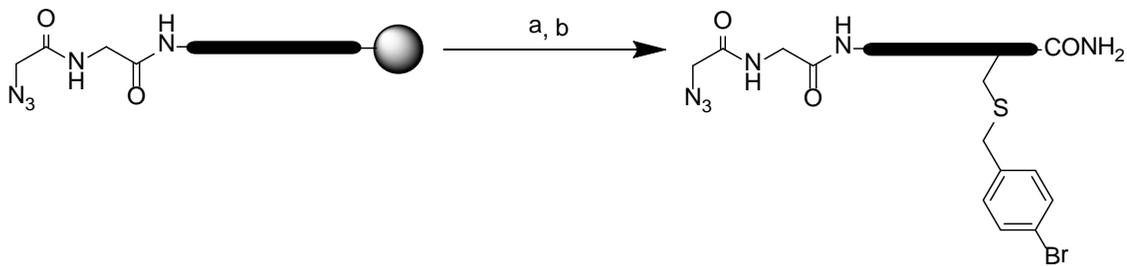
Figure 39: Scheme for the C-terminal cysteine modification of the $tau_{334\text{-}365}$ N-(N_3) peptide with 4-bromobenzyl bromide. (a) Cleavage with Reagent K in TFA (3 hr); (b) 4-bromobenzyl bromide, DTT, pH 7.6 phosphate buffer (1 hr).
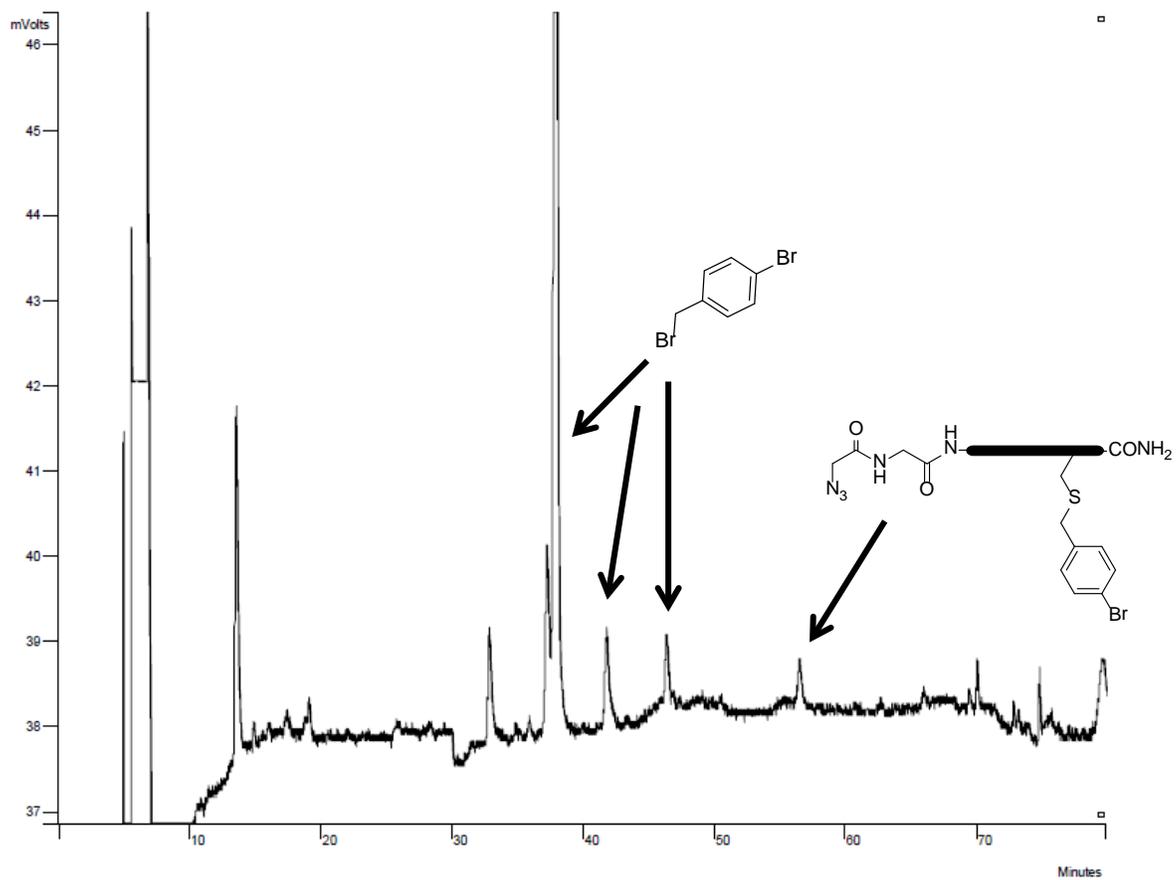
Figure 40: Crude HPLC chromatogram of the solution phase modification of the azide $\text{tau}_{334\text{-}365}$ peptide with 4-bromo benzyl bromide. The peak at $t_R = 56.4$ min represents the product. Peaks at $t_R = 37.9$, 41.5, and 46.1 min represent the propargyl bromide starting material.

2.5.3 $\text{tau}_{334\text{-}365}$ N-(Ac) Cys-(aldehyde)

The cysteine of the acetylated $\text{tau}_{334\text{-}365}$ peptide was modified with acrolein. The doubly modified peptide was synthesized by reacting the acetylated $\text{tau}_{334\text{-}365}$ peptide with acrolein (10 mM) in pH 7.6 phosphate buffer (100 mM) containing 200 mM NaCl, TCEP (0.2 mM), and guanidinium (0.05 mM). The reaction was conducted for 10 min with stirring at room temperature (Figure 41). The reaction proceeded with excellent conversion (Figure 42).
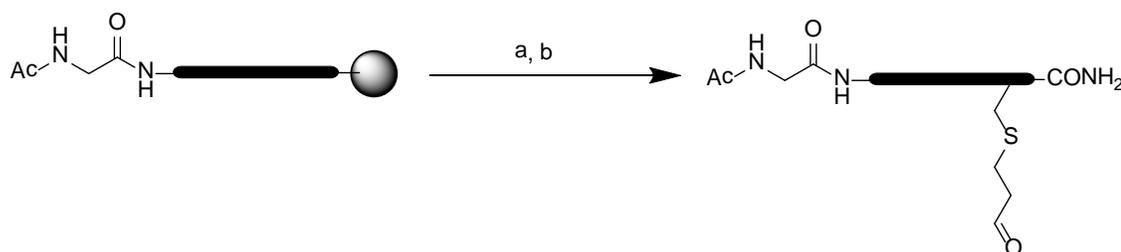
Figure 41: Scheme for the C-terminal cysteine modification of the acetylated $\text{tau}_{334\text{-}365}$ peptide with acrolein. (a) Cleavage with Reagent K in TFA (3 hr); (b) Acrolein, TCEP, guanidinium, pH 7.6 phosphate buffer (10 min).
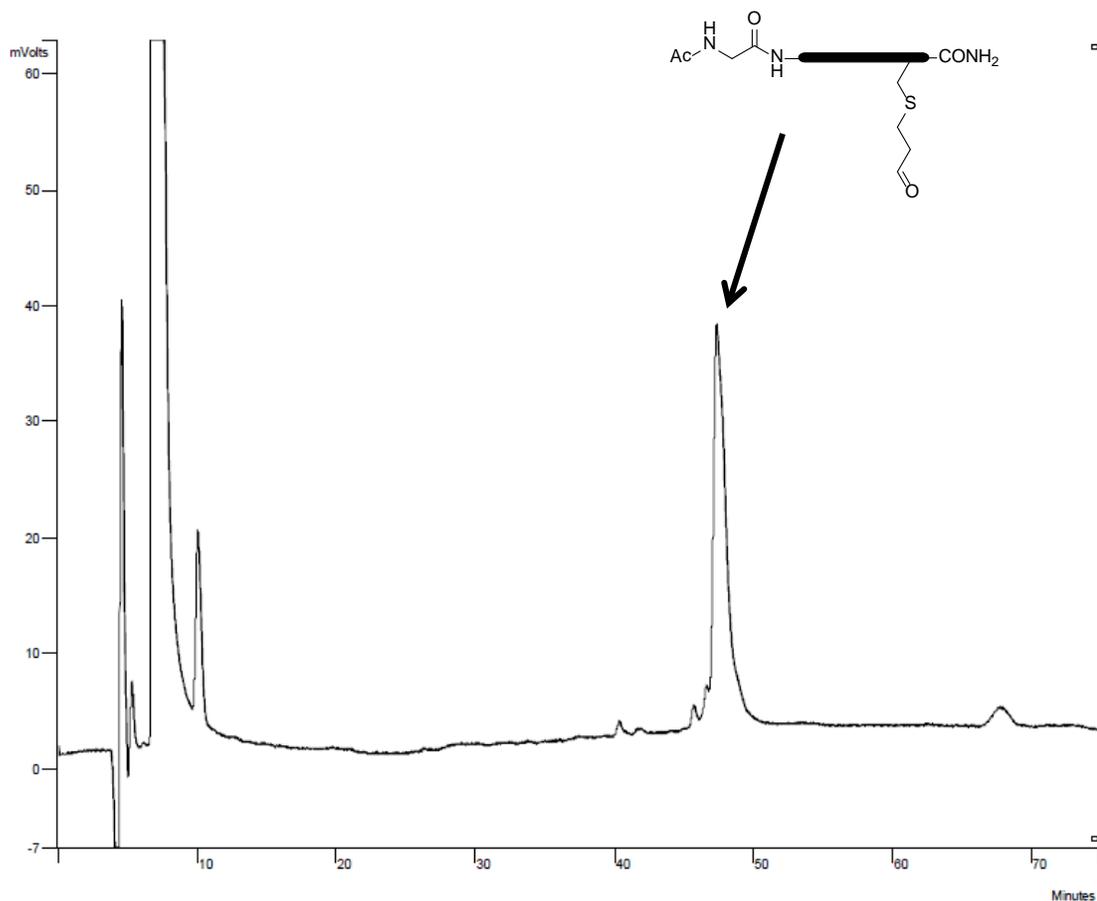
Figure 42:   Crude HPLC chromatogram of $\text{tau}_{334\text{-}365}$ N-(Ac) Cys-(Aldehyde). The large peak at t_R=47.3 corresponds to the product.

2.5.4 $\text{tau}_{334\text{-}365}$ N-(ONH$_2$) Cys-(alkyne)

The C-terminal cysteine of the $\text{tau}_{334\text{-}365}$ N-(O-phthalimide) peptide was modified using propargyl bromide. Tau$_{334\text{-}365}$ N-(O-phthalimide) was cleaved from the resin using Reagent K in TFA, purified via HPLC, and dried down. It was dissolved in pH 7.6 phosphate buffer (100 mM phosphate) containing 200 mM NaCl and allowed to react with TCEP (0.2 mM) for 15 minutes to reduce any disulfide bonds. Propargyl bromide (10 mM) and guanidinium (0.05 mM) were added and the peptide allowed to

react for 1 hour with stirring at room temperature to yield the tau$_{334\text{-}365}$ N-(O-phthalimide) Cys-(alkyne) peptide. (Figure 43, 44). The peptide was purified via HPLC. The deprotection of the phthalimide protecting group on peptide tau$_{334\text{-}365}$ N-(O-phthalimide) Cys-(alkyne) was conducted by reaction with hydrazine hydrate (4 mM) in pH 7.6 phosphate buffer (50 mM) containing 100 mM NaCl. The reaction was conducted for 1 hour with stirring to yield peptide tau$_{334\text{-}365}$ N-(ONH$_2$) Cys-(alkyne) (Figure 45).
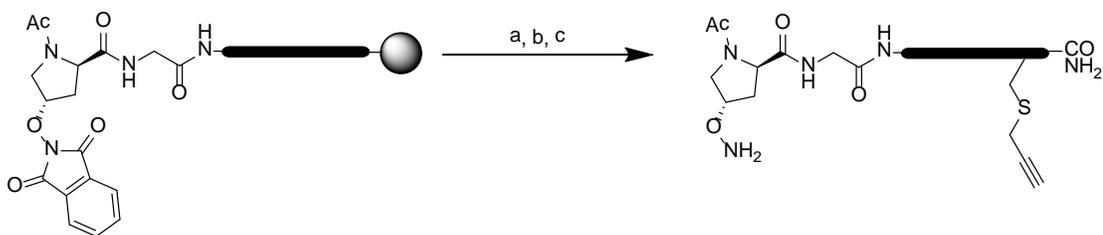
Figure 43:   Scheme for the C-terminal cysteine modification of the tau$_{334\text{-}365}$ N-(O-phthalimide) peptide with propargyl bromide. (a) Cleavage with Reagent K in TFA (3 hr); (b) Propargyl bromide, TCEP, guanidinium, pH 7.6 phosphate buffer; (c) NH$_2$NH$_2$•H$_2$O, pH 7.6 phosphate buffer.
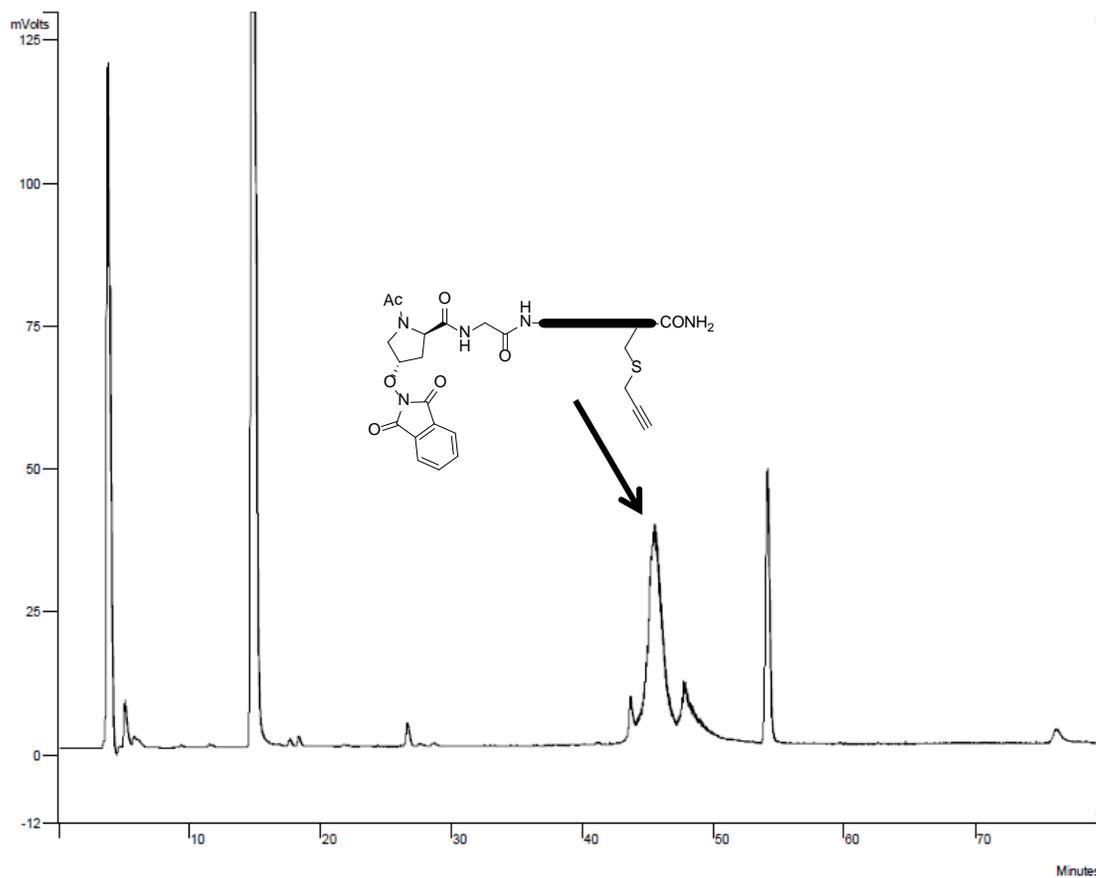
Figure 44:   Crude HPLC chromatogram of tau$_{334-365}$ N-(O-phthalimide) Cys-(alkyne).
The peak with t_R= 45.7 min corresponds to the product.
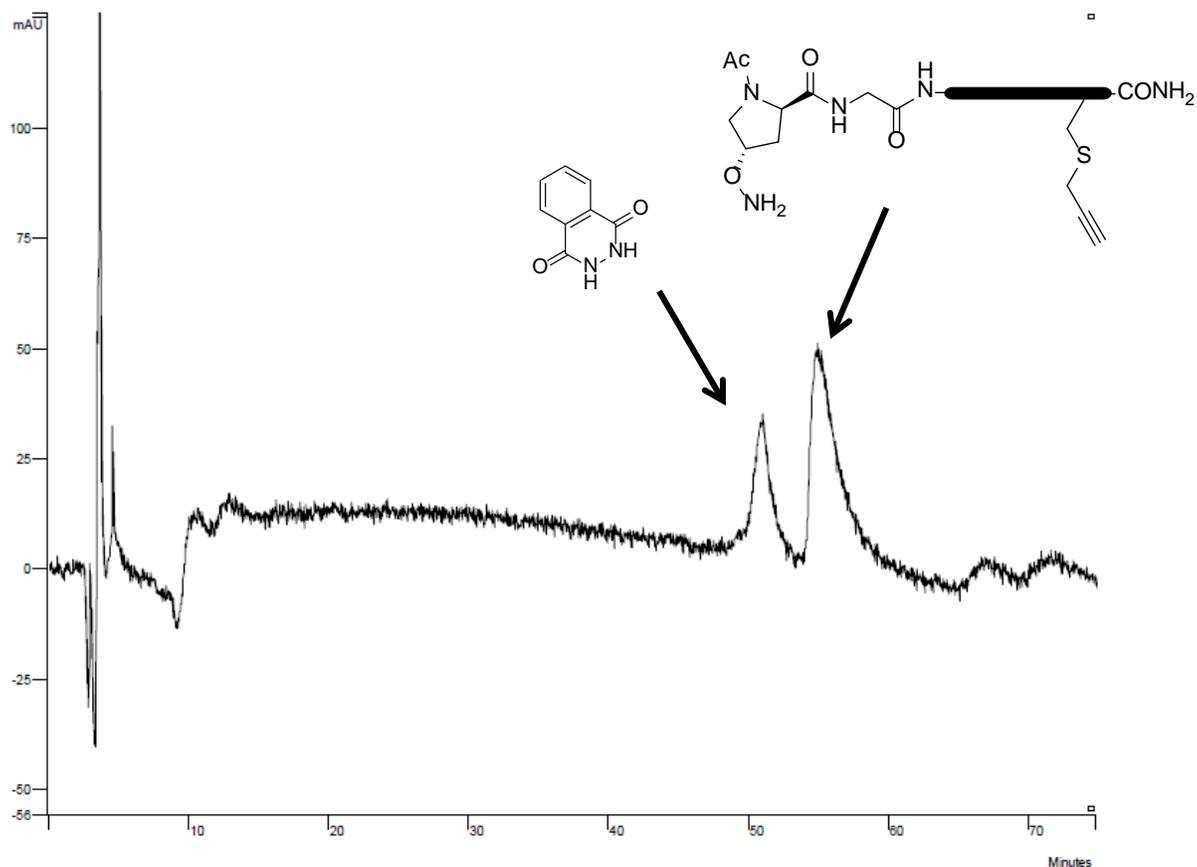
Figure 45:   Crude HPLC chromatogram of tau$_{334-365}$ N-(ONH$_2$) Cys-(alkyne). The large peak at t_R= 55 min corresponds to the product and the smaller peak at t_R= 51 min may correspond to the removed phthalimide group.

2.6 Analytical Data

Table 1:   Peptides synthesized in this study with their corresponding sequences and modifications. Peptide sequences are based on the largest isoform of tau (441 amino acids). X corresponds to 4-iodo-Phenylalanine.

	Peptide	Modification
1	tau$_{334-365}$	Unmodified, Fmoc-protected
2	tau$_{334-365}$	N-(COPhB(OH)$_2$)

3	tau$_{334-365}$	N-(N$_3$)
4	tau$_{334-365}$	N-(Acetyl)
5	tau$_{334-365}$	N-(O-phthalimide)
6	Ac-KKHMCX-NH$_2$	Unmodified
7	Ac-KKHMCX-NH$_2$	Cys-(Fluorescein)
8	Ac-KKHMCX-NH$_2$	Cys-(Aryl halide)
9	Ac-KKHMCX-NH$_2$	Cys-(Aldehyde)
10	Ac-KKHMCX-NH$_2$	Cys-(Alkyne)
11	tau$_{334-365}$	N-(COPhB(OH)$_2$) Cys-(Fluorescein)
12	tau$_{334-365}$	N-(N$_3$) Cys-(Aryl halide)
13	tau$_{334-365}$	N-(Ac) Cys-(Aldehyde)
14	tau$_{334-365}$	N-(O-phthalimide) Cys-(Alkyne)
15	tau$_{334-365}$	N-(ONH$_2$) Cys-(Alkyne)

2.6.1 Unmodified tau$_{334-365}$ (1)

Peptide **1** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 51 min, exp. 3818.9, obs. 1270 (M-3)$^{3-}$.

2.6.2 tau$_{334-365}$ N-(COPhB(OH)$_2$) (2)

Peptide **2** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 48.5 min, exp. 3745.8.

2.6.3 tau$_{334-365}$ N-(N$_3$) (3)

Peptide **3** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 44 min, exp. 3679.8, obs. 1222.9 (M-3)$^{3-}$.

2.6.4   tau$_{334-365}$ N-(Ac) (4)

Peptide **4** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 39.4 min, exp. 3639, obs. 1209.2 (M-3)$^{3-}$.

2.6.5   tau$_{334-365}$ N-(O-phthalimide) (5)

Peptide **5** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 41.7 min, exp. 3898, obs. 838 (M+Na$^+$)

2.6.6   Unmodified Ac-KKHMCX-NH$_2$ (6)

Peptide **6** was purified via semi-prep HPLC using a linear gradient of 0-70% buffer B in buffer A over 60 minutes: t_R= 26 min, exp. 959.9, obs. 959.7 (M).

2.6.7   Ac-KKHMCX-NH$_2$ Cys-(Fluorescein) (7)

Peptide **7** was purified via analytical HPLC using a linear gradient of 0-70% buffer B in buffer A over 60 minutes: t_R= 52 min, exp. 1333.

2.6.8   Ac-KKHMCX-NH$_2$ Cys-(Aryl halide) (8)

Peptide **8** was purified via analytical HPLC using a linear gradient of 0-70% buffer B in buffer A over 60 minutes: t_R= 29.3 min, exp. 998, obs. 998.3 (M).

2.6.9   Ac-KKHMCX-NH$_2$ Cys-(Aldehyde) (9)

Peptide **9** was purified via analytical HPLC using a linear gradient of 20 min iso A then 0-50% buffer B in buffer A over 60 minutes: t_R= 43 min, exp. 1016.

2.6.10 Ac-KKHMCX-NH$_2$ Cys-(Alkyne) (10)

Peptide **10** was purified via analytical HPLC using a linear gradient of 0-70% buffer B in buffer A over 60 minutes: t_R= 32 min, exp. 1129.

2.6.11 tau$_{334\text{-}365}$ N-(COPhB(OH)$_2$) Cys-(Fluorescein) (11)

2.6.12 tau$_{334\text{-}365}$ N-(N$_3$) Cys-(Aryl halide) (12)

Peptide **12** was purified via semi-prep HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 56.4 min, exp. 3822.

2.6.13 tau$_{334\text{-}365}$ N-(Ac) Cys-(Aldehyde) (13)

Peptide **13** was purified via semi-prep HPLC using a linear gradient of 20 min iso A then 0-50% buffer B in buffer A over 60 minutes: t_R= 47.3 min, exp. 3696, obs. 1228.4 $(M+3)^{3+}$.

2.6.14 tau$_{334\text{-}365}$ N-(O-phthalimide) Cys-(Alkyne) (14)

Peptide **14** was purified via analytical HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 45.7 min, exp. 3936.

2.6.15 tau$_{334\text{-}365}$ N-(ONH$_2$) Cys-(Alkyne) (15)

Peptide **15** was purified via analytical HPLC using a linear gradient of 0-50% buffer B in buffer A over 60 minutes: t_R= 55 min, exp. 3806, obs. 1266 $(M-3)^{3-}$.

2.7 SDS-PAGE Conditions

Sodium dodecyl sulfate (SDS) polyacrylamide gels were made to analyze the completion of the bioorthogonal ligation reactions. The Tris-HCl buffer, SDS solution, bis-acrylamide solution, and deionized water mixture were degassed. TEMED and APS were added to polymerize the gel. The conditions for the gels are listed below:

Table 2:   SDS polyacrylamide gel formulations.

Gel	Gel %	Tris-HCl Buffer (M/pH)	Tris-HCl Buffer (mL)	SDS (10%) (mL)	Deionized H_2O (mL)	Bis-Acrylamide (30%) (mL)	TEMED	APS (10%) (mL)
Resolving	15	1.5/8.8	2.5	0.1	6.1	1.3	5	50
Stacking	4	0.5/6.8	2.5	0.1	2.4	5.0	10	50

Gels were run at 200 mV in a BIO-RAD Mini-PROTEAN tetra system in pH 8.3 10 running buffer (0.1% SDS, 192 mM glycine, and 25 mM Tris in H_2O). After electrophoresis the gels were rinsed in deionized water (3), fixed for 30 minutes (40% MeOH, 10% Acetic Acid in H_2O), and stained with Coomassie Blue stain for 30 min. The gels were then destained for 1 hour.

Chapter 3

RESULTS AND DISCUSSION

The fourth repeat of the tubulin-binding domain of tau was synthesized on solid phase with a C-terminal cysteine to allow for the incorporation of bioorthogonal moieties on the N-terminus and C-terminal cysteine. Four bioorthogonal $tau_{334-365}$ peptides were synthesized to allow for ligations using oxime chemistry, click chemistry, and Suzuki-Miyaura chemistry. All peptides were purified to homogeneity using HPLC before ligation reactions were performed. Model peptides were used to examine the solution phase cysteine modifications for cross reactivity.

3.1 $Tau_{334-365}$ Peptide

Synthesizing the four bioorthogonal $tau_{334-365}$ peptides was challenging, with some difficulty due to aggregation and solubility issues. Poor yields resulted from solid phase N-terminal modifications because of poor aqueous solubility of the peptides. Multiple solvents were used to try and dissolve the peptide, such as: pH 3 phosphate buffer with 5% DMF, pH 3 phosphate buffer with Acetonitrile, pH 3 phosphate buffer with 5% DMSO, 0.5 M pH 4 Tris HCl buffer, pH 3 phosphate buffer with urea, 1 M guanidinium buffer with 5% DMF, and 1 M pH 7 phosphate buffer with 5% DMF.

The solubility problems were resolved by cleaving the tau peptides in a lower concentration and higher volume. The peptides were then dissolved and purified as described in the methods section.

3.2 Hyp-(4*R*-O-phthalimide) Synthesis

Initially the Hyp-(4*R*-O-phthalimide) was synthesized using protocol from a paper published by the Burke group (33). The reported synthetic protocol used a Cbz protecting group on the amine and a benzyl protecting group on the carboxylic acid of hydroxy proline. Poor reaction yields were observed for the Hyp-(4*R*-O-phthalimide) after the Mitsonobu reaction with the major product being the DIAD side product. Further analysis showed that when the proline ring is in the endo conformation (4S) the benzyl protecting group could potentially be sterically blocking the hydroxyl group, not allowing for proper conversion to the O-phthalimide derivative (Figure 46). Using a less sterically demanding methyl protecting group on the carboxylic acid led to better conversion to the desired product.
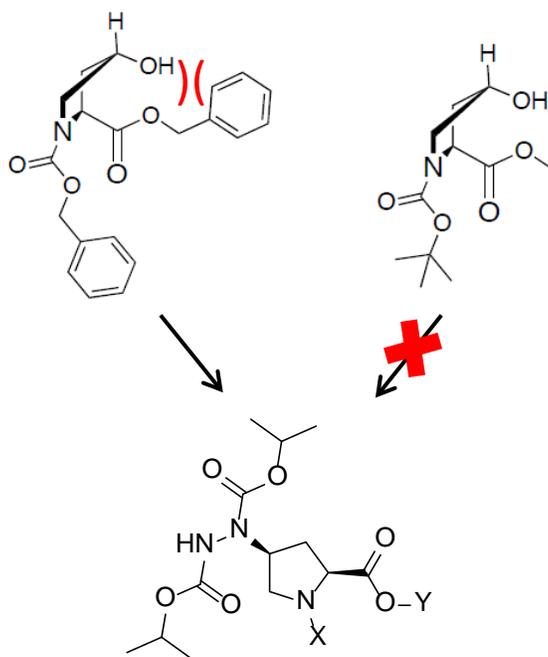
Figure 46:   Steric hindrance between the benzyl protecting group on the carboxylic acid and the hydroxyl nucleophile for the Mitsonobu reaction led to more DIAD side product. Changing the protecting group on the carboxylic acid to the smaller methyl group opened up the nucleophillic hydroxyl group and resulted in less side DIAD side product.

3.3   Ac-KKHMCX-NH$_2$ Model Peptide: Cross Reactivity

The KKHMCX model peptide was first modified with bromoacetaldehyde diethylacetal to incorporate an aldehyde onto the C-terminal cysteine. The reaction proceeded sluggishly and was replaced with acrolein to check for cross reactivity. Upon purification and analysis with MS it was shown that there was cross reactivity with the lysines in the model peptide. According to the MS the aldehyde moiety reacted with both the lysines and the cysteine in the peptide. This cross reactivity was resolved by running the reaction for a shorter period of time.

3.4   Bioorthogonal Ligation Reactions with tau$_{334-365}$ peptides

3.4.1   Oxime Reaction

The tau$_{334-365}$ N-(Ac) Cys-(Aldehyde) and tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne) peptides were ligated together to form an oxime (Figure 47). The reaction was run in pH 7.6 buffer with roughly equivalent concentrations of tau$_{334-365}$ N-(Ac) Cys-(Aldehyde) and tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne) for 30 minutes at 37 °C with mixing (Figure38). The reaction was purified via HPLC (Figure 48) and analyzed by SDS-PAGE (Figure 49). The 15% acrylamide gel was not capable of separating the starting material peaks from the final dipeptide product.
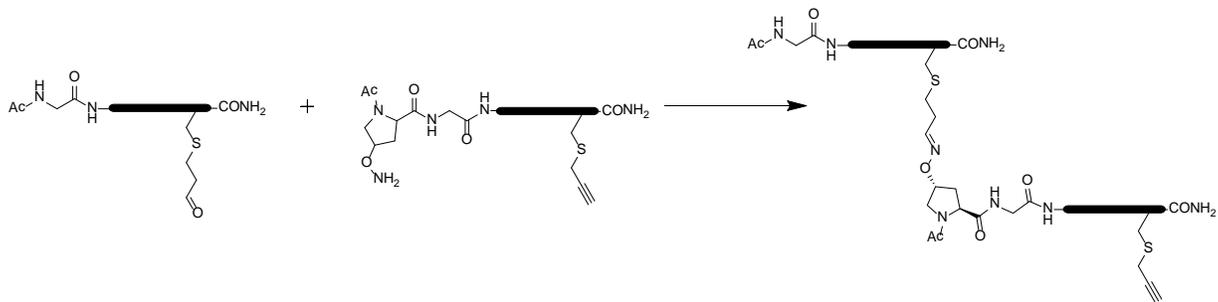
Figure 47: Oxime ligation between tau$_{334-365}$ N-(Ac) Cys-(Aldehyde) and tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne).
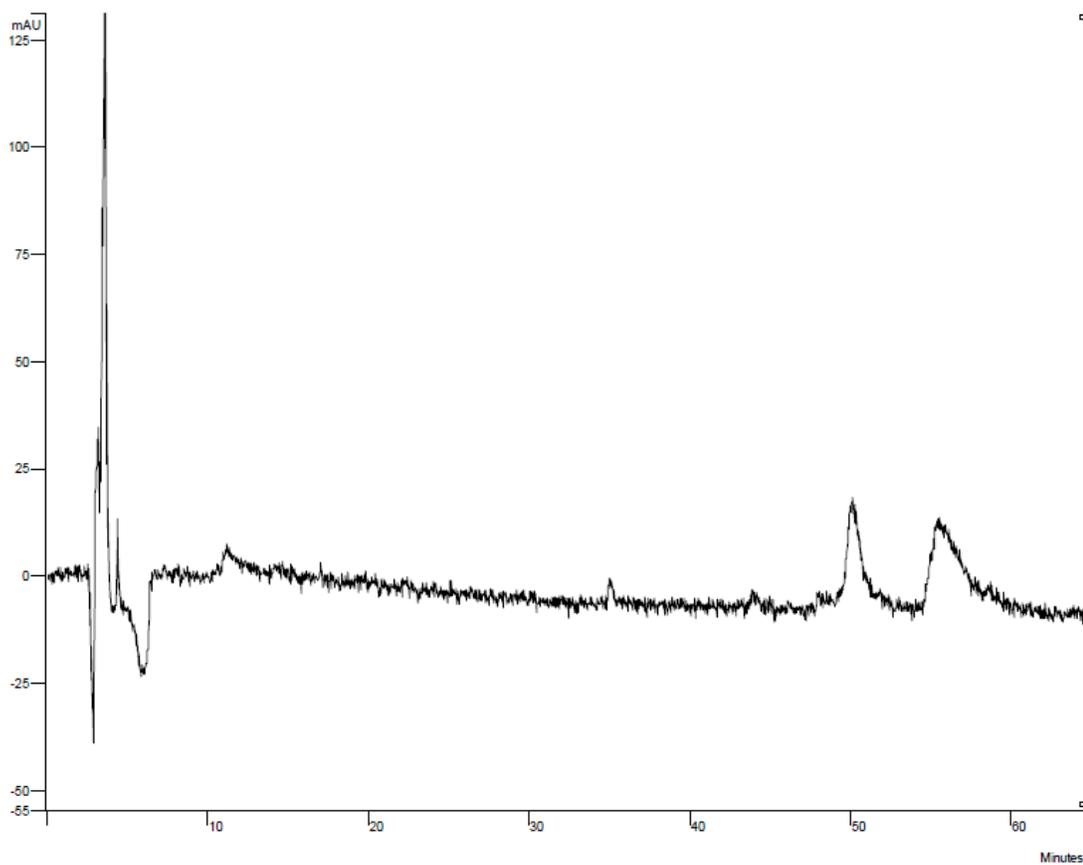
Figure 48: Crude HPLC chromatogram of oxime ligation between tau$_{334-365}$ N-(Ac) Cys-(Aldehyde) and tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne).
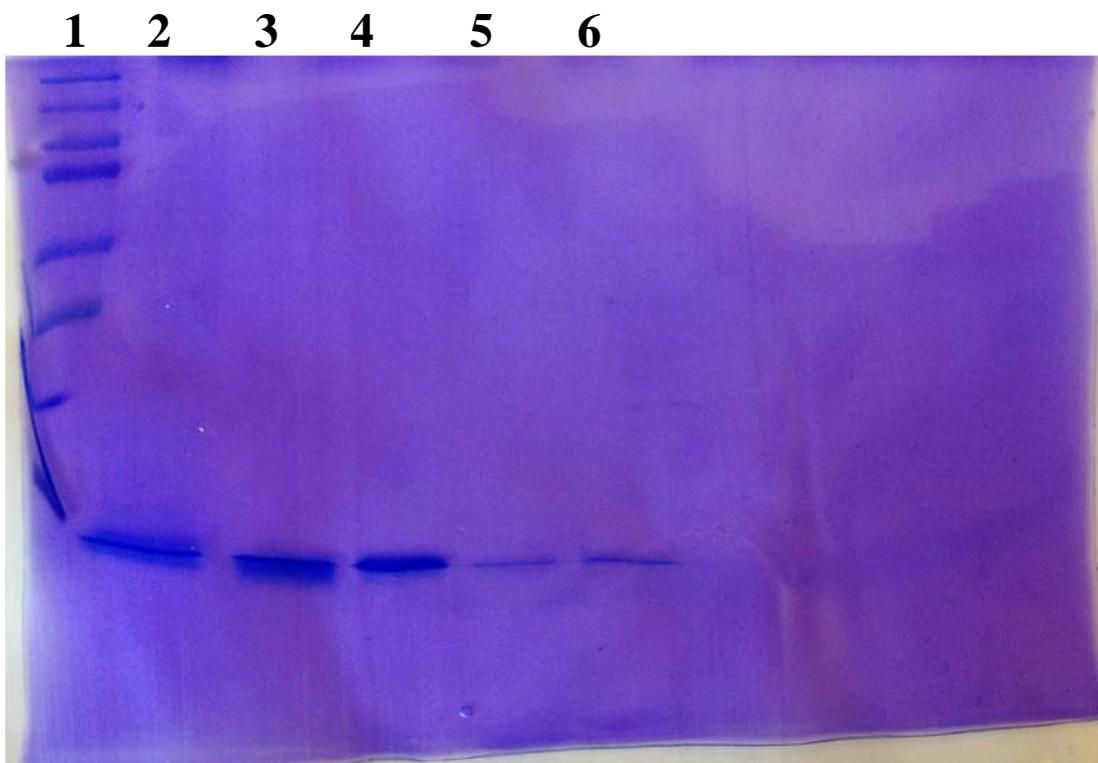
Figure 49:   SDS-PAGE of oxime ligation. Lane 1 contains the Kaleidscope ladder, lane 2 contains tau$_{334-365}$ N-(Ac) Cys-(Aldehyde), lane 3 contains tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne), lane 4 contains tau$_{334-365}$ N-(ONH$_2$) Cys-(Alkyne), lane 5 contains the oxime reaction, and lane 6 contains the oxime reaction.

3.4.2   Click Reaction

The tau$_{334-365}$ N-(ONH$_2$) Cys-(alkyne) and tau$_{334-365}$ N-(N$_3$) Cys-(aryl halide) peptides will be ligated together using click chemistry.
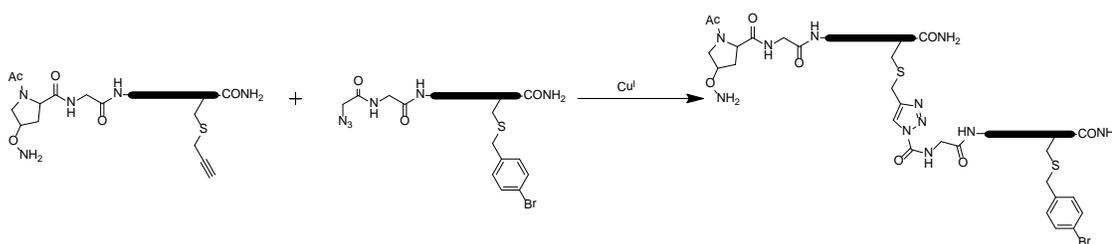
Figure 50: Ligation of tau$_{334-365}$ N-(ONH$_2$) Cys-(alkyne) and tau$_{334-365}$ N-(N$_3$) Cys-(aryl halide) peptides using click chemistry.

3.4.3 Suzuki-Miyaura Reaction

The tau$_{334-365}$ N-(B(OH)$_2$) C-(fluorescein) and tau$_{334-365}$ N-(N$_3$) C-(aryl halide) will be ligated together using Suzuki-Miyaura cross-coupling chemistry.

Figure 51: Ligation of tau$_{334-365}$ N-(B(OH)$_2$) Cys-(fluorescein) and tau$_{334-365}$ N-(N$_3$) Cys-(aryl halide) peptides using Suzuki-Miyaura chemistry.

Chapter 4

CONCLUSIONS

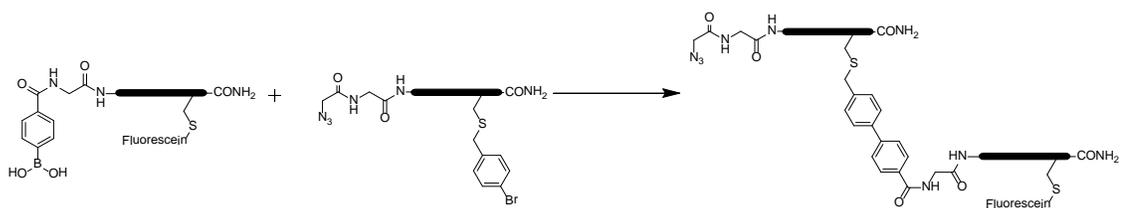A new way to synthesize selectively modified, functional proteins has been demonstrated using multiple bioorthogonal ligations. This method will be utilized to study the tubulin-binding domain of tau. Four bioorthogonal peptides were synthesized to allow for the synthesis of the tubulin-binding domain. The peptides were modified at the N-terminus as well as the C-terminal cysteine residues to incorporate bioorthogonal functionalities to allow for ligations via the oxime reaction, click reaction, and Suzuki-Miyaura reaction. Fmoc-$_L$-Hyp-(4R-O-phthalimide) was synthesized from *trans*-4-hdyroxy-$_L$-proline in an eight step synthesis. The proline derivative was used to incorporate an oxyamine into the tau peptide for use in the oxime reaction. Solution phase C-terminal modification reactions were checked for cross reactivity using model peptides. Fluorescent anisotropy will be used to analyze the synthetic tubulin-binding domain's ability to polymerize microtubules in the nonphosphorylated and phosphorylated state.

REFERENCES

1. Bennett, C.S. and Wong, C.H. *Chem. Soc. Rev*. **2007**, *36*, 1227-1238.

2. Broncel, M., Krause, E., Schwarzer, D., and Hackenberger, C.P.R. *Chem. Eur. J*. **2012**, *18,* 2488-2492.

3. Sletten, E.M. and Bertozzi, C.R. *Angew. Chem. Int. Ed*. **2009**, *48*, 6974-6998.

4. Lim, R.K.V. and Lin, Q. *Chem. Commun.* **2010**, *46*, 1589-1600.

5. Tarwade, V., Liu, X., Yan, N, and Fox, J.M. *J. Am. Chem. Soc*. **2009**, *131*, 5382-5383

6. *Alzheimer's Association*. Alz.org. (March 2013)

7. Buée, L., Bussiére, T., Buée-Scherrer, V., Delacourte, A., and Hof, P.R. *Brain Res. Brain Res. Rev*. **2000**, *33*, 95-130.

8. Du, J.T., Yu, C.H., Zhou, L.X., Wu, W.H., Lei, P., Li, Y., Zhao, Y.F., Naknishi, H., and Li, Y.M. *FEBS J*. **2007**, *274*, 5012-5020.

9. Mandelkow, E.M., Biernat, J., Drewes, G., Gustke, N., Trinczek, B., and Mandeklow, E. *Neurobiol. Aging*. **1995**, *16*, 335-363.

10. Litersky, J., Johnson, G.V., Jakes, R., Goedert, M., Lee, M., and Seubert, P. *Biochem. J*. **1996**, *316*, 655-660.

11. Wolfe, M.S. *J. Biol. Chem*. **2009**, *284*, 6021-6025.

12. Avila, J. *FEBS Letters.* **2006**, *580*, 2922-2927.

13. Mukrasch, M., Bergen, M., Biernat, J., Fischer, D., Griesinger, C., Mandelkow, E., and Zweckstetter, M. *J. Biol. Chem,* **2007**, *282*, 12230-12239.

14. Trinczek, B., Biernat, J., Baumann, K., Mandelkow, E.M., and Mandelkow, E. *Mol. Bio. Cell*. **1995**, *6*, 1887-1902.

15.   Waddell, W. J. *J. Lab. Clin. Med*. **1956**, *48*, 311-314.

16.   Miyaura, N., Yanagi, R., and Suzuki, A. *Synth. Commun*. **1981**, *11*, 513-519.

17.   Chalker, J.M., Wood, C., and Davis, B. *J. Am. Chem. Soc*. **2009**, *131*, 16346-16347.

18.   Pandey, A., Naduthambi, D., Thomas, K., and Zondlo, N.J. *J. Am. Chem. Soc*. **2013**, *135*, 4333-4363.

19.   Forbes, C. R.; Zondlo, N. J. *Organic Letters,* **2012**, *14*, 464-467.

20.   Wang, Q., Chan, T., Hilgraf, R., Fokin, V., Sharpless, K.B., and Finn, M.G. *J. Am. Chem. Soc*. **2003**, *125*, 3192-3193.

21.   Sletten, E.M. and Bertozzi, C.R. *Organic Letters*, **2008**, *10*, 3097-3099

22.   Forbes, C.R. and Zondlo, N.J.

23.   Drewes, G., Trinczek, B., Illenberger, S., Biernat, J., Schmitt-Ulms, G., Meyer, H.E., Mandelkow, E.M., and Mandelkow, E. *J. Biol. Chem*. **1995**, *270*, 7679-7688

24.   Gao, F. *The Design and Application of Protein Kinase-Inducible Domains as a General Approach for Detection of Kinase and Phosphatase Activities.*

25.   Crowther, R.A., Olesen, O.F., and Goedert, M. *FEBS*, **1992**, *309,* 199-202

26.   Drewes, G., Ebneth, A., Preuss, U., Mandelkow, E.M., and Mandelkow, E. *Cell*. **1997**, *89*, 297-308

27.   Meraz-Ríos, M.A., Lira-De León, K.I., Campos-Peña, V., Anda-Hernández, M.A., and Mena-López, R. *J. of Neurochem*. **2010**, *112*, 1353-1367.

28.   Barghorn, S. and Mandelkow, E. *Biochem*. **2002**, *41*, 14885-14896.

29.   Augustinack, J.C., Schneider, A., Mandelkow, E.M., and Hyman, B.T. *Acta. Neuropathol*. **2002**, *103*, 26-35.

30.   Schneider, A., Biernat, J., Bergen, M., Mandelkow, E., and Mandelkow, E.M. *Biochem*. **1999**, *38*, 3549-3558.

31. Jeganathan, S., Bergen, M., Brutlach, H., Steinhoff, H.J., and Mandelkow, E. *Biochem.* **2006**, *45*, 2283-2293.

32. Jeganathan, S., Hascher, A., Chinnathambi, S., Biernat, J., Mandelkow, E.M., and Mandelkow, E. *J. Biol. Chem.* **2008**, *283*, 32066-32076.

33. Liu, F., Kim,Stephen, A.G., Fisher, R.J., and Burke, T.R. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 10-96-1101.

34. Chen, Young K. et al. *PCT Int. Appl.*, **2009**. 2009097578.