1-1-2015

Minimally Perturbing Fluorescence Probe Pairs for the Study of Protein Folding and Misfolding

Rebecca Felice Wissner
University of Pennsylvania, wissner@sas.upenn.edu

Minimally Perturbing Fluorescence Probe Pairs for the Study of Protein Folding and Misfolding

Abstract

The development of new methods that provide mechanistic information on the structural dynamics of proteins represents a significant challenge in the field of biochemistry. Fluorescence spectroscopy is a highly sensitive technique that is ideally suited for monitoring protein movement in situ. However, the most commonly used fluorophores generally yield poor structural resolution, due to their relatively large size compared to the protein of interest. Research in our laboratory has demonstrated that a thioamide, a single atom-substitution of the peptide backbone, is capable of quenching a wide array of fluorophores in a distance-dependent fashion. We have shown that thioamide quenching of tyrosine and tryptophan can be used to monitor biological interactions, such as ligand binding to the protein Calmodulin (CaM). To expand the utility of the thioamide group as a spectroscopic probe, our laboratory has developed semi-synthesis techniques for its installation into full-length proteins. Having validated thioamide quenching of intrinsic protein fluorescence in our model system, we then applied this technique to monitoring the misfolding of the amyloidogenic protein Î±-synuclein (Î±S), implicated in the pathogenesis of Parkinsonâ??s Disease. In order to determine which of our spectroscopic pairs best behaves in accordance with our theoretical models, we also examined thioamide quenching of Cnf using our CaM model system. For intramolecular studies with Cnf/thioamide FRET pairs, we combined unnatural amino acid mutagenesis with native chemical ligation to access double-labeled Î±S using a minimum of chemical synthesis. We have also shown that we can combine unnatural amino acid mutagenesis with expressed protein ligation at methionine to incorporate the probe pair in an entirely traceless manner. Using thioamide/Cnf FRET of our constructs, we were able to monitor conformational changes of monomeric Î±S with unprecedented structural resolution. In addition to our work using thioamides, we have also developed efficient strategies for producing variants of singly and doubly-labeled Î±S containing red-shifted fluorophores for fluorescence polarization (FP) and other FRET based assays, respectively. Most recently, we have shown that we can use site specifically labeled Î±S in conjunction with FP to glean mechanistic insight into the processes of aggregation and disaggregation. Ultimately, these labeled constructs will allow us to study these processes in vivo.

MINIMALLY PERTURBING FLUORESCENCE PROBE PAIRS FOR THE STUDY OF PROTEIN FOLDING AND MISFOLDING

Rebecca Felice Wissner

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

_____
E. James Petersson
Assistant Professor of Chemistry & Assistant Professor of Biochemistry and Molecular Biophysics, Perelman School of Medicine

Graduate Group Chairperson

_____
Gary A. Molander, Hirschmann-Makineni Professor of Chemistry

Dissertation Committee:

Ronen Marmorstein, Professor of Biochemistry and Biophysics, Perelman School of Medicine
Ivan J. Dmochowski, Associate Professor of Chemistry
David M. Chenoweth, Assistant Professor of Chemistry

MINIMALLY PERTURBING FLUORESCENCE PROBE PAIRS FOR THE STUDY OF PROTEIN FOLDING AND MISFOLDING

DEDICATION

To my parents.

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my thesis advisor, Professor E. James Petersson. Without his mentorship, it is unlikely that I would be the scientist that I am today. I am deeply appreciative for having had the opportunity to join the group and to work on several exciting projects. I would also like to thank my committee members, Professor Ronen Marmostein, Professor David Chenoweth, and Professor Ivan Dmochowski for their kind support and encouragement. I am grateful for all of the helpful feedback and motivation that I received throughout the years.

Of course, I must acknowledge the incredible group of graduate, undergraduate, and postdoctoral students that I have had the opportunity to work with during my time in the Petersson group: Jacob Goldberg, Anne Wagner, John Warner, Mark Fegley, Lee Speight, Solongo Ziraldo, Jerri Wang, Stella Chen, Chris Walters, Miklos Szantai-Kis, Jack Ferrie, Ohm Sungwienwong, Christina Cleveland, Christina Grindley, Alyssa Klein, Eileen Moison, Keith Keenan, Anand Muthusamy, Colin Fadzen, Lily Owei, Dr. Tomohiro Tanaka, Dr. Moumita Samanta, Dr. Yun Huang, and Dr. Conor Haney. I have no doubt that my success was a direct result of the collaborative nature and positive spirit of the lab. I am incredibly lucky to have always been surrounded by helpful and upbeat labmates. Outside of the laboratory, I have made many wonderful friends in the Chemistry department who have helped to make my time at Penn enjoyable. In particular, I would like to thank the Dmochowski, Chenoweth, and Saven groups for their generosity and kindness. I am particularly indebted to Brittany Riggle, who has been my strongest supporter and best friend for the past six years.

Working with α-synuclein has been a great challenge, but we have come a long way thanks to our collaborative effort with Professor Virginia Lee at the Center for Neurodegenerative Disease Research. Professor Lee, Professor Kelvin Luk, and Dr. Dustin Covell have been

extremely generous with their materials, protocols, advice, feedback, and time. It is difficult to imagine where my current projects would stand without their invaluable help.

Finally, I would like to acknowledge my family. First, I must thank my parents for inspiring me to pursue a career in science and for taking me seriously when I decided to do so. And to my wonderful sisters - Jennie and Bonnie – thank you both for putting up with me for the for always taking me out when I needed it the most. And last, but certainly not least, I need to thank my fiancé, Michael Witten, for his endless love and encouragement and for supporting me in my commitment to a scientific career.

Table of Contents

ABSTRACT

MINIMALLY PERTURBING FLUORESCENCE PROBE PAIRS FOR THE STUDY OF PROTEIN FOLDING AND MISFOLDING

Rebecca Wissner

E. James Petersson

The development of new methods that provide mechanistic information on the structural dynamics of proteins represents a significant challenge in the field of biochemistry. Fluorescence spectroscopy is a highly sensitive technique that is ideally suited for monitoring protein movement *in situ*. However, the most commonly used fluorophores generally yield poor structural resolution, due to their relatively large size compared to the protein of interest. Research in our laboratory has demonstrated that a thioamide, a single atom-substitution of the peptide backbone, is capable of quenching a wide array of fluorophores in a distance-dependent fashion. We have shown that thioamide quenching of tyrosine and tryptophan can be used to monitor biological interactions, such as ligand binding to the protein Calmodulin (CaM). To expand the utility of the thioamide group as a spectroscopic probe, our laboratory has developed semi-synthesis techniques for its installation into full-length proteins. Having validated thioamide quenching of intrinsic protein fluorescence in our model system, we then applied this technique to monitoring the misfolding of the amyloidogenic protein α-synuclein (αS), implicated in the pathogenesis of Parkinson's Disease. In order to determine which of our spectroscopic pairs best behaves in accordance with our

theoretical models, we also examined thioamide quenching of Cnf using our CaM model system. For intramolecular studies with Cnf/thioamide FRET pairs, we combined unnatural amino acid mutagenesis with native chemical ligation to access double-labeled αS using a minimum of chemical synthesis. We have also shown that we can combine unnatural amino acid mutagenesis with expressed protein ligation at methionine to incorporate the probe pair in an entirely traceless manner. Using thioamide/Cnf FRET of our constructs, we were able to monitor conformational changes of monomeric αS with unprecedented structural resolution. In addition to our work using thioamides, we have also developed efficient strategies for producing variants of singly and doubly-labeled αS containing red-shifted fluorophores for fluorescence polarization (FP) and other FRET-based assays, respectively. Most recently, we have shown that we can use site-specifically labeled αS in conjunction with FP to glean mechanistic insight into the processes of aggregation and disaggregation. Ultimately, these labeled constructs will allow us to study these processes *in vivo*.

List of Tables

List of Illustrations

Chapter 1 : Introduction

§1.1 Preamble

Proteins, the most diverse and structurally varied macromolecules inside of cells, play a vital role in all of the biological processes that are required to sustain life. All living cells require thousands of proteins, linear chains of linked amino acids, for various complex functions such as catalysis of metabolic reactions, regulation of signaling processes, cellular architecture, and small molecule transport. The amino acids that comprise full-length proteins are connected to one another via amide bonds (composed of carbon, oxygen, and nitrogen) and the universal genetic code consists of only twenty naturally-occurring amino acids. These twenty building blocks are strung together in different orders to assemble all of the proteins required for life, from the simplest unicellular organisms up to humans.

Full length proteins generally twist and fold into three-dimensional structures. Pioneering studies performed by Christian B. Anfinsen and coworkers demonstrated that a protein's fold is dictated by the order of its constituent amino acids and the non-covalent interactions that occur between them.[1] Several decades of biochemical research have since established that the efficient formation and maintenance of the correct overall structure is required for proper protein function. Misfolded proteins, which fail to adopt their correctly folded or "native state" structure, are biologically inactive.[2] Furthermore, misfolded proteins can be detrimental to cell health in that they have a tendency to interact unfavorably with other proteins or components of the cell. In the crowded interior of a cell, properly folded proteins are only marginally stable, and misfolding occurs frequently. However, most cells possess a system of safeguards to prevent misfolded proteins from persisting. In some cases, misfolding is reversible. Certain proteins inside of cells, known as molecular chaperones, can specifically bind to misfolded proteins,

shield them from their environment, and enable their refolding. Alternatively, a misfolded protein can be selectively targeted for degradation.[3,4]

Although these distinct quality control mechanisms may render a protein misfolding event harmless, certain types of misfolded proteins can evade this protective machinery and cause disease. An increasing number of neurodegenerative disorders, such as Alzheimer's, Parkinson's, and Huntington's disease have been attributed to a specific kind of protein misfolding known as amyloidogenesis.[2,5] Amyloid formation has also been implicated in a number of other nonneuropathic diseases, including amyotrophic lateral sclerosis, type II diabetes, and familial renal amyloidosis.[2,5] Upon an initial misfolding event, single, monomeric amyloidogenic proteins associate with each other to form larger aggregates of multiple proteins in a self-templating fashion (Figure 1.1). These aggregates undergo further assembly to form highly-ordered structures known as amyloid fibrils, which ultimately deposit in the form of plaques in the brain.

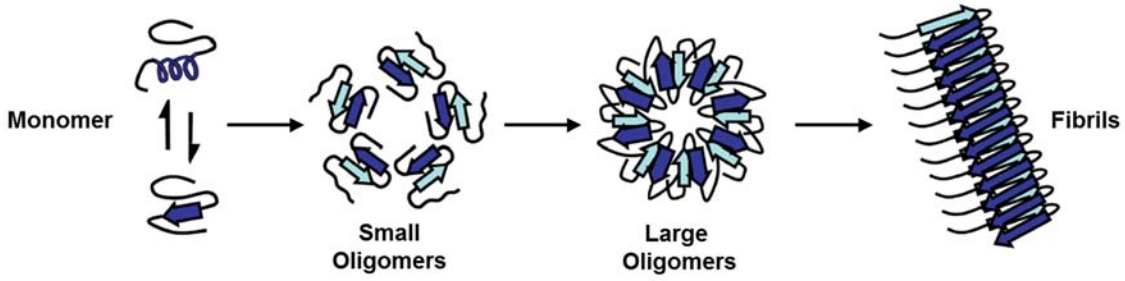

Figure 1.1 Proposed Model of Amyloid Formation.
Monomers under a conformational change that leads to oligomerization and the formation of insoluble amyloid fibrils.

Many species along this pathway (from small aggregates to fully mature fibrils) exhibit neurotoxic properties, but it remains unclear how and why these proteins misfold in the first place. A detailed structural understanding of the misfolding of amyloidogenic

proteins could ultimately result in the rational design of therapeutic strategies to combat the formation of these toxic intermediates.

§1.2 Application of Fluorescence Spectroscopy to the Study of Protein Folding

The development of new methods to ascertain mechanistic information on the structural dynamics of proteins remains an enduring challenge in the field of biochemistry. The most common techniques used to study protein structure are X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. These methods require large amounts of highly purified and well-folded proteins. Although these powerful tools allow one to determine protein structures at atomic resolution, the static images obtained cannot explain how a protein moves over time in order to perform its function. Additional dynamic experiments must be performed that either corroborate or disprove any mechanistic insight that is implied by the observed structure.

Fluorescence spectroscopy represents a highly sensitive technique that is ideally suited for analyzing protein movement in solution.[6,7] Since fluorescence data can be obtained on the sub-microsecond timescale, structural information from these experiments can be acquired in real-time. In order to perform fluorescence spectroscopy, the protein molecule of interest must contain a fluorophore. Two natural amino acids, tyrosine and tryptophan, are fluorescent. However, proteins typically contain multiple copies of these amino acids, which convolute data interpretation, as the fluorescent emissions from duplicate residues are relatively indistinguishable from one another. Therefore, it is often useful to attach to a protein an unnatural fluorescent tag that emits light with a unique wavelength compared to the other amino acids in the protein.

If a fluorescent tag is sensitive to changes in its local environment, it can be used to detect overall protein motion.[8] Such a fluorophore is defined as solvatochromatic.

Solvatochroism effects arise from changes in the polarity of the surrounding environment that lead to alterations in the fluorophore's emissive properties. For example, a fluorescent tag may emit more brightly when it faces the non-polar interior of a protein than when it is exposed to the hydrophilic exterior. In this case, when the protein undergoes a change in structure that places the tag towards the interior, a shift in fluorescence will be detected, and one can infer that protein motion has occurred (Figure 1.2). Although this type of experiment could be used to determine whether a global conformational change has taken place, it cannot reveal how one region of the protein moves with respect to another region of the protein (i.e., how the protein folds).

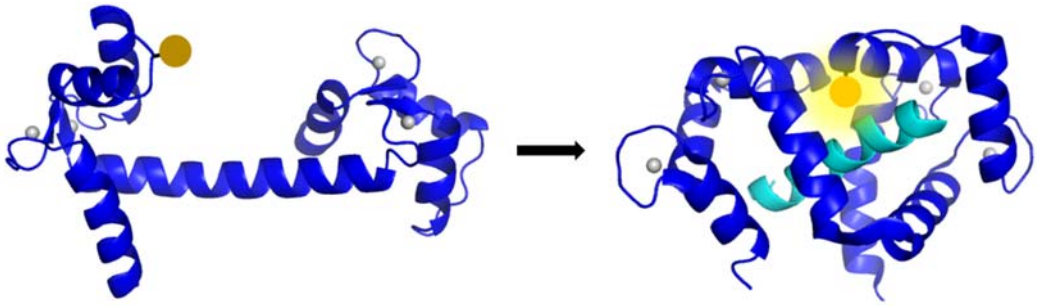

Figure 1.2 Cartoon Representation of Environmentally Sensitive Fluorophores for Monitoring Protein Motion.
The protein calmodulin (CaM) undergoes a substantial conformational rearrangement upon peptide binding. A strategically placed environmentally sensitive fluorophore could be used to report on this structural change.

In order to gain a more detailed depiction of movement, one must label the protein of interest with two tags that can interact in a distance-dependent manner. In some cases, the light emitted by one fluorophore can be attenuated by a second chemical species called a quencher (Figure 1.3).[6] The extent to which the quencher reduces the intensity of the emitted light can be used to measure the distance between the two probes.

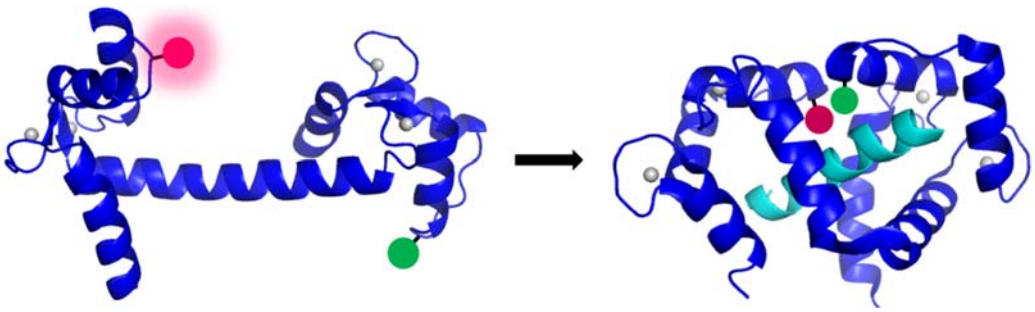

Figure 1.3 Cartoon Representation of Distance-Dependent Fluorescence Quenching. Two fluorophores that interact in a distance-dependent manner can be used to measure the distance between two different locations in a protein.

Clearly, the overall shape of a protein cannot be determined by obtaining a single distance measurement between two fixed locations. However, many distance measurements between several different pairs of points can be obtained by generating a diverse set of labeled proteins that contain the probe pair at different locations (Figure 1.4). Using this strategy, one can map out the key folding interactions that give rise to the protein's overall shape, then take these individual snapshots and compile them into a series in order to examine how these interactions change during protein movement.

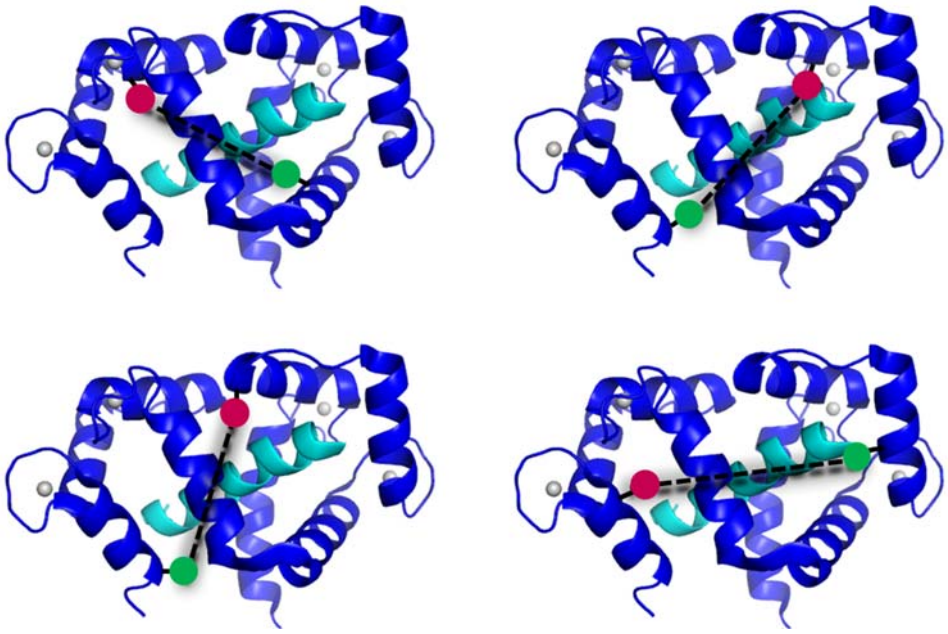

Figure 1.4 Cartoon Representations of Labeled Proteins for FRET Analyses.
Individual distance measurements obtained using a diverse set of labeled proteins can be used to map out protein structure.

There are multiple distinct mechanisms by which fluorescence quenching can occur between two probes in a distance-dependent manner. Förster resonance energy transfer (FRET) is a non-radiative process by which energy is transferred from an electronically excited donor fluorophore to a nearby acceptor chromophore (which may or may not be fluorescent) by means of long-range dipole-dipole coupling interactions between chromophores.[9] In order for FRET to occur, the fluorescence emission spectrum of the donor must overlap with the absorption spectrum of the acceptor chromophore. Additionally, the relative orientations of the donor and acceptor transition dipole moments must be favorably aligned. A further requirement for FRET is that the fluorescence lifetime of the donor must be long enough to permit the energy transfer event to occur. If these criteria are satisfied, the efficiency of energy transfer between a given FRET pair will strongly depend on their spatial separation. According to Förster theory, the efficiency of energy transfer (E) between a donor and an acceptor

chromophore depends on the inverse sixth power of the distance between them (r) by the following equation:

$$E = \frac{R_0^6}{R_0^6 + r^6} \qquad \text{(Eq. 1.1)}$$

where R_0 is the Förster radius, the distance at which energy transfer is 50% efficient. The R_0 of a particular FRET pair can be calculated using:

$$R_0^6 = \frac{9000(\ln 10)\kappa^2 Q_D J}{128\pi^5 n^4 N_A} \qquad \text{(Eq. 1.2)}$$

where Q_D is the quantum yield of the donor, n is the index of refraction of the solvent, N_A is Avogadro's number, and J is the spectral overlap integral. κ^2 is a geometrical factor (defined in the Materials and Methods sections of Chapter 3) that relates the relative orientation of the donor and acceptor transition moment dipoles. For a freely rotating FRET pair, this value is often approximated as $\kappa^2 = 2/3$. Conveniently, FRET can be used to measure distances in a range that is extremely useful for studies of protein folding (10 – 100 Å).

Quenching mechanisms that occur over shorter distances, such as Dexter electron exchange and photoinduced electron transfer (PET) can likewise be exploited for studies in protein folding.[10] Similar to FRET, Dexter and PET quenching mechanisms occur in a non-radiative fashion, meaning that energy transfer from the excited state donor to a nearby acceptor does not involve the emission or absorption of a photon. Dexter energy transfer occurs when an electronically excited donor exchanges an electron with a nearby acceptor in a concerted manner. This event returns the donor to its ground state while leaving the acceptor in the excited state. Since orbital overlap is a requirement of electron exchange, Dexter energy transfer only occurs over very short distances (< 10 Å) and

drops off exponentially at longer distances. Experimentally, FRET and Dexter mechanisms can be difficult to discern from each other. However, it is reasonable to assume that both processes are likely to occur at short donor/acceptor pair distances.

While spectral overlap between the donor and acceptor is a requirement for FRET and Dexter energy transfer, PET quenching is governed by redox chemistry.[11] PET occurs when an electronically excited fluorophore donates to or accepts an electron from a nearby molecule, leading to the formation of a charge-separated complex. This complex can return to the ground state without emission of a photon by spontaneous back electron transfer (also known as charge recombination). Alternatively, this charge-separated complex can return to the ground state by exciplex emission. The direction of electron transfer in the excited state is dictated by the oxidation and reduction potentials of the interacting pair. In order for PET quenching to occur, the formation of the charge separated complex must be energetically favorable upon electronic excitation. The change in Gibbs free energy (ΔG°) resulting from PET can be evaluated using the Rehm-Weller formalism:

$$\Delta G^{\circ}_{ET} = F\{E^{\circ}_{Ox}(D) - E^{\circ}_{Red}(A)\} - E_{0,0} + C \qquad \text{(Eq. 1.3)}$$

where F is the Faraday constant, $E^{\circ}_{Ox}(D)$ and $E^{\circ}_{Red}(A)$ are the oxidation and reduction potentials of the electron donor and acceptor molecules, respectively, $E_{0,0}$ is the zero vibrational electronic excitation energy of the fluorophore, and C is a term accounting for Coulombic interactions, which is typically assumed to be negligible in water.[12] Since molecular collisions are requisite for efficient electron transfer, quenching is only likely to occur when the donor and acceptor moieties are in close spatial proximity (< 15 Å). However, the distance-dependence of PET quenching is difficult to predict for a given

donor/acceptor pair. Although PET quenching can report on short-range interactions, this method is not appropriate for the accurate determination of interchromophore distances.

Both long and short-range fluorescence quenching mechanisms have been applied to the study of protein folding. Whereas FRET can be exploited to accurately measure the distances between donor and acceptor probes in the 10 to 100 Å range, PET can be used to determine whether two distinct regions of a protein approach molecular contact. In general, FRET is more useful for monitoring large-scale conformational changes. However, genetically encodable FRET labels (such as variants of the Green Fluorescent Protein, GFP) are often quite large relative to the protein being studied.[13] On the other hand, producing heterogeneous, double-labeled proteins for FRET studies containing extrinsic probes is a challenge in its own right. Fortunately, new methods for generating site-specifically double-labeled proteins for FRET studies are rapidly emerging.

§1.3 Recent Strategies for the Dual Incorporation of Fluorescent Labels into Proteins.

Over the past two decades, FRET has emerged as a powerful tool for studying protein dynamics.[14] Despite the widespread use of this technique, the site-specific labeling of proteins with a suitable pair of donor and acceptor dyes remains a significant challenge in the field of chemical biology. The genetically incorporable green fluorescent protein and its multicolor variants are commonly employed as labels for FRET studies. However, fluorescent protein tags are often quite large relative to the protein being studied and are therefore limited in the number of positions in which they can be placed, reducing the amount and quality of data that can be obtained from these experiments.[13]

In order to circumvent the disadvantages associated with fluorescent proteins, several post-translational strategies have been developed for transferring smaller

fluorescent labels onto proteins at specific amino acid sequences. Tsien and coworkers demonstrated that a genetically incorporable tetracysteine motif can be selectively labeled using a bisarsenical fluorescein derivative known as FlAsH.[15] FlAsH can be administered as a non-fluorescent complex with ethanedithiol (EDT). Displacement of EDT by the tetracysteine motif leads to a dramatic increase in the fluorescence quantum yield of the dye, rendering this system fluorogenic. Alternatively, several groups have developed chemoenzymatic protein labeling methods that capitalize on the selectivity of certain enzymes for specific amino acid sequences. For example, the enzyme sortase, which catalyzes a transpeptidase reaction at an internal recognition motif, can be used to conjugate fluorescent labels to the protein termini.[16,17] Other enzymes, such as biotin ligase and lipoic acid ligase, can be used to attach fluorophores to amino acid sidechains.[18,19] However, all of these methods require the insertion of a 5-15 amino acid recognition motif that may ultimately be disruptive to native protein structure and function.

Fluorescent dyes can also be introduced into a protein by chemical modification of amino acid functional groups such as the nucleophilic cysteine (Cys) thiol. Solvent accessible Cys residues can be readily labeled using fluorophores that are conjugated to alkyl halides or maleimide handles (Figure 1.5).[20]

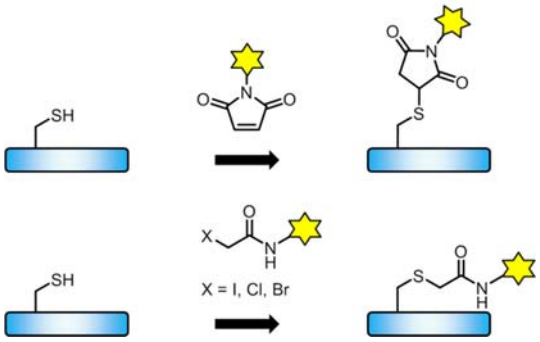

Figure 1.5 Chemical Modification of Cys Residues.
Cys can be readily labeled using probes attached to maleimide (top) or alkyl halide (bottom) handles. The yellow star represents a fluorescence probe.

Although Cys modification is extremely useful for introducing a single fluorescent label into a protein, attempts to place two labels on two independent, engineered Cys residues lead to heterogeneously labeled populations that are not optimal for use in FRET studies.[21] Moreover, this strategy is only appropriate for proteins that either lack endogenous Cys or for those that are tolerant to Cys replacement with an unreactive amino acid.

Recently, several groups have reported improved strategies for producing homogeneous, doubly-labeled proteins for FRET studies using unnatural amino acids (Uaas). The site-specific incorporation of Uaas into full-length proteins can be achieved using a method known as nonsense suppression. The nonsense codons, Opal (UGA), Amber (UAG), and Ochre (UAA), signal the termination of protein translation by serving as binding sites for ribosomal release factors. Pioneering studies performed by Schultz and coworkers demonstrated that an unnatural tRNA bearing an anticodon cognate to an intrinsic stop signal can suppress the termination of translation (Figure 1.6).[22] In order to facilitate the ribosomal incorporation of a Uaa in response to a stop codon, the suppressor tRNA must first be charged with the Uaa of interest.

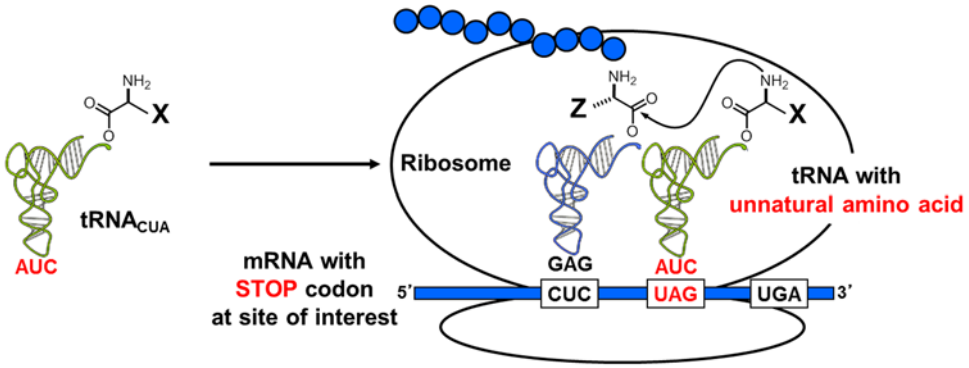

Figure 1.6 Ribosomal Incorporation of Uaas.
Uaa-tRNA$_{CUA}$ (X = Uaa side chain, Z = AA side chain) can suppress the termination of translation in response to UAG stop codons transcribed in the mRNA message.

Uaa-tRNA can be produced *in vitro* using semi-synthetic methods. However, semi-synthetic Uaa-tRNAs are generally restricted to use in cell-free translation systems that are costly and low-yielding in comparison to *in vivo* protein expression. Enzymatic acylation of a suppressor tRNA can be accomplished using an evolved mutant aminoacyl tRNA synthetase (aaRS) that is selective for the Uaa of interest.[23] In order to perform nonsense suppression *in vivo*, the aaRS/tRNA pair must be orthogonal to the endogenous cellular machinery. This requirement can be met by importing an evolved aaRS/tRNA pair derived from one organism into an alternative expression host. To date, over 150 Uaas have been incorporated into recombinant proteins for various structural and functional studies.[24] Two types of genetically incorporable Uaas are useful for generating doubly-labeled proteins for FRET studies: those that are intrinsically fluorescent and those that contain a selective reactive handle for further modification. A wide variety of Uaas that meet these criteria have been reported and several examples will be described herein.

Uaas bearing reactive moieties that can be selectively modified in the presence of biological functional groups are described as bioorthogonal.[25] Several research groups

have shown that bioorthogonal Uaas can be used in conjunction with Cys to produce doubly-labeled proteins for FRET studies. Previously, it had been shown that a mutant TyrRS/tRNA pair derived from the archea *Methanococcus jannaschi* (*Mj*TyrRS/tRNA$_{CUA}$) can be used to incorporate a variety of *p*-substituted Phe analogs into proteins in response to an amber stop codon in live *E. coli* cells.[26,27] In 2008, Brustad *et al.* used this pair to incorporate *p*-acetylphenylalanine (*p*AcF) into a variant of bacteriophage T4 lysozyme containing a single Cys residue.[28] *p*AcF contains a ketone group that can undergo a selective reaction with an alkoxyamine to form an oxime linkage. Labeling of protein-incorporated *p*AcF with Alexa488-alkoxyamine proceeded with high yields and was followed by Cys modification using Alexa545-maleimide. Single-molecule FRET measurements obtained using two distinct doubly-labeled T4 lysozyme constructs were in good agreement with distances observed in the reported crystal structure. Additionally, the authors examined denaturant-mediated unfolding of the protein, and found that expansion of the N- and C-termini occurs prior to the unfolding transition.

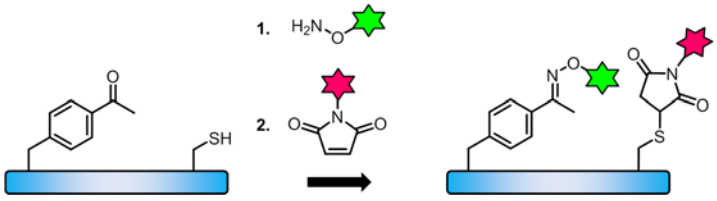

Figure 1.7 Dual Labeling of *p*AcF and Cys.
*p*AcF can undergo a selective reaction with an alkoxyamine to form an oxime linkage. Labeling of *p*AcF can be followed by Cys modification with a maleimide. The green star represents Alexa488 (donor). The red star represents Alexa545 (acceptor).

Seo *et al*. also used the *Mj*TyrRS/tRNA$_{CUA}$ pair to incorporate a different Tyr analog, *p*-azidophenylalanine (*p*AzF) into maltose binding protein (MBP) containing a single Cys residue.[21] In the presence of Cu(I), azides undergo 1,3-dipolar cycloaddition

with alkynes to form stable 1,4-disubstituted triazole linkages (commonly referred to as Cu(I)-catalyzed click chemistry).[29] However, the reported labeling yield of *p*AzF in this study was quite low (66%), which could likely be attributed to *in vivo* reduction of the aryl azide group to an unreactive amine (Figure 1.8).

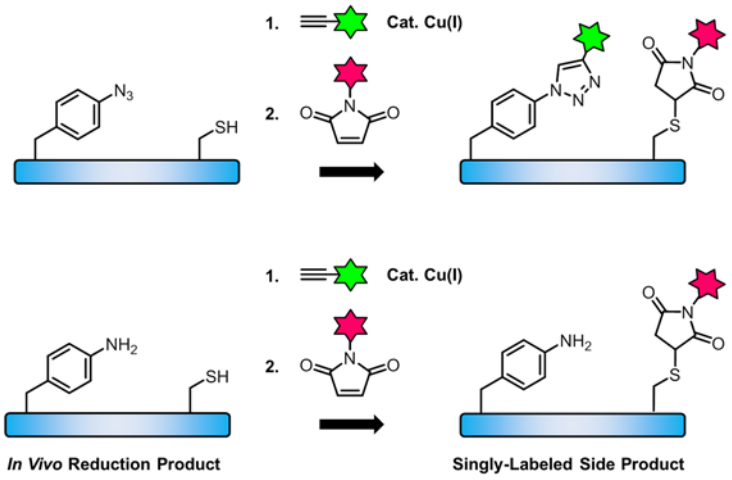

Figure 1.8 Dual Labeling of *p*AzF and Cys.
*p*AzF can undergo a selective reaction with an alkyne in the presence of Cu(I) to form a 1,4-disubstituted triazole linkage. Cu(I)-mediated cycloaddition can be followed by Cys modification with a maleimide. However, *in vivo* reduction of *p*AzF to the unreactive amine leads to a mixture of products. The green star represents fluorescein (donor). The red star represents tetramethylrhodamine (acceptor).

Recently, the pyrrolysyl aaRS derived from *Methanosarcina barkeri* (*Mb*PylRS)—an archea that naturally incorporates pyrrolysine into proteins—has emerged as an extremely useful aaRS that can be evolved to incorporate a variety of bioorthogonal Uaas.[24,30] In 2011, Nguyen *et al*. reported the genetic incorporation of a Uaa containing a masked 1,2-aminothiol, N^{ε}-L-Thiaprolyl-L-lysine. Upon deprotection of the thiazolidine ring, the 1,2-aminothiol can condense with a cyanobenzothiazole (CBT) in a bioorthogonal manner.[31] In order to demonstrate that 1,2-aminothiols can be used to produce double-labeled proteins, the authors introduced the masked Uaa into a mutant of the protein calmodulin (CaM) containing a single Cys. First, the Cys residue was labeled using rhodamine-maleimide. Next, the 1,2-aminothiol was deprotected using

methoxyamine and then treated with CBT-fluorescein. Although the ultimate doubly-labeled protein was indeed homogeneous, this method requires an additional deprotection step that may not be compatible with proteins that are sensitive to treatment with high concentrations of methoxyamine.

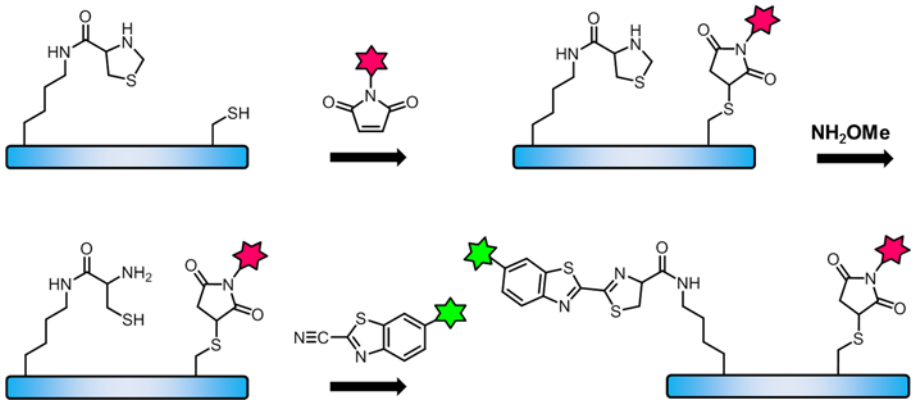

Figure 1.9 Dual Labeling of an Engineered 1,2-Aminothiol and Cys pair.
Cys labeling is performed prior to deprotection of N$^{\varepsilon}$-L-Thiaprolyl-L-lysine. The deprotected 1,2-aminothiol can undergo a selective condensation reaction with a cyanobenzothiazole. The green star represents fluorescein (donor). The red star represents tetramethylrhodamine (acceptor).

Uaa mutagenesis can be combined with labeling at Cys to afford doubly-labeled constructs, but this strategy is still not compatible with proteins that contain multiple Cys residues. In order to incorporate two distinct Uaas into a single protein, an additional blank codon must be introduced. In 2010, the Liu group demonstrated that the *Methanosarcina maize* PylRS-tRNA pair (*Mm*PylRS-tRNA) can be modified to suppress the ochre codon and that this protocol can be coupled with *Mj*TyrRS/tRNA$_{CUA}$ to incorporate two Uaas into a single protein in a site-specific manner.[32] As an initial proof of principle, Wan *et al*. used this strategy to incorporate an azide (*p*AzF) as well as an alkyne-containing amino acid (N$^{\varepsilon}$-propargyloxycarbonyl-L-lysine) into GFP. Dual labeling was achieved by performing two sequential copper mediated azide-alkyne cycloaddition reactions using propargyl-conjugated fluorescein to label the azide and 3-azido-7-hydroxycoumarin to label the alkyne. Although the cross-reactivity of the Uaa

pair was suppressed by using an excess of dye in the first labeling step, this pair may not be suitable for labeling positions that are in close spatial proximity.

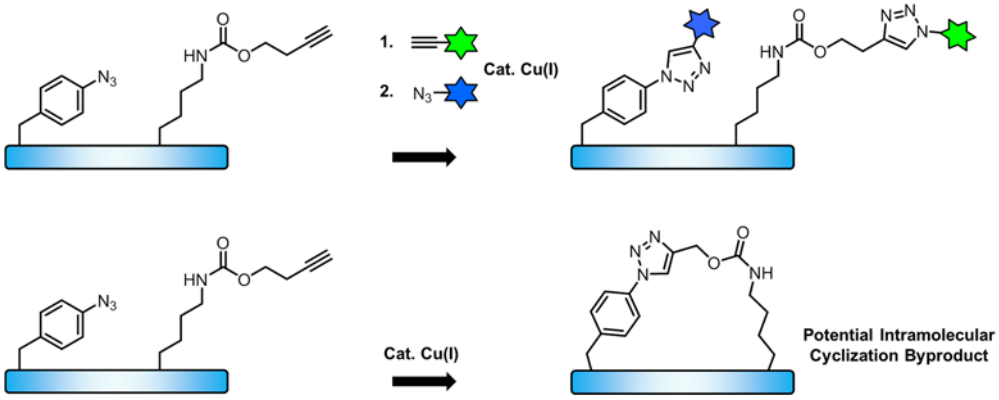

Figure 1.10 Dual Labeling of *p*AzF and N$^{\varepsilon}$-propargyloxycarbonyl-L-lysine.
Sequential Cu(I)-mediated 1,3-dipolar cycloaddition reactions can be performed to label the ribosomally incorporated azide and alkyne handles. However, if the probe pair is placed in close proximity, an intramolecular cyclization reaction may occur. The blue star represents hydroxycoumarin (donor). The green star represents fluorescein (acceptor).

To overcome this limitation, Wu *et al*. used a similar dual suppression system to incorporate a different bioorthogonal pair, *p*AzF and 2-amino-8-oxononanoic acid (an alternative ketone-containing Uaa), into glutamine binding protein (QBP).[33] In this case, strain-promoted azide-alkyne cycloaddition (SPAAC) could be used to effect the click reaction in the absence of a metal catalyst. Specifically, it had been reported that dibenzocyclooctyne (DBCO), which is highly strained about its carbon-carbon triple bonds, can undergo efficient 1,3-dipolar cycloadditions with azides, driven by the relief of ring strain that results upon triazole formation.[34] Catalyst-free double-labeling of QBP in one pot was performed by simply incubating the protein with fluorescein-DBCO and 7-diethylaminocoumarin-3-alkoxyamine. Quantitative conversion to the dual-labeled construct was achieved by allowing the reaction to proceed overnight. However, it is important to note that *in vivo* reduction of *p*AzF is typically problematic; it is unclear why higher labeling yields were observed in this specific case.

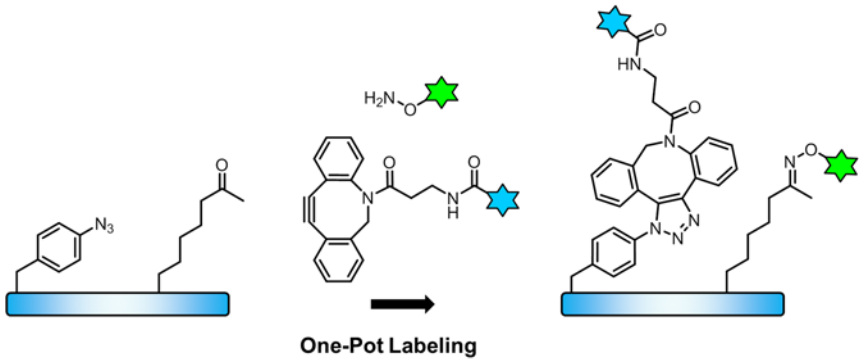

Figure 1.11 Dual Labeling of *p*AzF and 2-Amino-8-Oxononanoic acid.
Dual labeling was achieved by incubating the modified protein with alkoxyamine and DBCO-conjugated fluorophores in a single pot. The blue star represents diethylaminocoumarin (donor). The green star represents fluorescein (acceptor).

In order to demonstrate that dual-labeled QBP could be used for folding studies, denaturant-mediated unfolding of the protein was examined by FRET. Crystallographic studies have shown that upon binding to glutamine, QBP undergoes a significant structural rearrangement. Curiously, FRET signals of the dual-labeled protein in the presence and absence of glutamine were similar. The authors of the study proposed that the lack of observed changes in FRET could reflect the existence of a dynamic conformational ensemble in solution that cannot be detected using crystallographic techniques. However, further studies must be performed in order to fully explain this unexpected observation.

Recently, the Chin and Mehl laboratories reported an elegant alternative strategy for incorporating two Uaas into a single protein.[35] Previous studies have shown that Uaa-tRNAs with expanded four-base anticodons (tRNAxxxx) can read through a four-base codon inserted into the standard mRNA reading frame.[36] In theory, introduction of a four-base pair codon system should provide up to 256 blank codons. However, natural ribosomes do not efficiently process quadruplet codons as they are not well accommodated by the decoding center. Prior to this work, the Chin laboratory had

developed orthogonal ribosomes that translate mRNA messages with altered Shine-Dalgarno sequences, which cannot be recognized by natural ribosomes.[37] Since orthogonal ribosomes are not responsible for expressing or maintaining the cellular proteome, they can be engineered to perform new functions. In 2012, the Chin group evolved an orthogonal ribosome (ribo-Q1) capable of efficiently decoding quadruplet codons and demonstrated that two mutually orthogonal aaRS/tRNA pairs, *Mj*TyrRS-tRNA$_{UCCU}$ and *Mb*PylRS-tRNA$_{CUA}$, can direct the site-specific installation of two Uaas into a single protein.[38]

In 2014, Wang *et al*. reported an optimized orthogonal ribosome system for the dual incorporation of bioorthogonal Uaas.[39] In order to enhance the read-through of quadruplet codons, a library of Pyl tRNA$_{XXXX}$ mutants was screened to identify those exhibiting improved translation efficiencies. Presumably, the evolved tRNA$_{XXXX}$ mutants contain expanded anticodon loops that can bind more efficiency to quadruplet codons in the mRNA message. Although the authors were able to direct the incorporation of two distinct Uaas in response to two different quadruplet codons, higher incorporation efficiencies were observed by coupling the evolved *Mb*PylRS/tRNA$_{UACU}$ pair with *Mj*TyrRS/tRNA$_{CUA}$. With this improved system in hand, the authors were able to demonstrate that several chemically distinct pairs of Uaas can be site-specifically installed into full-length proteins with improved yields.

Next, the authors identified a Uaa pair that could be used to efficiently double-label a protein in a one-pot reaction in the absence of any additional catalyst. Prior work had shown that Uaas containing functional groups that participate in inverse electron-demand Diels-Alder reactions can be labeled in a bioorthogonal manner.[40-42] Tetrazine

groups rapidly undergo this type of cycloaddition reaction with strained alkenes and alkynes. Chin and coworkers identified a Uaa pair, norbornyl-lysine (NorK) and tetrazinyl-phenylalanine (TetPhe), that can be labeled using a tetrazine and alkyne-conjugated probe, respectively. Importantly, this Uaa pair is only minimally cross-reactive. CaM containing both Uaas was produced in *E. coli* cells harboring plasmids for ribo-Q1 as well as aaRS/tRNA pairs selective for NorK and TetPhe. Quantitative double-labeling was achieved by sequentially incubating the expressed protein with BODIPY-TMR-tetrazine and BODIPY-fluorescein-bicyclononyne.

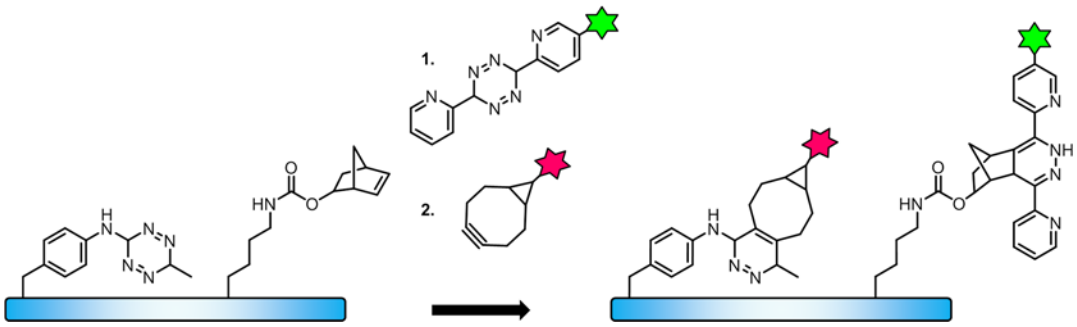

Figure 1.12 Dual Labeling of NorK and TetPhe.
Dual labeling was achieved by performing two sequential inverse electron demand Diels Alder reactions. The green star represents BODIPY-FL (donor). The red star represents BODIPY-TMR (acceptor).

CaM is a 140-amino acid neuronal signaling protein that undergoes a conformational rearrangement upon binding four Ca^{2+} ions.[43] In an initial proof of principle experiment, urea denaturation of folded CaM containing the donor/acceptor pair at the N- and C-termini was monitored using FRET. As expected, global unfolding of CaM resulted in a decrease in FRET efficiency between the two dyes. In order to monitor local conformational changes of the N-terminal domain in response to increasing amounts of Ca^{2+}, CaM was labeled with the donor/acceptor probe pair at positions 1 and 40, respectively. FRET measurements revealed that the N-terminal domain undergoes two distinct structural transitions upon the first and last equivalents of Ca^{2+}.

Recently, Sachdeva *et al*. used the quadruplet decoding orthogonal ribosome system to incorporate an optimized Uaa pair to double-label proteins with increased efficiency.[44] Prior to this work, the Chin group evolved PylRS to incorporate a 1,3 disubstituted cyclopropene containing amino acid that can be rapidly modified using a tetrazine probe.[45] This Uaa was installed into CaM along with *o*-propargyl Tyr (Ppy) such that the mutant protein could be labeled using a combination of Cu(I)-mediated azide/alkyne cycloaddition and inverse electron demand Diels-Alder reactions. Double-labeling was performed in a single pot by incubating the mutant CaM with BODIPY-FL-azide and BODIPY-TMR-tetrazine in the presence of click reaction catalysts. By using 20 equivalents of each dye, quantitative dual modification was complete within 30 minutes.

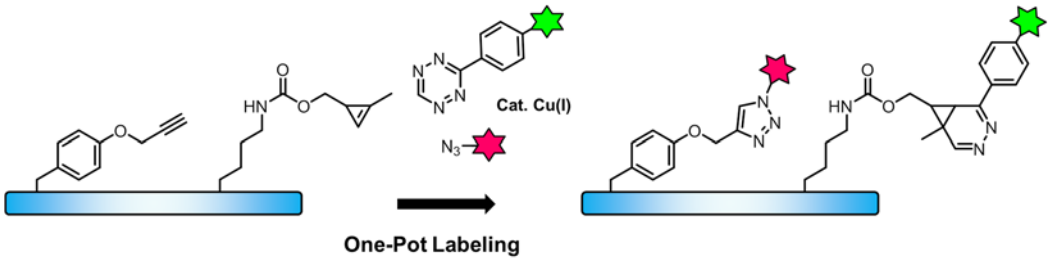

Figure 1.13 Dual Labeling of Ppy and a 1,3-Disubstituted Cyclopropene.
Efficient one-pot labeling was achieved by incubating the protein with tetrazine and alkyne-conjugated fluorescence probes in the presence of Cu(I). The green star represents BODIPY-FL (donor). The red star represents BODIPY-TMR (acceptor).

In 2013, Chen *et al*. reported an alternative efficient strategy for producing labeled proteins for FRET studies.[46] Prior to this work, other groups had shown that Uaas containing fluorescent side chains can be genetically incorporated into proteins and used as FRET partners with Trp or Tyr.[47,48] However, this method is not compatible with proteins containing multiple copies of these residues. In order to incorporate two fluorescent Uaas into a single protein, two distinct blank codons must be suppressed as

previously described. In this study, amino acyl tRNAs containing 4-biphenyl-L-phenylalanine (BpPhe-tRNA$_{CCCG}$) and L-(7-hydroxycoumarin-4-yl)ethylglycine (HCou-tRNA$_{CUA}$) were prepared using semi-synthesis. In this system, BpPhe serves as a low excitation (280 nm) donor fluorophore. Dual suppression of each cognate codon was achieved during *in vitro* translation of a message encoding the enzyme dihydrofolate reductase (DHFR). DHFR undergoes a modest conformational change upon binding its inhibitor trimethoprim (TMP). In order to monitor this transition, DHFR was prepared containing BpPhe and HCou at positions 17 and 115, respectively. Next, changes in FRET were monitored as a function of TMP concentration. Despite its elegance, the method suffers from various drawbacks. The production of semi-synthetic tRNAs is time consuming and the amount of protein generated by *in vitro* translation is usually limited to μgram quantities. Additionally, BpPhe cannot be selectively excited in the presence of Trp or Tyr. Despite these disadvantages, this important work has set the stage for devising an optimized strategy for producing proteins containing intrinsically fluorescent Uaa pairs.

To summarize, Uaa-based methods for the production of doubly fluorescently labeled proteins have undergone dramatic improvements in a short period of time. These new technologies are efficient, facile, and remarkably versatile. In the future, these strategies will enable researchers to generate vast libraries of doubly-labeled constructs for mapping out protein conformational changes in real-time with residue-level precision.

§1.4 Introduction to α-Synuclein.

Parkinson's disease (PD), a debilitating movement disorder whose sufferers exhibit resting tremors, muscular rigidity, bradykinesia, and postural instability, is the second most common aging-related neurodegenerative disease.[49] Histologically, PD is

characterized by dopaminergic nerve cell death in the *substantia nigra* of the brain and by the presence of insoluble proteinaceous inclusions in the remaining neurons of the same region.[50] These intracellular inclusions, known as Lewy bodies (LBs) and Lewy neurites, were found to contain a filamentous form of the protein α-synuclein (αS).[51] αS is a small (14.5 kDa) intrinsically disordered protein that is abundantly expressed in the brain. Although the precise physiological role of αS is unknown, several lines of evidence implicate a function for the protein in synaptic vesicle trafficking.[52] Moreover, αS has been clearly shown to play a role in the pathogenesis of PD. While most cases of PD are sporadic, point mutations (A30P, E46K, A53T, H50Q), duplications, and triplications of the αS gene have all been linked to familial forms of early-onset PD.[53] Expression of human wild-type or mutant αS in transgenic animal models produces age-related motor defects as a result of neuronal cell death.[54,55] Recently, it has been reported that intracerebral injections of αS fibrils induce the formation of intracellular inclusions in the brains of nontransgenic mice.[56] These studies revealed that misfolded copies of αS can recruit native protein and convert it into a pathogenic form, thus demonstrating that they are the causative agent in the spread of PD.

αS has three distinct domains: an N-terminus comprised of seven imperfect repeat sequences bearing the hexamer motif KTKEGV, a central hydrophobic region known as the nonamyloid-β component (NAC) domain, and a flexible, negatively-charged C-terminus (Figure 1.14).

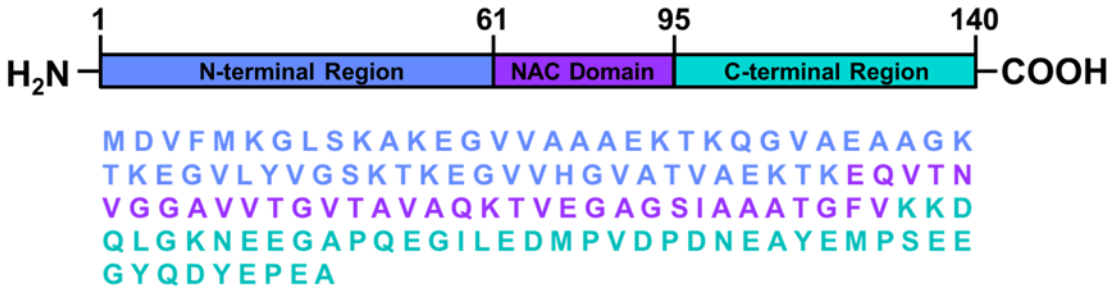

Figure 1.14 Primary Structure of αS.
Top: αS is comprised of three distinct domains. Bottom: αS FASTA sequence color coded to correspond to each domain above.

It is generally accepted that αS exists as a disordered monomer *in vivo*; however some controversial studies suggest that the physiologically relevant form of αS is actually a metastable tetramer.[57] Although it is possible that native αS is folded, the protein nonetheless exhibits remarkable structural plasticity.[58] Upon binding to membrane lipids, αS monomers can adopt α-helical conformations. Misfolding of monomeric αS leads to a conformational change that triggers aberrant oligomerization. Transient αS oligomers with β-sheet character eventually mature into insoluble cross-β strand fibrils. Despite increasing research in this area, the precise, molecular details of αS misfolding have yet to be described. Accordingly, a more complete picture of the dynamic structural changes of αS upon misfolding would greatly facilitate our understanding of the pathogenesis of PD.

§1.5 Structural Studies of Monomeric α-Synuclein.

αS was originally characterized as an intrinsically disordered protein using a variety of structural methods. Analysis of its CD spectrum revealed that the protein is predominantly random coil in solution.[59] FTIR measurements (which are generally more sensitive for the detection of β-sheet structure) confirmed that αS does not contain any significant secondary structure. However, small-angle X-ray scattering studies showed

that αS has a radius of gyration (R_g) that is smaller than expected for a fully extended random-coil conformation (40 vs. 52 Å), but larger than predicted for a folded globular protein of equivalent chain length (15 Å).[60]

Several groups have investigated the structure of monomeric αS using multidimensional heteronuclear NMR spectroscopy.[61,62] A variety of NMR experiments performed using ^{13}C and ^{15}N labeled αS confirmed that the monomer is largely unfolded and devoid of any stable secondary structure. However, it was also observed that the first 100 amino acids exhibit a slight preference towards adopting helical torsion angles, particularly between residues 6 and 37. Whereas this transient helical structure was detected in the N-terminus, the C-terminal region was less well defined and is therefore considered to adopt a fully extended conformation in solution.

Although it has been established that purified recombinant αS exists as a disordered monomer, recent reports have argued that the use of harsh purification protocols have prevented researchers from correctly identifying the native state of αS.[57,63] Moreover, it is possible that the structure of αS is influenced by macromolecular crowding in the cellular milieu. In a recent study, Binolfi *et al*. acquired high quality 2D NMR spectra of labeled αS in intact *E. coli* cells.[64] The obtained NMR spectra were remarkably similar to spectra acquired *in vitro*, indicating that αS is predominantly disordered in a crowded cellular environment. Although these studies do not rule out the possibility that native αS is folded in eukaryotic cells, they do confirm that αS is capable of adopting extended, disordered conformations *in vivo*.

Despite the fact that αS lacks fully formed α-helices and β-sheets, the hydrodynamic radius of αS is smaller than that expected for a random coil structure,

suggesting that long-range tertiary contacts may have a role in the compaction of the protein. In order to investigate this phenomenon, Dedmon *et al*. performed paramagnetic relaxation enhancement (PRE) NMR experiments.[65] PRE measurements are based on the altered relaxation rates of nuclei that are in close proximity to paramagnetic probes (also referred to as spin labels) and can be used to identify interatomic distances up to 25 Å apart. A library of spin-labeled αS mutants was prepared by selectively labeling engineered Cys mutants with a nitroxide paramagnetic probe known as MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate). By applying distance constraints to molecular dynamics (MD) simulations, the authors found that the native αS ensemble samples a variety of non-random, compact conformations.

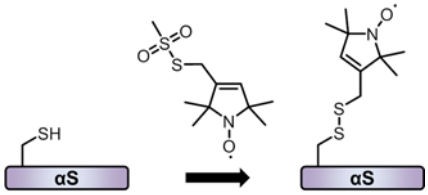

Figure 1.15 Labeling of αS with a Nitroxide Spin Label.
Cys residues react with MTSL to form a disulfide linkage.

Intriguingly, long-range contacts were observed between the C-terminal tail and the central region of the protein. PRE experiments performed by a different laboratory confirmed this initial finding and also revealed significant contact between N- and C-terminal residues.[66] Thus, the authors proposed that long-range interactions, especially those mediated by the C-terminus, may function to shield the highly hydrophobic region of the protein and thereby prevent its assembly into oligomeric species. This notion is consistent with the experimental observation that C-terminally truncated variants of αS form fibrils more rapidly than full-length αS.[67]

Furthermore, FRET experiments have been used to probe the structure and dynamics of monomeric αS. In 2004, Lee *et al.* measured distance distributions in αS using Trp/3-NitroTyr (W/Y-NO_2) as a donor/acceptor (DA) pair.[68] The Förster distance of this DA pair is approximately 26 Å. Since native αS does not contain any Trp residues, site-directed mutagenesis can be performed to produce single Trp-containing proteins. In order to convert particular Tyr residues into their nitrated forms, the expressed protein was treated with tetranitromethane. To prevent chromophore cross-talk, all Tyr residues aside from the chosen acceptor site were mutated to Phe. Distance measurements were obtained using a total of six DA pairs spanning the entire length of the protein. Distance distributions extracted from time-resolved Trp fluorescence measurements revealed that while the protein is highly dynamic, compacted conformers also contribute to the population of the ensemble.

The Rhoades laboratory has shown that single-molecule FRET (smFRET) can also be used to study the conformational behavior of monomeric αS.[69-71] The use of smFRET is particularly advantageous for studying monomeric αS as it allows one to directly probe the structure of αS at extremely low concentrations. Therefore, the acquired FRET data are unlikely to be complicated by intermolecular interactions arising from populations of oligomeric species. Heterogeneously double-labeled αS for smFRET experiments can be produced by sequential labeling of engineered Cys residues with maleimide-conjugated Alexa Fluor dyes. While the smFRET data suggest that αS undergoes substantial conformational rearrangements in solution, significant interactions between the N- and C-termini were observed at physiological pH. Notably, significant compaction of αS was observed at low pH.[71] Since low pH conditions are known to

accelerate aggregation, the authors of this study postulated that this collapsed structure represents the pathogenic, aggregation-prone conformation. smFRET data were later used to guide computational studies of αS structure. By applying the experimentally determined distances as constraints in MD simulations, the authors were able to generate a model of the disordered ensemble in solution. Remarkably, the dimensions of the computationally-derived ensemble were in excellent agreement with experimental measurements of R_g.

§1.6 Lipid Induced Folding of Monomeric α-Synuclein.

Upon binding to negatively charged lipid membranes or membrane-mimetic detergent micelles, αS adopts a well-folded helical structure. Detailed NMR analyses revealed that the micelle-bound form of αS consists of two non-interacting antiparallel helices in the N-terminal region connected by a short linker between residues 38-44 (Figure 1.16).[72] However, in the presence of unilamellar vesicles, αS has been shown to adopt an extended helical conformation by smFRET.[73] In either conformation, the C-terminal domain of αS is predicted to remain disordered. The lipid-bound conformations of αS have been extensively probed by several groups using a variety of biophysical techniques including NMR and smFRET.[62] While the details of these reported structures slightly vary, there appears to be a general consensus that the nature of the fold is likely to be dependent upon the size of the vesicles employed.

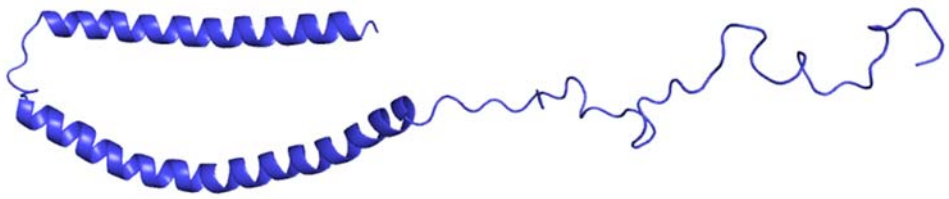

Figure 1.16 Solution NMR Structure of αS Bound to SDS Micelles.
PDB Code: 1XQ8

§1.7 TMAO Induced Folding of Monomeric α-Synuclein.

Osmolytes are naturally-occurring small organic molecules that have the capacity to alter the physicochemical environment of biological systems. Broadly, these molecules are categorized by their effect on protein folding. A protecting osmolyte is known to enhance protein stability, whereas a destabilizing osmolyte causes proteins to unfold. Trimethylamine *N*-oxide (TMAO) is the most efficient protecting osmolyte studied to date. Previous studies have shown that the addition of TMAO can induce thermodynamically unstable proteins to fold and regain their functional activity *in vitro*.[74,75] Additionally, TMAO has been shown to counteract the destabilizing effects of the denaturant urea.[76] More recently, Bandyopadhyay *et al.* have shown that TMAO can even promote the folding of mutation-destabilized globular proteins in live *E. coli* cells.[77]

The exact molecular mechanism by which TMAO and other protecting osmolytes enhance protein folding is currently under investigation. One mechanistic scenario has been presented by Zou *et al.*, who measured the transfer free energy of a model peptide backbone from water to TMAO.[78] The authors proposed that an unfavorable interaction of TMAO with the protein backbone results in a stabilizing effect on proteins. In other words, TMAO destabilizes the unfolded state of a polypeptide chain and shifts the folding equilibrium to the native state. This phenomenon is referred to as the

"osmophobic effect". Others have proposed that TMAO induces protein folding simply by acting as a macromolecular crowding reagent through excluded volume effects.[79]

Since it had been reported that TMAO can promote the folding of other IDPs, Uversky *et al*. examined the effect of TMAO on the structural properties of αS.[80] CD measurements revealed that αS adopts a predominantly helical conformation in high concentrations of TMAO. Next, the authors probed the TMAO-induced structure of αS using fluorescence spectroscopy. 8-Anilinonaphthalene-1-sulfonic acid (ANS) is an environmentally-sensitive dye that binds to hydrophobic surfaces of partially folded proteins, resulting in a concomitant increase in fluorescence intensity. In the presence of αS, ANS fluorescence increased as a function of TMAO concentration. Although this result may be consistent with the formation of a partially folded intermediate, the authors did not examine the effect of TMAO on ANS fluorescence in the absence of αS. Additionally, folding of αS was examined using acrylamide as a contact-dependent quencher of Tyr fluorescence. In TMAO, acrylamide quenching of Tyr fluorescence was reduced, indicating that these residues are somewhat protected from solvent in the TMAO-induced conformation. However, acrylamide quenching of free Tyr in the presence and absence of TMAO was not reported. Although these results are interesting, they must be considered with caution, as both experiments lack an appropriate control that would be required to fully interpret the reported data.

In 2012, Ferreon *et al*. used smFRET to study the effect of osmolytes on the structure of monomeric αS.[81] Double-labeled αS for smFRET experiments was produced by sequential labeling of engineered Cys residues at positions 7 and 84 with maleimide-conjugated Alexa Fluor dyes. In buffer, the average E_{FRET} value measured was 0.47,

corresponding to a distance of approximately 61 Å. The measured E_{FRET} between the two labels decreased as a function of TMAO concentration (Figure 1.17). At 4 M TMAO, the average E_{FRET} value measured was 0.86, which corresponds to a distance of approximately 44 Å. These results corroborate earlier findings suggesting that TMAO induces folding of αS.

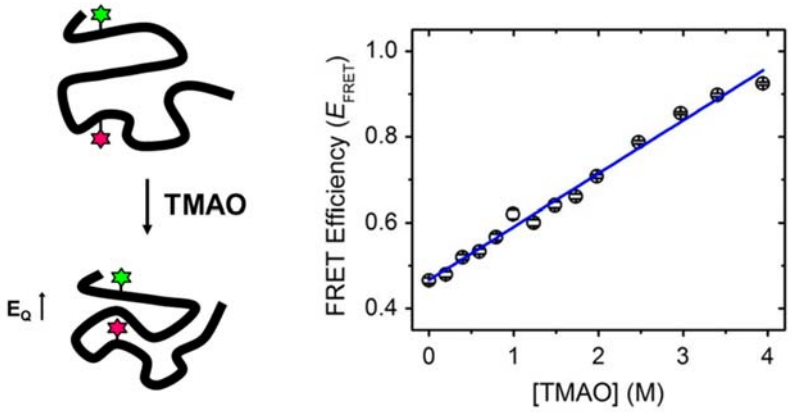

Figure 1.17 Compaction of αS in TMAO.
Left: Cartoon representation of αS compaction in the presence of TMAO. Right: FRET efficiency as a function of TMAO concentration. Figure adapted from Ferreon *et al*.[81]

In a subsequent publication, Moosa *et al.* performed a more rigorous titration of TMAO with double-labeled αS.[82] A plot of FRET efficiency as a function of TMAO concentration revealed a sigmoidal transition profile suggestive of a two-state folding process. In addition to monitoring folding by smFRET, the authors of this study performed CD spectroscopy measurements, which were also consistent with a cooperative two-state folding transition to a helical structure. Although it is unknown whether the TMAO-induced structure is related to the native conformation of αS *in vivo*, these studies demonstrate that αS can be induced to fold into a compact helical structure in the absence of lipids.

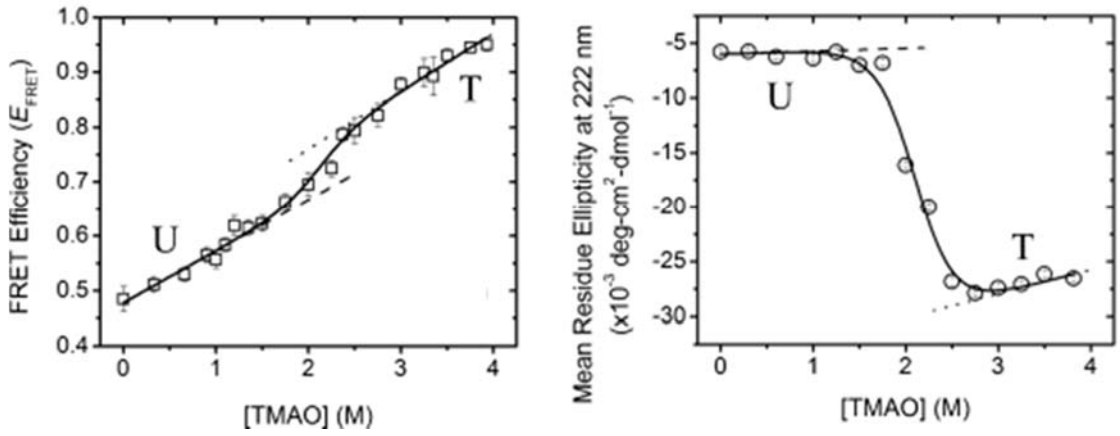

Figure 1.18 TMAO-Induced Two-State Folding of αS.
Left: FRET efficiency as a function of TMAO concentration for to a two-state linear extrapolation model. Right: TMAO-induced folding measured by CD spectroscopy fit to a two-state linear extrapolation model. U represents the unfolded protein and T represents the protein in the presence of TMAO. Figure adapted from Moosa *et al.*[82]

§1.8 Structural Studies of α-Synuclein Fibrils.

The observation that LBs consist of an insoluble, filamentous form of αS prompted numerous investigations aimed at characterizing the molecular details of this structure. Early transmission electron microscopy (TEM) studies revealed that purified recombinant αS readily assembles into fibrils with overall morphologies that are similar to those extracted from PD patients post-mortem.[83] By TEM, αS fibrils appear long (1-10 μm), straight, and unbranched with an average diameter of roughly 5-10 nm. CD measurements revealed that αS undergoes a conformational change from random coil to β-sheet upon aggregation. X-ray and electron fibril diffraction patterns confirmed that αS fibrils consist of a β-sheet structure and that the individual β-strands are stacked perpendicular to the long fiber axis. [84] High-resolution cryoelectron microscopy images revealed that αS filaments are comprised of two individual fibril strands that assemble into both straight and twisted structures. Whereas the strands of the fibrils in the straight

filaments are separated by 1.5 nm, the fibrils in the twisted filaments are separated by up to 3 nm at their widest points.[85]

A variety of experimental approaches have been taken to obtain more detailed information on the fold of αS fibrils. In 2003, Der-Sarkissian *et al*. investigated the structural organization of αS fibrils using electron paramagnetic resonance (EPR) spectroscopy.[86] Like PRE, EPR studies require the incorporation of a nitroxide spin label. Whereas PRE experiments are based on the relaxation rates of nuclei that are in close proximity to the nitroxide probe, EPR measurements directly report on the structural environment of the label itself. In this initial study, fibrils were prepared from a total of 36 spin-labeled αS mutants. Analysis of the obtained EPR spectra suggested that the core region of the fibrils consists of 68 amino acids spanning residues 34 to 101. These data were consistent with proteolytic digestion studies suggesting that the core region extends into the N-terminal region beyond the hydrophobic NAC domain.[87] A thorough analysis of intermolecular spin-spin interactions between adjacent nitroxide labels revealed that β-sheets in the core region are organized in an in-register, parallel fashion. Although the fibril core appears to be well-folded by EPR, the N-terminal region showed a heterogeneous, less-ordered structure, and the C-terminal region was determined to be entirely disordered. In a subsequent study, Chen *et al*. prepared a library of 83 spin-labeled αS mutants for EPR studies.[88] In addition to confirming the previous findings, the authors were able to detect flexible regions within the tightly-packed core that likely correspond to loop regions between neighboring β-strands.

Several groups have studied the structure of αS fibrils using advanced NMR techniques. In 2005, Heise *et al*. reported the first solid-state NMR-derived structure of

the αS fibril core.[89] In order to obtain resonance assignments corresponding to specific amino acid residues, the authors performed a variety of multidimensional cross-correlation experiments using isotope-labeled αS. Overall, the results of the NMR experiments were in good agreement with the EPR studies. Residues 39–95 were determined to comprise the β-sheet rich fibril core, whereas the C-terminal tail was characterized as disordered. Although seven distinct β-strands were identified, their overall topology was not described.

In 2008, Vilar *et al.* performed hydrogen/deuterium (H/D) exchange as well as solid-state NMR experiments to further characterize the fold of αS fibrils.[85] Backbone amide protons can readily exchange with solvent deuterons provided that they are sufficiently solvent exposed. Otherwise, the process of H/D exchange occurs slowly (seconds compared to hours). H/D exchange experiments revealed that the amide residues between amino acids 30-110 are protected from solvent, whereas residues at both termini were determined to be solvent exposed. Within the identified core, five segments of amino acid residues were found to undergo H/D exchange more slowly than any other observed regions. Taking these results as well as previous EPR studies into account, the authors proposed that the fibril core consists of five distinct β-strands connected by small loops that fold into a β-sandwich (Figure 1.19).

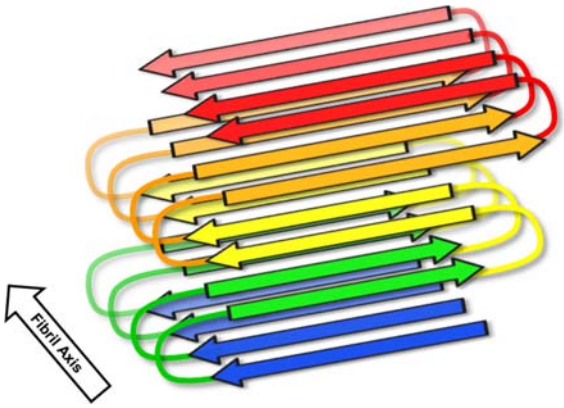

Figure 1.19 Proposed Fold of αS Fibrils.
The fibril core is predicted to consist of at least five β-strands (shown as individual colors) connected by small loops that fold into a β-sandwich.

Since these initial reports, other groups have used solid-state NMR to investigate the structure of αS fibrils.[90,91] Although the details of each reported structure vary, each study predicts the formation of a core region that contains at least five β-strands.

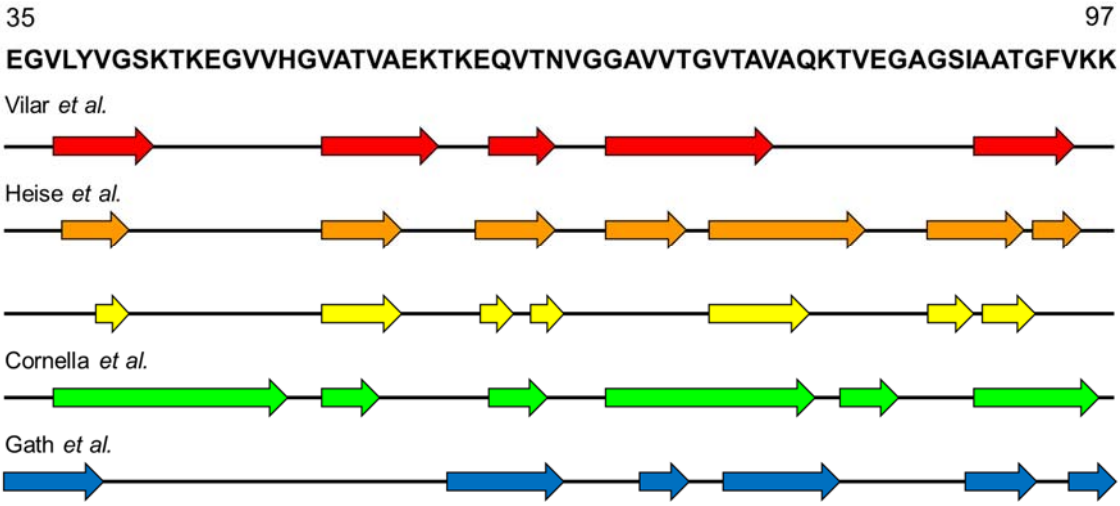

Figure 1.20 Summary of Solid State NMR Measurements of αS Fibrils.
Top: Amino acid residues residing within the fibril core. Bottom: Arrows represent β-strand regions identified by solid state NMR spectroscopy. Note that is Heise *et al.* two different fibril forms (orange and yellow) were identified. Figure adapted from Pornsuwan *et al.*[92]

Recently, Pornsuwan *et al.* carried out a series of double electron-electron resonance (DEER) experiments with dually spin-labeled mutants of αS to ascertain the distances between individual β-strands.[92] These preliminary inter-strand measurements

suggest that the overall fold may be more complex than the predicted β-sandwich model. Although EPR and NMR studies have helped to shed light on the fold of αS fibrils, neither technique has been used to describe structural details involving the termini. Moreover, since these measurements can only be obtained at the end of the aggregation process, they cannot be used to resolve conformational changes that occur in real-time.

§1.9 Structural Studies of α-Synuclein Oligomers.

Although fibrils are known to induce pathology, an increasing number of reports suggest that soluble oligomeric intermediates of αS are toxic to neurons. However, the transient nature of these heterogeneous oligomers renders them exceedingly difficult to isolate and study. Although several different methods have been reported for generating αS oligomers *in vitro*, none have been proven to produce species that are physiologically relevant. Regardless, several groups have examined their structure using a variety of experimental techniques.

Apetri *et al*. examined the structural features of oligomeric αS using Raman and atomic force microscopy (AFM).[93] In this study, αS oligomers were produced by simply allowing the protein to incubate at 37 °C without agitation for several weeks. After 21 days, spheroidal oligomers ranging from 1.4 – 7.5 nm in height were observed by AFM. Spheroidal oligomers do not enhance the fluorescence of ThT, indicating that these intermediates lack the crossed β-sheet structure that is characteristic of amyloid fibrils. Longer periods of incubation led to the formation of filamentous αS (Figure 1.21). In order to gain insight into the changes in secondary structure accompanying the formation of spheroidal oligomers, Raman microscopy measurements were performed at each stage of the aggregation process. Upon oligomerization, the observed Raman amide I band

profiles were consistent with increasing β-sheet content and decreasing extended β-strand and polyproline II structure.

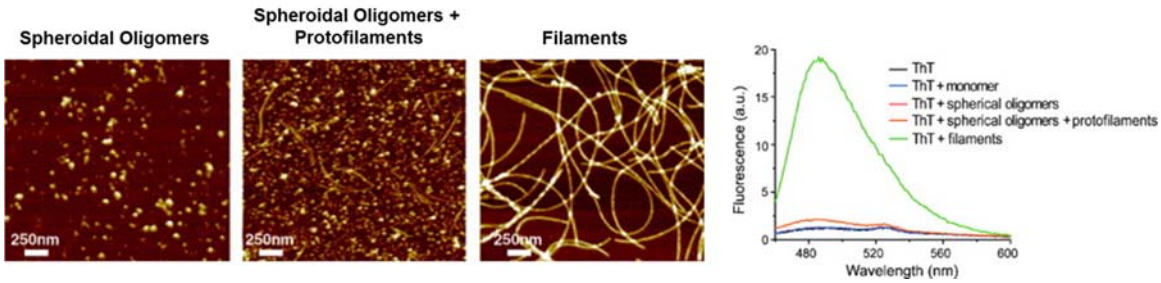

Figure 1.21 AFM Images of αS at Various Stages of Aggregation.
Left: AFM images obtained at 21 (left), 32 (middle), and 42 (right) days of incubation at 37 ºC. Right: ThT binding assays corresponding to each stage of the aggregation process demonstrate that spheroidal oligomers lack amyloid content. Figure adapted from Apetri *et al.*[93]

In order to probe the conformation of αS oligomers using fluorescence spectroscopy, Rooijen *et al*. prepared αS mutants containing single Trp residues (F4W, Y39W, A69W, A90W, A124W, and A140W).[94] Solvent exposed Trp residues typically emit from 340 – 355 nm. In a more hydrophobic environment, Trp emission undergoes a blue-shift to yield a maximum closer to 335 nm. Stable oligomers were prepared by re-suspending dried monomeric αS in buffer. Oligomers formed from F4W, Y39W, A69W, and A90W exhibited significantly blue-shifted Trp fluorescence in comparison to each respective monomer (~348 nm to ~337 nm). No blue-shift in Trp emission was detected for oligomers formed from A124W and A140W, indicating that the C-terminus of the protein remains solvent exposed. Upon binding to lipid vesicles, a similar pattern in Trp fluorescence was observed. Overall, these studies suggest that the N-terminal and NAC domains constitute the folded core of the spheroidal oligomer.

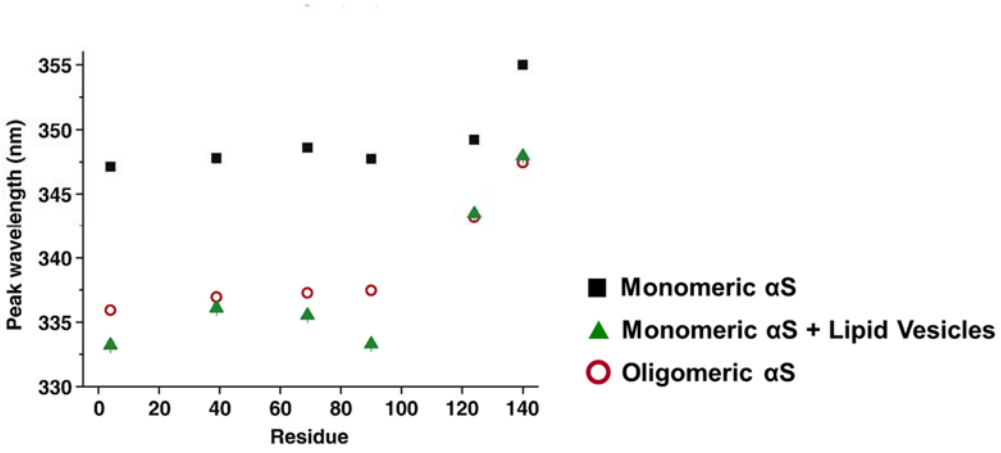

Figure 1.22 Spectral Analysis of αS Trp Mutants.
The observed emission maximum of αS was plotted as a function of the position of the inserted Trp residue. Figure adapted from Rooigen *et al.*[94]

Recently, Mysling *et al*. analyzed the structure of αS oligomers using H/D exchange monitored by mass spectroscopy (HDX-MS).[95] In this work, oligomeric αS was isolated using size-exclusion chromatography. First, H/D exchange was performed by diluting oligomeric αS into D_2O. Following the exchange reaction, the oligomers were digested with pepsin, and the resultant peptide fragments were analyzed using LC/MS. Efficient backbone exchange was observed for the C-terminal domain (residues 94-140), consistent with Trp fluorescence studies suggesting that the acidic tail remains solvent exposed in αS oligomers. In the remainder of the protein, three regions that were strongly protected from H/D exchange (residues 4-17, 39-54, and 70-89) were interspersed by two relatively more dynamic segments (residues 18-38 and 55-76). These results are consistent with the notion that intermolecular associations between αS monomers are driven by hydrophobic amino acids residing in the NAC domain (residues 65-90).

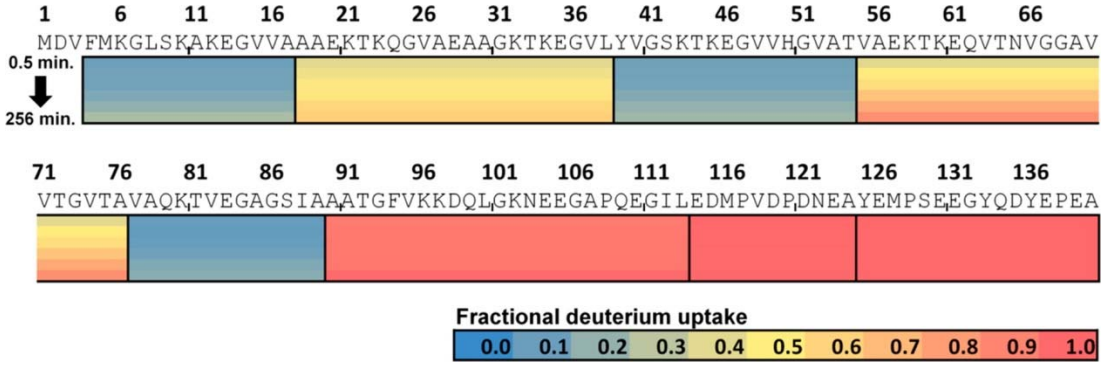

Figure 1.23 Heat Map of Local H/D Exchange Patterns Observed in αS Oligomers. Peptides are color-coded according to fractional deuterium uptake. Figure adapted from Mysling *et al.*[95]

Solid-state NMR measurements have shown that the N-terminus is conformationally dynamic and solvent exposed in αS fibrils. However, both Trp fluorescence and HDX-MS experiments suggest that the N-terminus of αS resides in a less polar, less solvent exposed environment in αS oligomers. Thus, it is possible that structural rearrangements of the N-terminal domain play a critical role in the maturation of αS fibrils.

§1.10 Monitoring α-Synuclein Dynamics using Fluorescence Spectroscopy.

It is well established that misfolding of αS leads to the formation of β-sheet rich fibrils, but the molecular mechanisms underlying the conversion of αS into neurotoxic species remain poorly understood.[96] Fluorescence spectroscopy is a highly sensitive technique that is ideally suited for monitoring dynamic protein behavior *in situ*. Several fluorescence-based approaches have been applied to the study of αS. Commonly used amyloid dyes, such as Thioflavin T (ThT), can be used to detect the presence of αS fibrils. ThT is a small benzothiazole dye that exhibits a dramatic increase in fluorescence upon binding to amyloid fibrils. In solution, the emission of ThT is quenched by a process known as twisted intramolecular charge transfer (TICT). TICT occurs when

rotation about the central C-C bond orients the benzothiazole and dimethylaminobenzene rings perpendicular to one another. Upon binding to fibrils, this bond can no longer rotate, resulting in a substantial increase in the quantum yield of the dye.[97] Although ThT and other extrinsic dyes are routinely employed to monitor the kinetics of amyloid formation, they cannot be used to yield structural details of this process at the residue level. Moreover, the most commonly used extrinsic probes are insensitive to the formation of early αS oligomers.

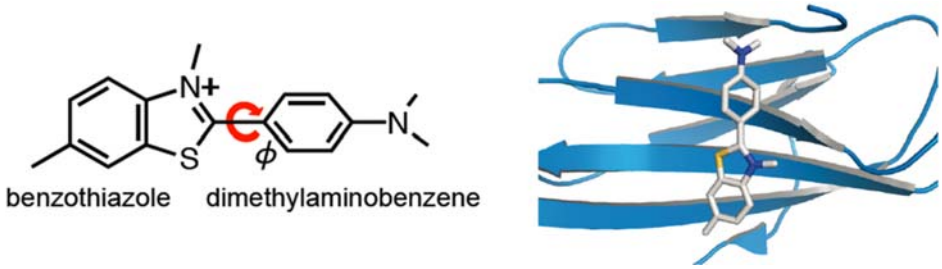

Figure 1.24 Thioflavin T Binds to Amyloid Fibrils.
Left: Chemical structure of Thioflavin T. Rotation about the central C-C bond leads to twisted intramolecular charge transfer. Right: Structure of Thioflavin T bound to amyloid fibrils comprised of β-2 microglobulin (PDB Code 3MYZ).

In order to directly monitor αS dynamics using fluorescence spectroscopy, a fluorescent moiety must be installed into the protein. Several groups have incorporated environmentally-sensitive fluorophores into αS for monitoring the various stages of aggregation. Dusa *et al*. introduced a Trp mutation into position 39 to generate an intrinsic fluorescence probe for monitoring αS dynamics.[98] Upon aggregation, Trp fluorescence increased with a rate that was similar to ThT binding. Curiously, the emission maximum was similar in the monomeric and fibrillized forms (346-355 nm), indicating that residue 39 remains somewhat solvent exposed upon aggregation. This result is surprising in light of NMR and EPR studies predicting that residue 39 is embedded within the fibril core.

In order to incorporate unnatural fluorescence probes into αS, several groups have turned to site-directed labeling of engineered Cys residues. In 2008, Thirunavukkuarasu *et al*. prepared three mutants of αS labeled with a pyrene group at positions 18, 90, and 140.[99] The high extinction coefficient, long lifetime, and solvent-sensitivity of pyrene render it a particularly useful probe for studying αS dynamics. Additionally, pyrene groups that are in close spatial proximity have a tendency to form pi-stacked complexes; upon electronic excitation, these complexes can form excimers that exhibit unique emission bands. Since near-molecular contact is required for excimer formation to occur, it can be used to evaluate short distances within or between proteins. Over the course of aggregation, several parameters of pyrene fluorescence were examined including steady-state, lifetime, and anisotropy measurements. Together, these experiments allowed the authors to detect the formation of early αS oligomers. Intriguingly, the emission properties of pyrene-labeled αS containing familial PD point mutations were different from those of labeled wild-type. Although this observation is interesting, the authors were unable to rationalize the origin of these distinctive features.

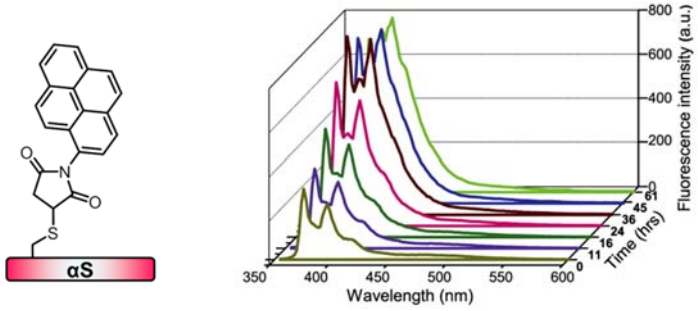

Figure 1.25 Site-Specific Pyrene Labels for Monitoring Aggregation.
Left: Cys-containing labeled with pyrene maleimide. Right: Example of aggregation monitored by pyrene fluorescence. Figure adapted from Thirunavukkuarasu *et al*.[99]

In 2010, Yushchenko *et al*. introduced 3-hydroxychromones (3HCs) as probes for monitoring αS oligomerization.[100] Upon excitation, 3HCs undergo a process known as

excited-state intramolecular proton transfer (ESPIT) that results in emission from two chemically distinct excited states referred to as the normal excited state N* and the ESPIT phototautomer product T*. The ratio (T*/N*) of the emission band intensity is extremely sensitive to the polarity and hydrogen bonding capacity of the surrounding environment. Therefore, the authors reasoned that 3HCs are uniquely well-suited for detecting changes in the local microenvironment of αS residues during early stages of its aggregation. In this study, αS was labeled with 2-(2-furyl)-3-hydroxychrome (FC) at position 140. During aggregation, the FC emission band ratio underwent a dramatic change that was consistent with the probe moving into a less protic and less polar environment. Kinetics measured by the change in FC fluorescence preceeded those obtained using ThT, indicating that FC was more sensitive for detecting the formation of αS oligomers. Although the C-terminal domain of αS is characterized as conformationally-mobile and solvent exposed, this study suggests that its local environment is markedly different in fibrils.

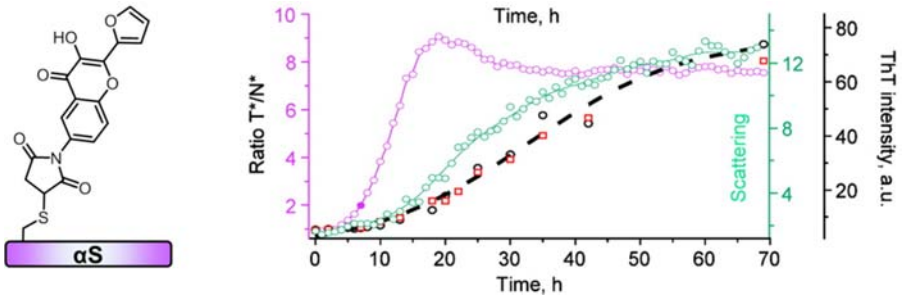

Figure 1.26 Ratiometric Hydroxychomone Probe for Monitoring αS Aggregation.
Left: Cys-containing αS labeled with 2-(2-furyl)-3-hydroxychomone maleimide (MFC). Right: Aggregation of αS_{140}MFC. Figure adapted from Yushchenko *et al.*[100]

In order to detect structural changes involving the N- and C-termini upon aggregation, Yap *et al.* generated a small library of αS mutants labeled with an environmentally sensitive dansyl (Dns) probe.[101] Aggregation was monitored with

mutants labeled at the termini (7 and 136) as well as with mutants labeled at sites next to the predicted fibril core (26 and 100). Upon aggregation, the fluorescence of each Dns probe changed in a manner that was consistent with being transferred into a more hydrophobic environment. Surprisingly, kinetic studies revealed that changes in Dns fluorescence at the termini preceded those occurring at sites closer to the fibril core. In order to determine whether this effect was nucleation-dependent, these assays were repeated in the presence of a wild-type αS seed, and a similar trend was observed (Figure 1.27). Together, these data suggest that the first steps in αS fibrillization involve conformational rearrangements occurring at the N- and C-termini.

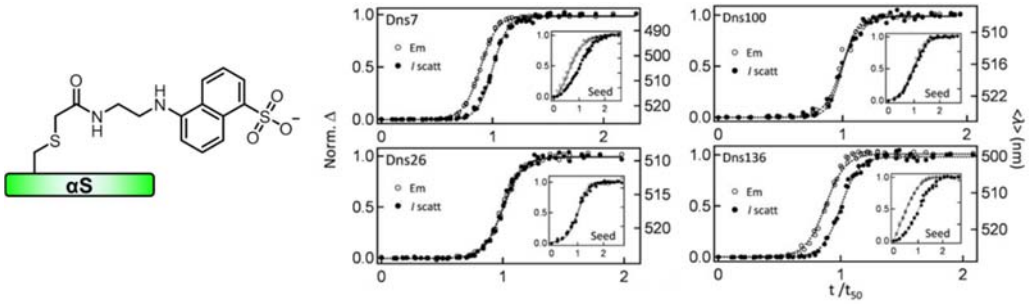

Figure 1.27 Site-Specific Dansyl Labels for Monitoring αS Aggregation.
Left: Cys-containing αS labeled with dansyl probe. Right: Aggregation kinetics monitored by Dns fluorescence and light scattering.

Although environmentally-sensitive fluorophores are useful for monitoring aggregation, the information obtained from these experiments is limited. Alternatively, FRET can be employed to define the structural details regarding the oligomeric conformations of αS. Whereas smFRET experiments can be performed using mixtures of doubly-labeled proteins, ensemble FRET measurements require the preparation of homogeneous site-specifically labeled material. Due to the challenging nature of producing doubly-labeled αS containing optimal labels for FRET studies, there are only a few reports in the literature describing the use of this technique at the ensemble level.

Perhaps the simplest way to produce homogeneous doubly-labeled proteins for FRET studies is to take advantage of intrinsic protein fluorescence. In 2005, Kaylor *et al*. generated a mutant of αS containing a single Tyr donor and Trp acceptor (Y125W/Y133F/Y136F) at positions 39 and 125, respectively.[102] Significant energy transfer from Tyr to Trp was observed in monomeric αS, indicating that these regions are separated by an average 25 Å distance in the disordered ensemble. Over the course of aggregation, an increase in FRET was observed in the soluble fractions, suggesting that early oligomers adopt more compact conformations. However, since the authors did not dilute the doubly-labeled protein into a label-free mixture (such as an all Phe mutant), it is not possible to distinguish inter vs. intramolecular FRET interactions from the obtained spectra. Although FRET was detected at the oligomeric stage, energy transfer between Tyr and Trp was not observed in the fibrillar state. The authors attribute this finding to a possible lack of FRET between the pair, but they also note that Tyr emission is significantly quenched at position 39 in fibrils lacking the Trp acceptor. Despite these areas of concern, this report is still often cited as evidence for the compaction of αS upon aggregation.

In 2010, Nath *et al*. prepared αS containing a red-shifted donor/acceptor probe pair for FRET studies.[103] First, a mutant of αS containing an engineered Cys residue at position 140 was labeled with tetramethylrhodamine (TMR) maleimide. Since the pKa of the N-terminus is lower than that of lysine (~ 7.7 vs ~ 10.5) semi-selective modification of the N-terminus can be achieved at low pH (~ 6.5). Following an initial round of purification, the singly labeled protein was reacted with Alexafluor-488 succinimidyl ester to afford the doubly-labeled construct. However, the authors note that the

implemented labeling strategy afforded heterogeneous material; absorbance measurements revealed that the donor/acceptor ratio in the final product was nearly 2:1. Additionally, fluorescence correlation spectroscopy measurements revealed that oligomers had formed during the course of the labeling reactions. Despite these issues, intramolecular misfolding events were monitored by diluting the labeled construct into a molar excess of unlabeled protein. Steady-state and time-resolved fluorescence measurements suggested that compaction of the termini increased upon further oligomerization. However, due to the heterogeneity of the sample, the authors were unable to use FRET to quantify the actual change in distance that occurred.

§1.11 Conspectus

Distance-dependent fluorophore interactions such as FRET and PET are well-suited for studying the misfolding of αS. However, the production of homogeneously double-labeled αS is a challenge in its own right. Furthermore, in order for the labels to be useful, they must not perturb the misfolding and/or aggregation process. The Petersson laboratory develops new fluorescence-based technologies to study the dynamic behavior of proteins using a combination of synthetic chemistry, molecular biology, and biophysical techniques. In this thesis, I will describe how we have applied our unique, minimally perturbing fluorescence probe – the thioamide – to the study of protein folding. Next, I will discuss our development of novel semi-synthetic strategies for installing thioamide/fluorophore probe pairs into full-length αS for folding studies. Finally, I will describe alternative methods for producing labeled αS for various aggregation and disaggregation experiments. Together, these studies have generated unique insight into the conformational dynamics of αS that will be described in detail herein

Chapter 2 : Applications of Thioamide Quenching of Intrinsic Protein Fluorescence to Studies of Folding and Binding

§ 2.1 Thioamide Quenching of Protein Fluorescence

Intrinsic protein fluorescence, which chiefly arises from excitation of tryptophan (Trp) or tyrosine (Tyr) residues, has been used extensively to probe protein folding, conformational rearrangements, and ligand binding.[6] Many of these studies employ distance-dependent quenching of Trp and Tyr through energy transfer to an extrinsic probe, but data interpretation in these experiments is often limited by the size of the acceptor chromophore. If a chromophore were sufficiently compact to be incorporable at any position in a protein without severely perturbing native structure, the structural resolution of these experiments could be greatly improved. Here, we show that a thioamide – a single atom substitution of the peptide backbone - can quench Trp and Tyr fluorescence in a distance-dependent fashion. We also show that this technique can be used to monitor protein/protein interactions *in vitro*.

Thioamide Properties. Thioamides are nearly isosteric with the natural oxoamides found in the peptide backbone. However, this single-atom substitution results in subtle differences between the two functional groups. Sulfur has a larger van der Waals radius than oxygen (1.85 Å vs. 1.40 Å) and is considerably less electronegative (2.58 vs 3.44).[104] The thiocarbonyl bond is longer than the oxocarbonyl bond (1.71 Å vs. 1.23 Å), whereas the carbon-nitrogen (C-N) bond is slightly shorter in a thioamide (1.35 Å vs 1.37 Å).[105] The NH pK_a is substantially lower for the thioamide than the oxoamide

(12 vs 17).[106] Additionally, the barrier to rotation about the C-N bond is higher for thioamides (22 kcal/mol vs. 17 kcal/mol).[107]

In order to rationalize these properties, Wiberg *et al*. have argued that the relative contribution of resonance structure B is greater for thioamides than for oxoamides (Figure 2.1).[107] Since oxygen is more electronegative than sulfur, dipolar structure B may seem more important for oxoamides. However, computational electron density maps revealed greater transfer of π electron density from N to S in thioformamide than from N to O in formamide. It can therefore be reasoned that the energetic penalty for further polarization due to charge transfer from nitrogen is higher for oxoamides than for thioamides. Furthermore, the larger size of sulfur allows it to better accommodate negative charge, and weakens overlap with carbon p orbitals in the C=S bond.

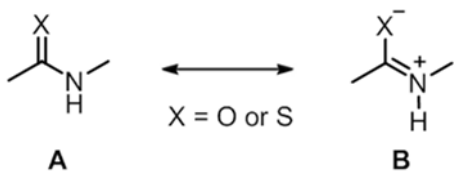

Figure 2.1 Dipolar Resonance Structure of Oxoamides and Thioamides.
The relative contribution of resonance structure B is greater for thioamides than for oxoamides.

In general, the thioamide NH is a stronger hydrogen bond donor than the oxoamide NH, while the sulfur is a slightly weaker hydrogen bond acceptor than the corresponding oxygen.[108,109] Although thioamides are weaker hydrogen bond acceptors, they are superior electron pair donors. Raines and coworkers have shown that thioamide interactions may have stabilizing effects via n-π* interactions (in which a thiocarbonyl sulfur donates lone pair electron density into another carbonyl group) in certain turn geometries.[110] Despite the larger size of sulfur, the longer length of the thiocarbonyl bond, and the potential for altered stabilizing and/or destabilizing interactions, it has been

shown that a single thioamide substitution incurs only minor perturbation to peptide secondary structure (discussed below).

The unique properties of the thioamide that are most relevant to our spectroscopic studies are the red shift in the carbonyl π - π * absorption and the lower oxidation potential of the thioamide.[111,112] The change in the absorption maximum to ~ 270 nm allows the thioamide to have spectral overlap resulting in FRET interactions with fluorophores such as Cnf and Tyr. The ~ 2 eV change in oxidation potential allows the thioamide to be competitive in fluorophore PET quenching processes with other easily oxidized natural and unnatural amino acids.[113]

Chemical Synthesis of Thiopeptides. Thioamide-containing peptides can be synthesized on solid phase using pre-activated thioamide precursors and standard Fmoc-chemistry procedures.[114] Fmoc-protected thiocarboxybenzotriazoles can be generated from the corresponding Fmoc-protected amino acid using 4-nitro-1,2-phenylenediamine (nitro route) or *N*-Boc-phenylenediamine (Boc route) (Figure 2.2). In general, the nitro route is preferable for several practical reasons. The presence of the electron-withdrawing nitro group leads to deactivation of the *p*-amino group. As a result, the initial coupling of the amino acid proceeds in a regioselective manner. Since no additional protecting group is required, the overall scheme requires fewer steps. Additionally, the nitro route is compatible with amino acids bearing acid-labile protecting groups. Although generally lower-yielding, an important advantage of the Boc route is that the final benzotriazole products are more stable and can potentially be stored for longer periods of time. To date, our group has reported successful thioamide incorporation from benzotriazole precursors for the following amino acids: Ala, Asp, Glu, Leu, Phe, and Val.[115-118] We have also

demonstrated that Arg, Ile, and Pro can be similarly installed into peptides (unpublished results). Conceivably, a thioamide analog of any of the natural amino acids can be incorporated into peptides using appropriate protecting group manipulations.

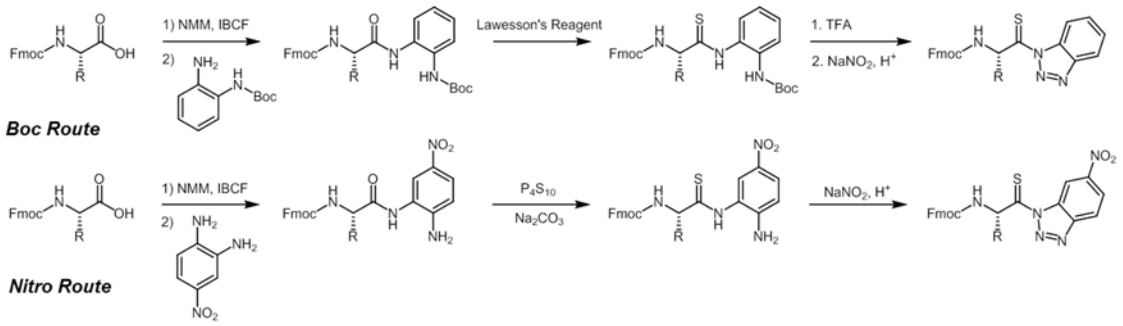

Figure 2.2 Synthesis of Thioamide Precursor Compounds.

Thioamide Effects on Protein Folding. Several independent studies have examined the effect of thioamide substitution on the thermodynamic stability of α-helices and β-sheets. In α-helices, thioamide substitution at the C-terminus (within the last four residues) is expected to be tolerated, as the thiocarbonyl would not participate in hydrogen bonding at these sites. Other positions within a tightly packed α-helix were initially proposed to be unable to accommodate the increase in the (thio)amide-amide hydrogen bond distance or the larger size of the sulfur atom.[119] In order to examine the effect of thioamide substitution within α-helical structures, Miwa *et al.* prepared thioamide analogs of GCN4, a 35-residue helical peptide that dimerizes to form a parallel, two-stranded coiled-coil.[120] Circular dichroism (CD) measurements confirmed that thioamide substitution (at either the C-terminus or in the middle of the helix) yielded structures that were similar to those formed by the native peptide. The melting temperatures (T_ms) of the folded coiled-coils were derived by measuring the temperature-dependent change in the CD signal corresponding to helical content. A slight increase in T_m was observed for the peptide containing a thioamide near the center of the helix,

suggesting that the increase in hydrogen bond strength conferred by the thioamide NH may enhance thermal stability in a position-dependent manner. In a separate study, Fischer, Kiefhaber, and coworkers introduced single thioamide substitutions at a central or an N-terminal position in an alanine-based model helical peptide.[121] CD measurements revealed that incorporation of an Ala′ residue at either location led to helix destabilization similar in magnitude to that observed for a Gly substitution at each respective position (roughly 1.7 kcal/mol). It should be noted that this study differs significantly from the Miwa α-helix study in that the dimeric nature of GCN4 may also be influenced by thioamide incorporation, whereas the artificial α-helices in this study are monomeric and may provide a more direct assessment of the impact of thioamides on the α-helix itself. Taken together, these results demonstrate that the effect of thioamide incorporation on α-helix stability may ultimately be dependent on the position in question.

To determine whether a thioamide linkage is compatible with β-sheet secondary structure, Miwa *et al*. prepared a thioamide-containing β-hairpin peptide, comprised of two antiparallel strands joined by a type II′ β-turn.[122] To minimize perturbation of the hairpin structure, the thioamide residue was inserted into the $i + 2$ position of the β-turn. In this position, the thiocarbonyl points toward the exterior of the hairpin, while the thioamide NH can still participate in an interstrand hydrogen bond. NMR and CD experiments confirmed that the hairpin conformation and stability were similar to that of the corresponding oxoamide peptide. More recently, Gai, Degrado, and coworkers used thioamides to site-specifically examine the role of interstrand hydrogen bond formation in the folding of a stable monomeric β-hairpin known as tryptophan zipper (Trpzip). [123] In these studies, thioamide residues were placed in a manner such that the thiocarbonyl

would be forced to participate in backbone-backbone hydrogen bonding throughout the β-hairpin sequence. Analysis of CD thermal melting curves for the mutant peptides revealed that the degree of β-hairpin destabilization depended on the position of the interstrand hydrogen bond being perturbed, with the effect far less pronounced at the terminal region of the hairpin. To further examine the mechanistic basis of thioamide-induced destabilization, the folding and unfolding rates of the mutant β-hairpins were determined by measuring their relaxation rates in response to a laser-induced temperature-jump using time-resolved IR spectroscopy. Thioamide incorporation at the turn region slowed down the rate of folding, whereas incorporation at the terminal region dramatically increased the rate of unfolding. Similar to observations in thioamide-containing α-helices, the effect of thioamide substitution into β-sheets is likely to be dependent on the position of incorporation and on whether the residue is involved in the formation of a key hydrogen bond in the native sequence or a folding intermediate.

Although thioamide substitution has been investigated in model peptides, there are few examples describing how thioamide incorporation into full-length proteins affects their ability to fold or function. To date, only one protein is known to naturally contain a thioamide linkage. Methyl-coenzyme M reductase, an enzyme that contains a thioglycine residue near its active site, catalyzes the final methane-forming step in methanogenic archaea.[40, 41] It has been proposed that the embedded thiocarbonyl moiety may work as a redox active one-electron relay system to facilitate the oxidation and reduction of cofactor and substrate.[124] The X-ray crystal structure of Methyl-coenzyme M reductase has been determined, providing insight as to the geometry that thioamide residues adopt in complex folded proteins. The Gly′ residue is not located in a canonical α-helix or β-

sheet. The thiocarbonyl length is extremely long (1.8 Å) and the C-N bond length is very short (1.3 Å), consistent with the predominance of resonance structure B as a representation of the thioamide. As one might expect, the hydrogen bond to sulfur is unusually long (3.5 Å) as compared to a typical backbone hydrogen bond in a protein (~ 3.0 Å). However, no increase in thioamide NH hydrogen bond donor strength is apparent, as the NH hydrogen bond is not unusually short (3.1 Å). Choudhary *et al.* note that the thiocarbonyl group engages in a strong n-π* interaction with the side-chain carbonyl of a nearby asparagine residue.[125] However, since Methyl-coenzyme M reductase has presumably evolved to take advantage of this specialized Gly′ residue, it is not clear that this information has general implications for the introduction of thioamides into existing secondary structures.

Fischer and coworkers have been able to demonstrate that a thioamide can be chemically installed into a different enzyme without perturbing its inherent activity.[126,127] Ribonuclease (RNase) S, originally generated by cleaving RNase A into two fragments using subtilisin, is comprised of the separated S-peptide (residues 1 to 20) and the S-protein (residues 21 to 124). In solution, the S-protein is inactive but folded, whereas the S-peptide is disordered. Upon binding to the S-protein, residues 3-13 of the S-peptide adopt an α-helical structure, yielding a catalytically active bimolecular complex.[128] To monitor backbone hydrogen bond formation in the transition state of the S-peptide/S-protein association, single thioamide substitutions were introduced into different positions between residues 1 and 15 of the S-peptide. All of the S-peptide mutants bound to the S-protein to generate catalytically active complexes. However, isothermal titration calorimetry (ITC) experiments revealed that the thermodynamic stability of each mutant

RNase S complex was reduced to varying extents, ranging from 0.6 to 4.7 kcal/mol. These results indicate that some thioamide substitutions can be tolerated in full-length proteins without grossly perturbing their native state or function, but that one must use caution in their placement.

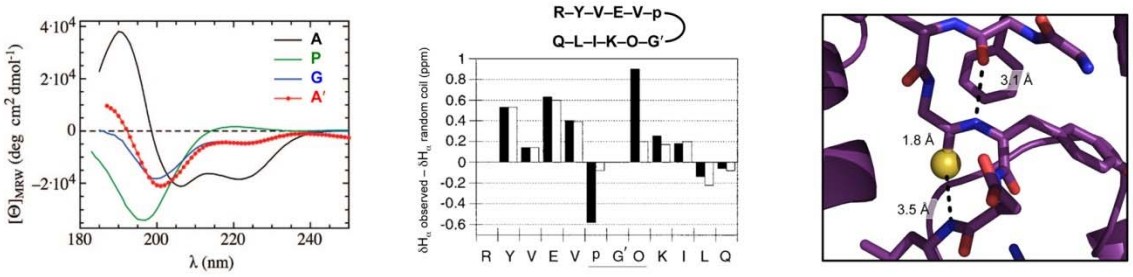

Figure 2.3 The Impact of Thioamide Substitution on Protein Folding.
Left: Circular dichoism (CD) analysis of α-helical peptide sequences by Reiner *et al*. showed that Ala′ substitution can be comparably destabilizing to glycine substitution. Middle: NMR analysis of a α-turn sequence by Miwa and coworkers showed that Gly′ was compatible with turn formation (p indicates D-proline). Right: The geometry of a hydrogen-bonded Gly′ residue can be seen in the sole structurally-characterized example of a full-length protein containing a thioamide residue, Methyl-coenzyme M reductase (PDB ID 1MRO).Figures adapted from references 116 (left), 124 (middle) and 125 (right).

Thioamide Quenching of Cnf. In 2010, our laboratory reported that thioamide quenching of the unnatural amino acid *p*-cyanophenylalanine (Cnf or F*) can be used to monitor protein folding.[115] Cnf was chosen for these initial experiments because of its large extinction coefficient (ε = 13,000 M^{-1} cm^{-1} at 240 nm) and substantial spectral overlap with thioamides. The theoretical value of R_0 was calculated to be 15.6 Å, indicating that the thioamide/Cnf FRET pair can be used to measure distance changes in the 8 to 32 Å range. In order to experimentally verify the distance-dependence of this interaction, a series of peptides was synthesized that contained thioleucine (Leu´) at the N-terminus and Cnf at the C-terminus, separated by an increasing number of proline residues. Quenching efficiency was determined as

$$E_Q = 1 - F_{Thio} / F_{Oxo} \qquad \text{(Eq. 2.1)}$$

where F_{Thio} is the fluorescence intensity of a given Leu´Pro_nCnf peptide, and F_{oxo} is the fluorescence of the oxoamide control peptide $LeuPro_2Cnf$. By plotting the quenching efficiency as a function of distance (R) and fitting the data to a $1/R^6$ function, the half-maximal change in fluorescence was determined to occur at a distance of 16.5 Å. Although this result is in good agreement with the calculated prediction based on Förster theory, it is also possible that Dexter energy transfer, a fluorescence quenching mechanism that depends on direct overlap of donor and quencher molecular orbitals, may contribute to Cnf quenching at short distances.[10] Dexter transfer has a $1/e^R$ distance dependence, which also provides a good fit to the Pro series data. Therefore, it is likely that both mechanisms contribute to thioamide quenching.

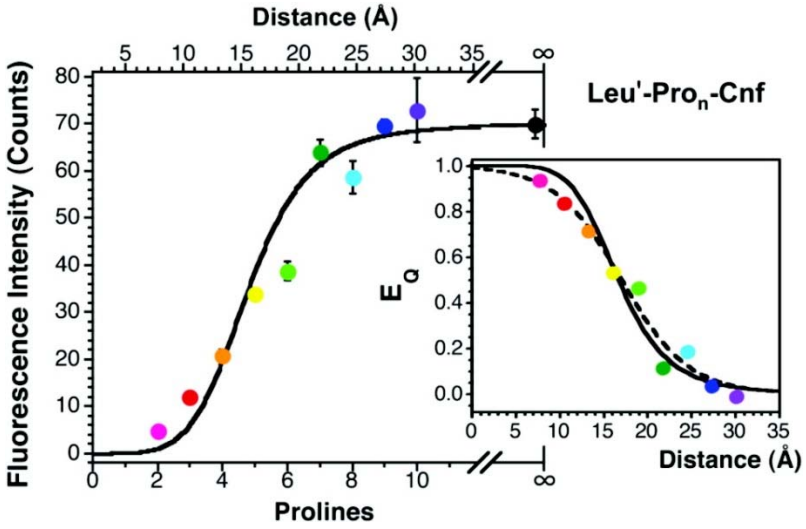

Figure 2.4 Fluorescence Emission as a Function of Chomophore Spacing.
The fluorescence emission of Leu′-Pro_n-Cnf (n = 2−10) at 293 nm is shown. The "∞" data point indicates the fluorescence of Leu-Pro_2-Cnf. The solid line indicates the distance dependence predicted by Förster theory with R_0 = 15.6 Å (~5.1 prolines). The inset shows E_Q as a function of the computed Leu′−Cnf interchomophore distance for the proline series. The data were fit to Förster (solid line) and Dexter (dashed line) distance dependences.

In order to demonstrate that thioamide quenching of Cnf can provide useful information on conformational changes, a proof of principle experiment was performed using a small, well-folded peptide prepared by chemical synthesis. HP35, a variant of the

villin headpiece subdomain, is known to adopt a compact helical structure in which the N- and C-termini are in close proximity (within 20 Å).[129] Since global unfolding of HP35 leads to a significant increase in the distance between the termini, we chose to incorporate Leu′ at the N-terminus and Cnf at the C-terminus for FRET experiments. Additionally, because Cnf is excited at 240 nm and fluoresces with a maximum at 295 nm, Trp_{23} was mutated to Phe to ensure that Trp fluorescence (arising from direct excitation or from Cnf/Trp FRET) would not convolute the spectral data. By obtaining temperature-dependent CD measurements, we were able to show that the T_m of the double-labeled HP35 variant (HP35-$L'_1F_{23}F^*_{35}$) was similar to that of the corresponding oxoamide control protein (HP35-$F_{23}F^*_{35}$), indicating that the thioamide replacement did not exert a measureable effect on overall stability. We were then able to demonstrate that the thermal denaturation of HP35 can be accurately monitored by observation of Cnf/thioamide FRET. Quenching efficiency (E_Q) was determined over a range of temperatures (5 to 75 ºC) by comparing the fluorescence intensity of the thioamide protein to the oxoamide control. Upon converting E_Q to distance using Förster theory, we determined that the separation between the termini of HP35 increased from 21 Å in the folded form to a distance beyond that measurable by our probe pair (32 Å) in the unfolded state.

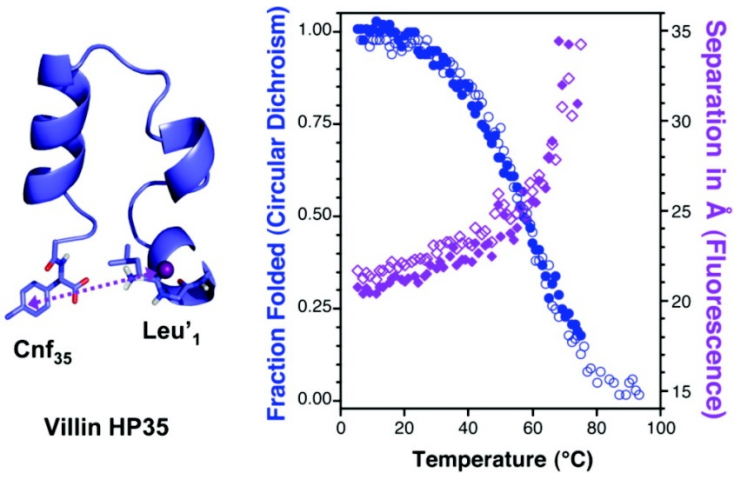

Figure 2.5 Villin HP35 unfolding monitored by Cnf/thioamide FRET.
Left: Villin HP35 structure modified with the Cnf nitrile on Phe_{35}. Right: Fraction folded as determined from temperature-dependent CD spectroscopy for HP35-Leu'_1Cnf_{35} (●) and HP35-Cnf_{35} (○) and temperature dependence of the Leu'_1/Cnf_{35} separation determined by comparison of E_Q computed from the HP35-Leu'_1Cnf_{35}/HP35-Cnf_{35} fluorescence ratio with the proline series distance dependence computed using either the Förster (◆) or Dexter (◇) equation. Figure adapted from Goldberg *et al.*[115]

Thioamide Quenching of Intrinsic Protein Fluorescence. Although electron-transfer-induced quenching of Trp and Tyr fluorescence by many functional groups is well documented, there are only a few reports of the effect of a thioamide on protein fluorescence. The first observation of fluorophore quenching by thioamides was reported by Wiczk *et al.* In this study, thioamide quenching of Trp in a short model peptide was examined in neat propylene glycol.[130] In non-polar solvent, thioamide absorbance is red-shifted and Trp emission is blue-shifted, resulting in significant spectral overlap and an enhanced likelihood of FRET. It was unclear from this study whether thioamide quenching of Trp would occur under aqueous conditions. Additionally, Rownicka *et al.* reported quenching of bovine serum albumin fluorescence by 2-*n*-propyl-6-thiouracil, which contains a thioamide.[131] Despite these initial reports, neither group carried out further systematic studies on thioamide quenching of intrinsic protein fluorescence.

The Cnf/thioamide pair is not generally compatible with Trp- and Tyr-containing proteins because FRET between the residues can complicate data interpretation. Since global replacement of these aromatic residues is undesirable, we chose to explore thioamide quenching of Trp and Tyr fluorescence in order to understand the effects of thioamide incorporation on the near-UV fluorescence of typical proteins.

§ 2.2 Results and Discussion

We began our investigation by considering a possible FRET mechanism for thioamide quenching of Trp and Tyr. According to Förster theory, the efficiency of energy transfer between a fluorophore and quencher depends on the spectral overlap of donor emission and acceptor absorbance.[9] Although the Tyr emission spectrum has moderate overlap with the absorbance of a thioamide, the Trp emission spectrum has almost no overlap (Figure 2.6). These observations are borne out in the theoretical calculations of R_0, the distance at which energy transfer is 50% efficient. Using the previously reported quantum yields for Tyr ($\Phi = 0.14$) and Trp ($\Phi = 0.13$), and assuming the transition moment dipoles to be randomly oriented during the course of energy transfer ($\kappa^2 = 2/3$), we calculated R_0 to be 13.4 Å for the Tyr/thioamide FRET pair, and 4.0 Å for the Trp/thioamide pair in aqueous solutions. For small values of R_0, such as those that we calculated for Trp, Förster theory may not provide an adequate prediction of quenching, since other short-range mechanisms will contribute significantly.

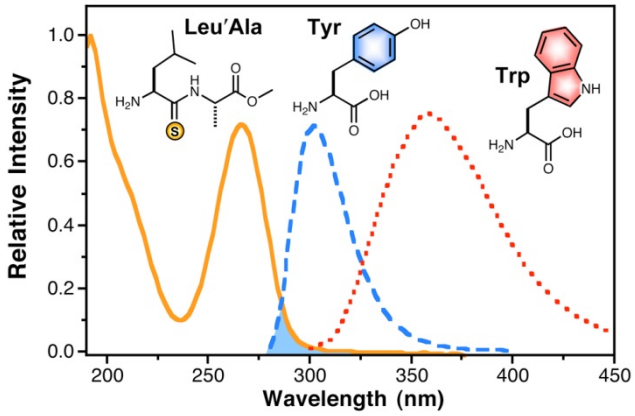

Figure 2.6 Thioleucylalanine ester (Leu'Ala), Tyrosine (Tyr) and Tryptophan (Trp) Spectra. Absorption spectrum of Leu'Ala (solid orange line) shown with relative absorption intensity normalized to extinction coefficient (12,400 $M^{-1}cm^{-1}$ at 266 nm). The fluorescence spectra of Tyr and Trp (dashed blue and dotted red lines, respectively) are arbitrarily normalized to emission maxima. The shaded area indicates the spectral overlap that contributes to FRET between a thioamide and Tyr.

To verify the distance dependence predicted by Förster theory, we synthesized groups of peptides that contained either Trp or Tyr at the C-terminus and thioleucine (Leu´) at the N-terminus, separated by an increasing number of proline residues. Fluorescence spectra of dilute samples of each peptide were taken in phosphate buffer at pH 7.00. In both cases, the fluorescence intensity showed a strong dependence on distance, at length scales much greater than that predicted by Förster theory (Figure 2.7). This is particularly noticeable for Trp, where a half-maximal change in fluorescence is observed at 17 Å, far exceeding the predicted R_0 of 4.0 Å. The distance dependence of both Tyr and Trp are shown fit to a sigmoidal expression, which provides a reasonable description of the distance dependence without assigning a mechanistic interpretation. Fits of the data to other mechanism-based equations, including Forster, Dexter, and electron transfer are provided below with accompanying discussion. The sigmoidal fit in Figure 2.7 serves as a working, empirical "ruler" based on quenching efficiency to be used while we investigate the mechanism of quenching further.

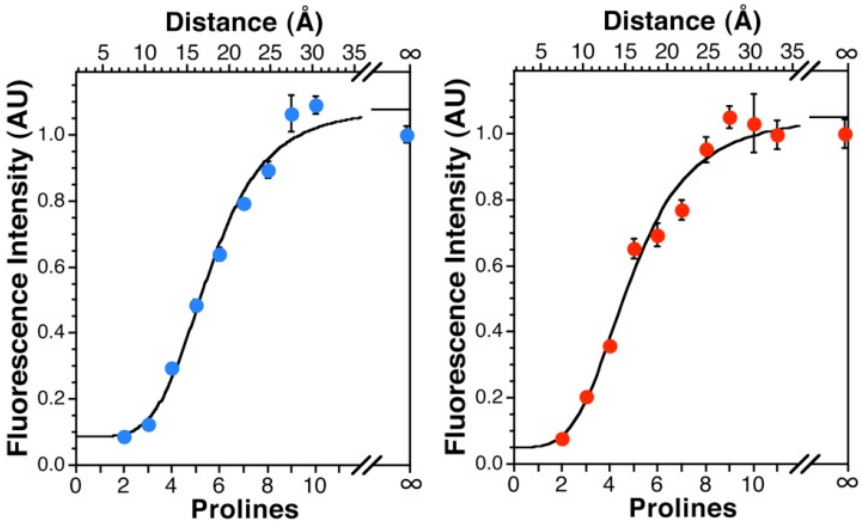

Figure 2.7 Fluorescence Intensity as a Function of Chomophore Separation.
Left: The fluorescence emission at 305 nm of Leu'- Pro_n-Tyr (n = 2-10) is shown. Right: Fluorescent emission at 355 nm of Leu'-Pro_n-Trp (n = 2-11). In both plots, the "∞" data point indicates the fluorescence of Leu-Pro_2-Tyr or Leu-Pro_2-Trp. Sigmoidal fits to the data shown (3 or more trials per peptide, bars represent standard error). Distances determined from molecular dynamics simulations.

To test the utility of our spectroscopic ruler and verify that quenching takes place in a molecular context other than the Pro series, we conducted protein-binding studies in which the thioamide was incorporated into peptides that bind to the protein calmodulin (CaM). We chose CaM because it expresses well in *E. coli* and it has been thoroughly studied from a structural and biophysical standpoint. CaM has been shown to bind a variety of helical peptides derived from its endogenous protein interaction partners.[132] These peptides bind with high affinity in a central groove formed by the folding of the N- and C-terminal CaM lobes. In particular, we studied the binding of a peptide derived from the intracellular domain of an olfactory cyclic nucleotide-gated ion channel (pOCNC).[133] This peptide binds with a low nM affinity, and its CaM-bound structure has been determined previously by NMR.[133] Therefore, we expected that 1:1 mixtures of 10 μM each CaM and pOCNC would be stably bound and that the NMR structures could be

used to determine interatomic distances. These measurements would then be compared to the corresponding distances derived from our FRET measurements.

Chicken CaM contains two tyrosine residues, Y_{100} and Y_{139}, and, in the presence of Ca^{2+}, binds helical peptides with diverse sequences. We prepared derivatives of pOCNC in which thiophenylalanine (Phe′) was incorporated at the N-terminus of the peptide (pOCNC-F_1′). As a control for any changes in fluorescence due to factors other than thioamide quenching, we prepared an oxoamide version of the peptide (pOCNC) with phenylalanine at the N-terminus. From NMR structural data, we estimated the distance between the thiocarbonyl carbon and the center of the phenol ring to be 19 Å for Y_{100} and 16 Å for Y_{139}.[133] Titrations of each peptide into CaM solutions are shown in Figure 2.8. A 28% decrease in fluorescence is seen at saturating concentrations of pOCNC-F_1′, whereas essentially no change in fluorescence is seen with the oxoamide control pOCNC. Since these peptides differ only in the O-to-S substitution in the F_1/R_2 amide bond, we assign the quenching to the Tyr/thioamide interaction. Both Trp and Tyr fluorescence are sensitive to their local environment with respect to solvent accessibility, pH, neighboring residues, and other chromophores.[134-136] CaM undergoes a substantial conformational rearrangement upon peptide binding, and we were concerned that these factors might alter Tyr fluorescence independent of thioamide quenching.[137,138] Therefore, the inclusion of an oxoamide control is essential to assign quenching specifically to the Tyr/thioamide interaction. Corrected pOCNC-F_1′ fluorescence values can be fit to a 1:1 binding model to obtain a K_D of less than 280 nM.

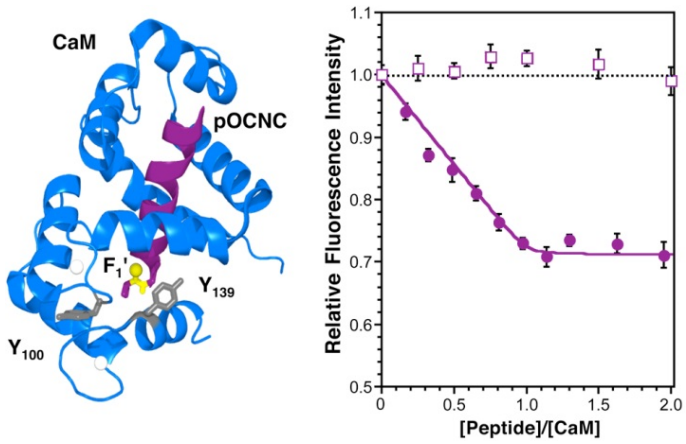

Figure 2.8 CaM Binding pOCNC.
Left: CaM structure taken from PDB 1SY9. Image created in PyMOL. Y_{100} and Y_{139} are rendered in grey; the thioamide bond is highlighted in yellow. Right: Titration of pOCNC (filled circles) and pOCNC-F_1' (open squares) into a solution of 10 μM CaM in 15 mM HEPES buffer, 140 mM KCl, and 6 mM $CaCl_2$, pH 6.70, monitored by fluorescence spectroscopy. Fluorescence data are normalized to CaM in the absence of peptide and corrected for peptide fluorescence.

Prediction of the efficiency of quenching (E_Q) from Förster theory is complicated by the fact that the two donor Tyr have different extinction coefficients (ε) and quantum yields (Φ), both of which contribute to FRET (i.e. quenching) efficiency.[137,139-141] Using reported ε and F values for each Tyr, we calculated E_Q values from Förster theory and an empirical calibration based on Figure 2.7. Since Phe is not noticeably fluorescent at the excitation wavelength of Tyr, we generated $Y_{100}F$ and $Y_{139}F$ mutants to verify the effect of thioamide quenching on each Tyr residue independently. Table 2.1 summarizes the observed quenching of each mutant in a stoichiometric ratio with pOCNC-F_1′, as well as theoretical predictions from Förster theory (E_Q Förster) and our empirical spectroscopic ruler (E_Q Calculated) using distances from the NMR structure. There is reasonable agreement between our empirical E_Q and the observed E_Q for the single Tyr mutants, but the WT values deviate, perhaps because we are not weighting the fluorophore contributions or orientation effects (κ^2) properly.

Table 2.1 Quenching of CaM Mutant Fluorescence by Thioamide Peptides.

CaM Mutant (Donor, DQ Distance)[a]	E_Q (%) Experimental[b]	E_Q (%) Calculated[b,c]	E_Q (%) Förster[b,c]
WT	28 ± 1	42	10
$Y_{100}F$ (Y_{139}, 16 Å)	50 ± 2	56	14
$Y_{100}W$ (W_{100}, 19 Å)	27 ± 3	30	0
$Y_{139}F$ (Y_{100}, 19 Å)	22 ± 2	34	8
$Y_{139}W$ (W_{139}, 17 Å)	58 ± 2	40	0

[a] Donor/quencher (DQ) distance measured in PyMol from aromatic ring of Tyr or Trp to thiocarbonyl carbon. [b]Calculations described in the Supplementary Information. [c] We estimate the error to be ± 15% for Tyr predictions and ± 20% for Trp predictions.

To further examine quenching of Trp, we mutated either Y_{100} or Y_{139} to Trp. Stimulation at 295 nm allowed us to selectively excite the Trp residue. The observed E_Q of the Trp mutants at stoichiometric concentrations of pOCNC-F_1′ are reported in Table 1 along with predicted quenching efficiencies. Several observations are worth noting from these data. First, given the predicted and observed distance dependencies, a FRET mechanism is clearly not responsible for Trp quenching, and is at best partially responsible for Tyr quenching. Second, the exponential distance-dependence one might expect for though-bond electron transfer is seen both in the Pro series and the CaM-binding experiments, which shows that the rigidity of the Pro σ-bond framework is not necessary for efficient transfer over 15-20 Å distances. Third, the observed quenching efficiency in the CaM binding experiments is in moderate agreement with the quenching level predicted from our polyproline rulers, indicating that they may be useful as empirical standards. Further investigation in other systems (below), allowed us to gain a greater understanding of local environment and orientation effects.

We demonstrated that thioamides quench Trp and Tyr fluorescence in a distance-dependent fashion that can be used to monitor biological interactions, such as

macromolecular binding events. It is well established that Trp and Tyr are also quenched by many other functional groups. However, since we can compare our synthetic thioamide proteins to all-oxoamide equivalents, we should be able to separate thioamide-quenching from quenching by protein sidechains or backbone in interpreting our data. The Trp data presented here as well as other data collected in our laboratory show that thioamides are capable of quenching red-shifted fluorophores with no spectral overlap through an electron transfer process. Given the substantial difference in the oxidation potentials between oxoamides (3.25 eV) and thioamides (1.21 eV), it is not surprising that thioamide-specific quenching can occur.[142] We are investigating this phenomenon further in the context of Tyr and Trp as well as other fluorophores. The methods for the semi-synthesis of large thioamide-containing proteins described below may be used to extend our findings to mapping interactions among two full-sized proteins.

§ 2.3 Native Chemical Ligation of Thioamide-Containing Peptides: Development and Application to the Synthesis of Labeled α-Synuclein for Misfolding Studies

In order to expand the utility of the thioamide as a spectroscopic probe, our laboratory has developed semi-synthesis techniques for its installation into full-length proteins.[116] Backbone thioamides cannot currently be installed into a protein by means of ribosomal expression. Therefore, we incorporate thioamides into full-length proteins using native chemical ligation (NCL), a fragment condensation reaction that typically takes place between peptides bearing a C-terminal thioester and an N-terminal cysteine.[143] Ligation proceeds when the thiol of the N-terminal Cys attacks the C-terminal thioester to form an S-acyl intermediate. This intermediate undergoes an irreversible S to N acyl shift to generate a native peptide bond containing Cys at the ligation site. NCL reactions have been shown to be tolerant of protein functional groups

as well as a variety of non-native functional groups. Thus, we expected thioamides to be compatible with NCL conditions, although we had some concerns regarding Edman-type side reactions or Cys desulfurization through attack of the nucleophilic thioamide (Scheme 2.1). Another potential concern was desulfurization of the thioamide via transient water attack at the thiocarbonyl carbon followed by expulsion of H_2S from the tetrahedral intermediate. However, initial investigations of NCL reactions between short peptides demonstrated that thioamides can be placed at nearly any position in the thioester or Cys peptide fragments.

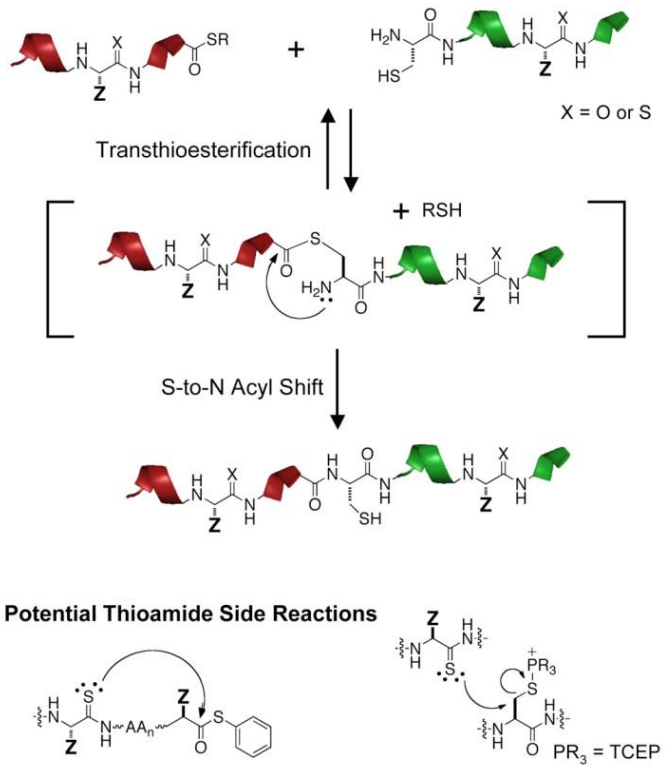

Figure 2.9 Native Chemical Ligation Mechanism and Potential Thioamide Side Reactions.

Misfolding of the abundant neuronal protein α-Synuclein (αS) has been implicated in the pathogenesis of several debilitating neurodegenerative disorders including Parkinson's disease (PD).[53] However, the molecular details underlying the

conversion of αS into neurotoxic species remain poorly understood. Accordingly, a more complete picture of the structural changes of αS upon misfolding would greatly facilitate our understanding of the spread of PD.[58] Many fluorescence-based studies have attempted to determine structures for aggregation intermediates, but they often employ fluorophores that can perturb the structure, particularly when substituted at sites that do not natively contain an aromatic amino acid.[144] Our thioamide probe, which has been shown to quench natural and unnatural amino acids, should offer much greater freedom, particularly in the placement of the acceptor (i.e., the thioamide) in a donor/acceptor pair. Here, we demonstrate that we can use semi-synthesis to generate thioamide-labeled αS and that thioamide fluorescence quenching can be used to track conformational changes during the course of its aggregation.

§ 2.4 Results and Discussion

Note: Semi-Synthesis of αS was performed by Solongo Batjargal and Yanxin Wang.[116]

In order to make labeled αS through semi-synthesis, we generated constructs for producing a version of αS with an N-terminal His tag which could be removed by proteolysis to reveal an N-terminal Cys for ligation. We chose to use commercially available Factor Xa protease, which selectively cleaves an IEGR sequence and had been used previously in expressed protein ligation.[145] Since αS has no native Cys, we introduced one by altering the expression construct to remove the first eight residues and leave an N-terminal Cys ($\alpha S_{9–140}C_9$) after proteolysis.

Solution-phase thioesterification of protected peptides can lead to epimerization of the α-carbon of the C-terminal residue. We therefore examined an on-resin thioesterification method for synthesizing the thioamide-containing fragment of αS for

ligations. Our method uses a C^b-Pro-G_o linker, where C^b is a t-Bu-protected Cys, and G_o is glycolic acid. Following acidic cleavage from the resin, the C^b t-Bu group is removed by treatment with TCEP, initiating a cascade of reactions resulting in a peptide thioester attached to the diketopiperazine (Dkp) product of C^b-Pro-G_o rearrangement (Scheme 2.2). The Dkp thioester can then either react directly with an N-terminal Cys or undergo transesterification with a thiol additive.

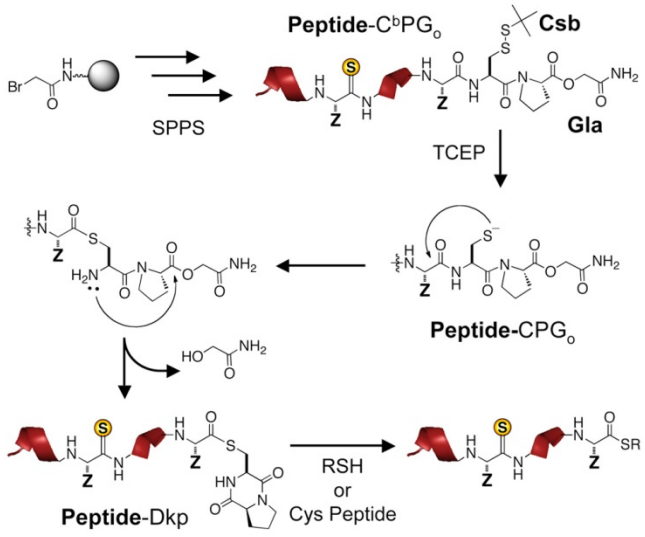

Figure 2.10 On-Resin Thioester Generation by N-to-S Rearrangement.

The scheme for the semi-synthesis of Val_3-labeled αS is shown in Figure 2.10. The C-terminal fragment, H_{Tag}-α$S_{9–140}C_9$, was expressed in *E. coli*, purified by Ni-NTA chromatography, and proteolyzed with Factor Xa to give α$S_{9–140}C_9$. The N-terminal fragment was synthesized on solid phase using the C^bPG_o linker, with a thioamide inserted at Val_3 during SPPS (α$S_{1–8}V'_3$-C^bPG_o). Activation of the peptide by TCEP yielded α$S_{1–8}V'_3$ Dkp thioester, which reacted with α$S_{9–140}C_9$ *in situ* to give full-length αSV'_3C_9. Although ligations reached only 50% completion, ligated αSV'_3C_9 could be purified by HPLC. The integrity of the thioamide bond after ligation was assessed by

MALDI MS of the full-length protein and trypsinized products, as well as by UV absorption spectroscopy.

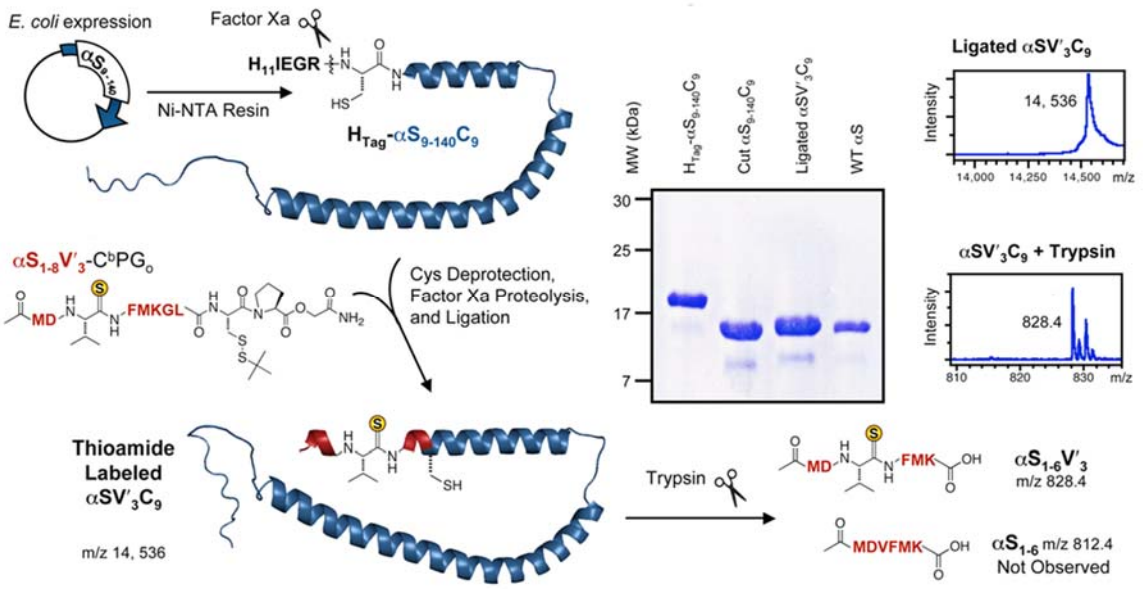

Figure 2.11 Synthesis of Labeled αS by Ligation of a Thioamide Peptide to an Expressed Protein Fragment.
Left: C-Terminal fragment of $\alpha S_{9\text{-}140}$ with an N-terminal His tag, an S_9C mutation, and IEGR cleavage sequence expressed and purified on Ni-NTA resin. After Factor Xa cleavage to generate an N-terminal Cys, the protein fragment is ligated to synthetic $\alpha S_{1\text{-}8}$ with a thioamide at position 3 ($\alpha S_{1\text{-}8}V'_3$), generated *in situ* from precursor $\alpha S_{1\text{-}8}V'_3\text{-}C^bPG_o$. Top Right: PAGE gel analysis of ligation. Far Right: MALDI MS of full length aSV'_3C_9 (calcd m/z 14, 536) or trypsinized $\alpha SV'_3C_9$, giving $aS_{1\text{-}6}V'_3$ (calcd m/z 828.4). There is no peak corresponding to desulfurized oxopeptide $\alpha S_{1\text{-}6}$ (calcd m/z 812.4).

In order to demonstrate the value of incorporating thioamide probes into a protein such as αS, we carried out an aggregation experiment in which intramolecular conformational changes were monitored in an αS construct. Here, we employ an αS construct with a single Trp donor fluorophore (W_{94}) and a single thioamide quencher (V'_3). The native αS sequence has no Trp residues, so selective excitation of an introduced Trp can be achieved with 295 nm light. Comparison to an oxoamide control (i.e., αSW_{94}) allows us to isolate thioamide quenching effects from other environmental effects on Trp fluorescence. Since through-space quenching does not occur over distances longer than 30 Å, observation of Trp fluorescence quenching normalized to oxoamide

controls shows that the Trp must be within 30 Å of the thioamide residue. However, in an oligomer or fibril, quenching can arise through either inter- or intramolecular interactions. In order to isolate intramolecular misfolding events, we carry out aggregation experiments with our $\alpha SV'_3C_9W_{94}$ construct present in a 1:30 ratio with WT αS. A ratio of less than one Trp/thioamide-labeled αS per 26 WT αS should ensure that, on average, no two labeled αS are directly next to each other in aggregates. This derives from a simple statistical packing model in which each monomer unit would be at the center of a cube with 26 partners along the centers of faces and vertices, or at corners.

The construct used in aggregation studies was synthesized by ligation of the same synthetic fragment ($\alpha S_{1–8}V'_3$) with an expressed C-terminal fragment ($\alpha S_{9–140}C_9W_{94}$). For an aggregation experiment, the presence of non-native Cys in the sequence could be problematic as the formation of disulfide-bonded dimers would alter the aggregation mechanism. Therefore, we incubated αS with β-mercaptoethanol (BME) to prevent disulfide bond formation. No significant increase in aggregation rate was observed in the Cys-containing mutants.

Aggregation experiments were carried out by shaking the 1:30 $\alpha SV'_3C_9W_{94}/\alpha S$ mixtures at 37 °C in phosphate-buffered saline with BME. Although overall Trp fluorescence increased over 4 days during protein aggregation, the relative fluorescence of the thioamide (F_{Thio}/F_{Oxo}) decreased. (See section 2.6, Materials and Methods for additional primary fluorescence data.) This indicates that positions 3 and 94 approach each other to within 30 Å in the oligomeric and fibrillar states, but that they are more separated in the monomeric state. Since we used a 1:30 $\alpha SV'_3C_9W_{94}/\alpha S$ ratio, we interpret the quenching as arising primarily from intramolecular interactions within the

labeled molecule. We normalize our data to an equivalent oxoamide protein (1:30 $\alpha SW_{94}/\alpha S$) in order to account for other local effects on fluorescence, which are necessarily complex in an aggregate. We monitored overall fibrillization by ThT fluorescence, as well as the independent metrics of Congo Red (CR) staining and sedimentation PAGE gel analysis (see below).

Comparing the ThT data to the Trp fluorescence data clearly indicates that we can observe quenching in oligomers, which occur at early time points (<24 h), before fibrillization and ThT fluorescence (Figure 2.12). It is unlikely that multiple labeled αS molecules are present in oligomers of less than 30 monomer units. Therefore, quenching of Trp fluorescence provides evidence of close intramolecular approach of residues 3 and 94 when bound in oligomers, consistent with some previous studies showing folding of the N- and C-termini before fibrillization.[101] Our experiments show that thioamide fluorescence quenching should be useful in monitoring intermediates in the aggregation process, which are silent in the conventional ThT protocol. Conducting many such experiments in conjunction with positional scanning of the thioamide probe should provide mechanistic insight into the aggregation process and the mode of toxicity of these metastable intermediates.

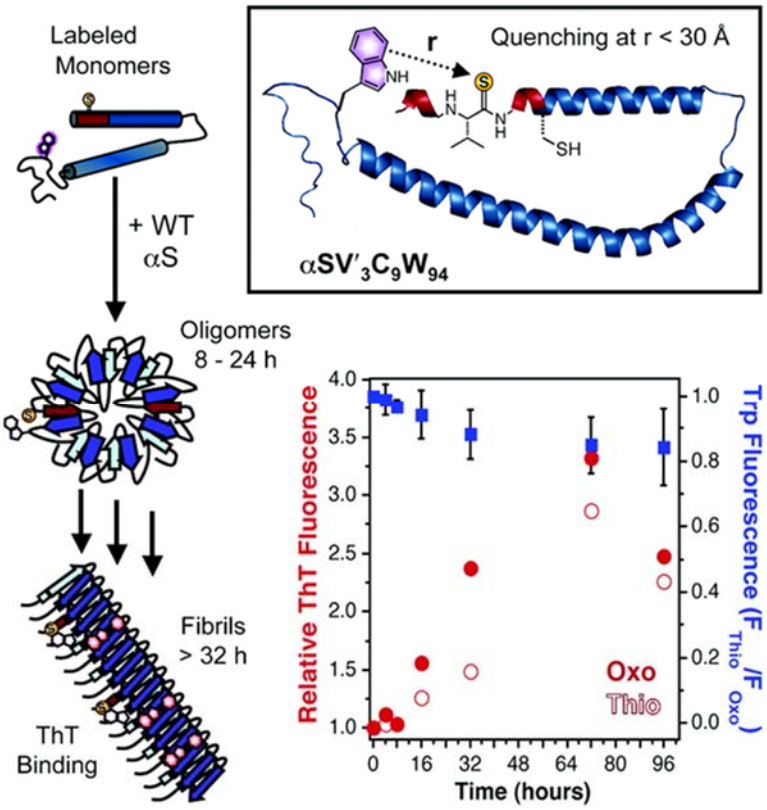

Figure 2.12 Monitoring Intramolecular Misfolding During αS Aggregation Using Thioamide Quenching.
Left: Monomeric $\alpha SV'_3C_9W_{94}$ mixed in a 1:30 ratio with WT αS. Trp fluorescence quenching is observed in oligomers and fibrils. ThT fluorescence is only observed for fibril-bound ThT. Top Right: Schematic representation of $\alpha SV'_3C_9W_{94}$ construct showing Trp, thiovaline, and Cys. Bottom Right: Normalized Trp (squares) and ThT (circles) fluorescence data for four aggregation experiments with $\alpha SV'_3C_9W_{94}$ (Thio) and control experiments with αSW_{94} (Oxo).

§ 2.5 Conclusion

We have recently shown that thioamide quenching of Trp or Tyr could be used to monitor interactions for native proteins.[146] These experiments would be limited to thiopeptides synthesized on solid phase without the ability to ligate thioamide-containing peptide fragments. Thus, the methods developed here will be used to synthesize thioamide-labeled proteins for use in the study of protein folding and protein–protein interactions. Misfolded or aggregating proteins are particularly attractive as they are often refractory to higher resolution structural techniques such as crystallography or NMR. We envision generating a large number of labeled αS constructs and using thioamide

fluorescence quenching to monitor conformational changes during aggregation and misfolding. Our proof-of-principle experiment indicates that these studies are feasible and should reveal information that cannot be accessed with common methods such as ThT binding. Of course, our methods are potentially applicable to a great number of protein dynamics questions, and we will continue to explore thioamide compatibility with NCL reactions for the synthesis and study of these proteins.

§ 2.6 Materials and Methods

General Information. Boc-L-thionoleucine-1-(6-nitro)benzotriazolide and Boc-L-thionophenylalanine-1-(6-nitro)benzotriazolide were purchased from Bachem (Torrance, CA). 2-chlorotrityl resin, Fmoc-Ala-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Leu-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Gly-OH, Fmoc-Val-OH, Fmoc-Phe-OH, Fmoc-Ile(OtBu)-OH, Fmoc-Tyr(OtBu) and Fmoc-Trp(Boc)-OH were purchased from Novabiochem (San Diego, CA). Piperidine and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from American Bioanalytical (Natick, MA). Sigmacote®, *N*,*N*-diisopropylethylamine (DIPEA), hen egg-white lysozyme, and phenyl-sepharose CL-4B resin were purchased from Sigma-Aldrich (St. Louis, MO). *E. coli* BL21(DE3) cells and a QuickChange® site-directed mutagenesis kit were purchased from Stratagene (La Jolla, CA). DNA oligomers were purchased from Integrated DNA Technologies, Inc (Coralville, IA). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA). Milli-Q filtered (18 MΩ) water was used for all aqueous solutions (Millipore; Billerica, MA). All peptide synthesis reaction vessels (RVs) were treated with Sigmacote® prior to use. DNA sequencing was performed at the University of Pennsylvania DNA sequencing facility. Matrix-assisted

laser desorption ionization (MALDI) mass spectra were collected with a Bruker Ultraflex III MALDI-TOF-TOF mass spectrometer (Billerica, MA). UV absorbance spectra were obtained with a Hewlett-Packard 8452A diode array spectrophotometer (currently Agilent Technologies; Santa Clara, CA). Fluorescence spectra were collected with a Varian Cary Eclipse fluorescence spectrophotometer fitted with a Peltier multicell holder (currently Agilent Technologies; Santa Clara, CA). Circular dichoism experiments were conducted with an Aviv 410 CD spectrometer (Aviv Biomedical; Lakewood, NJ).

Peptide Synthesis was performed by Jacob Goldberg and Alyssa Klein.

Peptide Synthesis and Purification. Leu-Pro_2-Trp (LP_2W) and Leu-Pro_2-Tyr (LP_2Y) were each synthesized on a 12.5 μmol scale on 2-chlorotrityl resin. For each synthesis, 2-chlorotrityl chloride resin (100-200 mesh; 0.6 mmol substitution/g; 12.5 μmol) was added to a dry glass RV. The resin was swollen by two successive 15 min incubations with 5 mL dimethylformamide (DMF) and magnetic stirring. After swelling, DMF was removed with vacuum suction and Fmoc-Trp(Boc)-OH or Fmoc-Tyr(OtBu)-OH was coupled to the resin. The amino acid in DMF (5 equiv; 42 mM, 1.5 mL) and DIPEA (10 equiv; 22 μL) were added to the RV and the mixture was allowed to react for 30 min with magnetic stirring. Spent solution was removed with vacuum suction and the resin beads were washed thoroughly with DMF. Excess DMF was removed with vacuum suction and the resin beads were deprotected by treatment with 20% piperidine in DMF (5 mL) for 20 min with magnetic stirring. The deprotection solution was drained from the RV and the beads were rinsed extensively with DMF. Subsequent amino acid couplings and deprotections proceeded as described above, with the exception that the remaining Fmoc-protected amino acids were activated with HBTU (5 equiv) prior to addition to each

reaction. The N-terminal Fmoc group was removed before the peptide was cleaved from the resin. After the beads were washed extensively with DMF and dried with CH_2Cl_2, peptides were cleaved by successive 60 min and 30 min incubations on a rotisserie with 2.5 mL of a fresh cleavage cocktail of trifluoroacetic acid (TFA), water, and triisopropylsilane (TIPS) (10:9:1 v/v). After each treatment, the resulting solution was expelled from the RV with nitrogen and reduced to a volume of less than 1 mL by rotary evaporation. The combined residue was diluted with 6 mL of CH_3CN/H_2O (2:1 v/v).

The peptides were purified by reverse-phase HPLC on a Vydac 218TP C18 semi-prep column (Grace/Vydac; Deerfield, IL) using a linear solvent gradient that ranged from 98% to 60% aqueous phase over 19 min, then to 0% aqueous phase over 5 min, then returning to 98% aqueous phase during a 10 min wash out period. Trp peptides eluted at approximately 23 min with this method and Tyr peptides eluted at approximately 19 min. MALDI-MS was used to confirm identities (Table 2.2). Purified peptides were dried in a vacuum centrifuge.

Polyproline peptides Leu'-Pro_n-Trp (L'P_nW), n = 2-11, and Leu'-Pro_n-Tyr (L'P_nY), n = 2-10 were each synthesized on a 10 μmol scale using a procedure analogous to that described above. These peptides were synthesized from a common pot of resin (100 μmol). After coupling the appropriate number of proline residues, a 10 μmol portion of resin was removed from the RV and transferred to a separate, clean RV for thioleucine coupling. Since the thioleucine was introduced through the pre-activated Boc-L-thionoleucine-1-(6-nitro)benzotriazolide, HBTU was not added for these couplings. As the synthesis progressed, reagents were scaled accordingly. Peptide cleavage, purification, and characterization followed that described above. The CaM binding

peptide pOCNC has the sequence FRRIARLVGVLREFAFR; in the thioamide derivative pOCNC-F_1', the N-terminal phenylalanine is replaced with thiophenylalanine. The general sequence of these peptides is derived from the bOCNCp peptide fragment described by Contessa *et al*. (Sequence: GGFRRIARLVGVLREWAYR). Each peptide was synthesized on a 100 μmol scale using the same general procedures as those described above. Briefly, a peptide consisting of residues R_2-R_{17} was synthesized on a 200 μmol scale on 2-chlorotrityl resin with 5 equiv amino acid, 10 equiv DIPEA, and 5 equiv HBTU when needed. Before removal of the Fmoc group from R2, the resin beads were divided into 2 equal portions and transferred to clean RVs for the final coupling of either Fmoc-L-Phe-OH or Boc-L-thionophenylalanine-1-(6-nitro)benzotriazolide. After removal of the N-terminal Fmoc group from pOCNC with 20% piperidine, the resin beads were incubated on a rotisserie twice for 60 min with 10 mL portions of fresh TFA/H_2O/TIPS (38:1:1 v/v). The thioamide-containing peptide was cleaved from the resin and deprotected with two successive 60 min incubations of the resin beads with 10 mL portions of fresh TFA/H_2O/TIPS (18:1:1 v/v). After each incubation period, the cleavage cocktail for each peptide was expelled from the RV with nitrogen and dried by rotary evaporation. The residue was diluted to 10 mL with CH_3CN/H_2O (3:2 v/v) before purification. Crude peptides were purified by reverse-phase HPLC with linear solvent gradients that ramped from 98% to 70% aqueous phase over 6 min, then to 56% aqueous phase over 14 min, then to 0% aqueous phase over 5 min, followed by a 10 min wash out period, during which the solvent system returned to 98% aqueous phase. Using this method, the peptides eluted at approximately 19.5 min (Figure 2.13). Peptides were analyzed by MALDI-MS (Table 2.2).

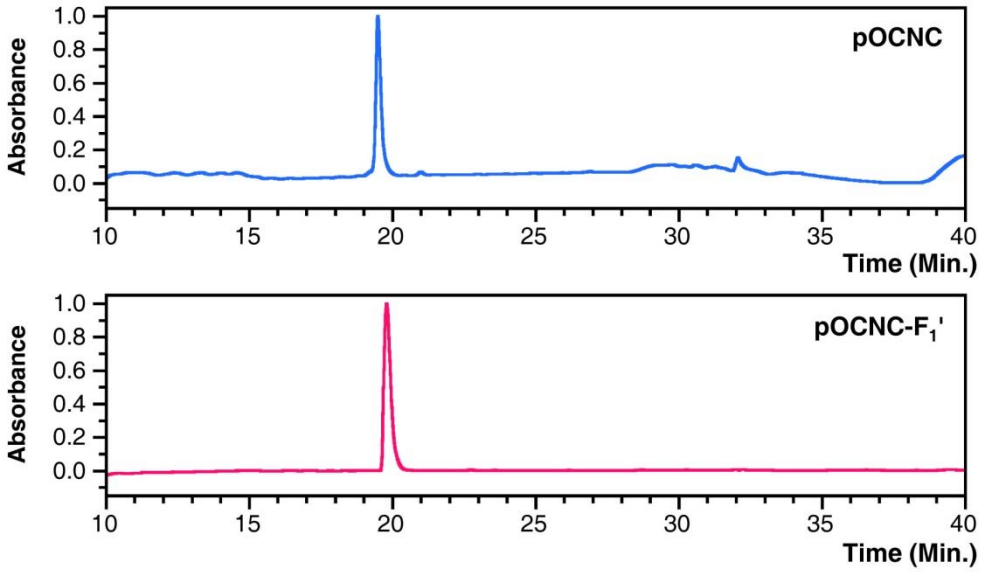

Figure 2.13 Analytical HPLC Chromatograms of Purified Peptides.
Top: pOCNC monitored at 215 nm. Bottom: pOCNC-F_1' monitored at 277 nm. Absorbance is normalized. Solvent gradients given in text.

Table 2.2 Calculated and Observed Peptide Masses.

Peptide	Calculated m/z $[M+H]^+$	Observed m/z $[M+H]^+$	Calculated m/z $[M+Na]^+$	Observed m/z $[M+Na]^+$
LP_2Y	489.27	489.15	511.25	511.12
$L'P_2Y$	505.25	505.14	527.23	527.11
$L'P_3Y$	602.30	602.12	624.28	624.08
$L'P_4Y$	699.35	699.16	721.36	721.14
$L'P_5Y$	796.41	796.28	818.39	818.20
$L'P_6Y$	893.46	893.29	915.44	915.27
$L'P_7Y$	990.51	990.21	1012.49	1012.18
$L'P_8Y$	1087.57	1087.24	1109.55	1109.21
$L'P_9Y$	1184.62	1184.60	1206.60	1206.57
$L'P_{10}Y$	1281.67	1281.66	1303.65	1303.66
LP_2W	512.29	512.02	534.27	534.01
$L'P_2W$	528.26	528.28	550.25	550.26
$L'P_3W$	625.32	625.10	647.30	647.07
$L'P_4W$	722.37	722.14	744.35	744.12
$L'P_5W$	819.42	819.33	841.40	841.30
$L'P_6W$	916.48	916.41	938.46	938.39
$L'P_7W$	1013.53	1013.28	1035.51	1035.27
$L'P_8W$	1110.58	1110.27	1132.56	1132.25
$L'P_9W$	1207.63	1207.58	1229.62	1229.51
$L'P_{10}W$	1304.69	1304.57	1326.67	1326.57
$L'P_{11}W$	1401.74	1401.69	1423.72	1423.67
pOCNC	2106.26	2106.78	2128.24	2128.75
pOCNC-F_1'	2122.23	2122.31	2144.22	2144.28

Förster Distance Calculation. The Förster distance is given in Å by equation 2.1,

$$R_0^6 = \frac{9000(\ln 10)\kappa^2 Q_D J}{128\pi^5 n^4 N_A} \qquad \text{(Eq. 2.1)}$$

where κ^2 is a geometrical factor that relates the orientation of the donor and acceptor transition moments, Q_D is the quantum yield of the donor, n is the index of refraction of the solvent, N_A is Avogadro's number, and J is the spectral overlap integral defined in units of $M^{-1}\cdot cm^{-1}\cdot nm^4$. Combining constants and rearranging gives R_0 as

$$R_0 = 0.211\{Q_D \kappa^2 n^{-4} J\}^{1/6} \qquad \text{(Eq. 2.2)}$$

J is formally defined as

$$J = \int_0^\infty f_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \qquad \text{(Eq. 2.3)}$$

where $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor at each wavelength λ and $f_D(\lambda)$ is the normalized donor emission spectrum given by

$$f_D(\lambda) = \frac{F_{D\lambda}(\lambda)}{\int_0^\infty F_{D\lambda}(\lambda) d\lambda} \qquad \text{(Eq. 2.4)}$$

where $F_{D\lambda}(\lambda)$ is the fluorescence of the donor at each wavelength λ. Fluorescence spectra of LP_2W and LP_2Y in water were integrated with KaleidaGraph from 300 to 450 nm for LP_2W and 280 to 400 nm for LP_2Y to calculate $f_D(\lambda)$. UV-Vis spectra of thioleucylalanine ester in water were used to determine $\varepsilon_A(\lambda)$. The literature value of ε_{273} = 12,400 $M^{-1}\cdot cm^{-1}$ was used to prepare solutions of known concentration. J was calculated to be 1.63×10^9 $M^{-1}\cdot cm^{-1}\cdot nm^4$ for the Trp/thioamide pair and 2.21×10^{12} $M^{-1}\cdot cm^{-1}\cdot nm^4$. Substituting these results into equation 2.2, as well as the quantum yields (0.13 for Trp and 0.14 for Tyr), 1.33 for the index of refraction of water, and 2/3 for κ^2

gives Förster distances of 4.0 Å for the Trp-thioamide pair and 13.4 Å for the Tyr-thioamide pair.

Polyproline ruler experiments were performed by Jacob Goldberg.

Fluorescence Spectroscopy. Dry proline series peptides were brought up in a minimal volume of pH 7.0 phosphate buffer (150 mM NaCl, 10 mM Na_2HPO_4, pH adjusted with HCl). Tryptophan-containing peptides were diluted to concentrations of approximately 10 μM, as determined by absorbance at 280 nm (ε_{280} = 13,270 M^{-1}•cm^{-1} for thioamide-containing peptides; ε_{280} = 5,600 M^{-1}•cm^{-1} for all others). Tyrosine-containing peptides were diluted to concentrations of approximately 20 μM, as determined by absorbance at 274 nm (ε_{274} = 14,420 M^{-1}•cm^{-1} for thioamide-containing peptides; ε_{274} = 1,420 M^{-1}•cm^{-1} for all others). Corrected fluorescence spectra were collected at 25 °C in triplicate for each peptide using quartz fluorometer cells with path lengths of 1.00 cm. For all Trp experiments, the excitation wavelength was 295 nm and emission data was collected from 300 - 500 nm as the average of three scans. For all Tyr experiments, the excitation wavelength was 275 nm and emission data was collected from 280 – 400 nm. For all experiments, the excitation and emission slit widths were 5 nm, the scan rate was 120 nm/min, the averaging time 0.5 s, and the data interval 1.0 nm. Examples of fluorescence spectra and associated UV spectra are shown in Figure 2.14.

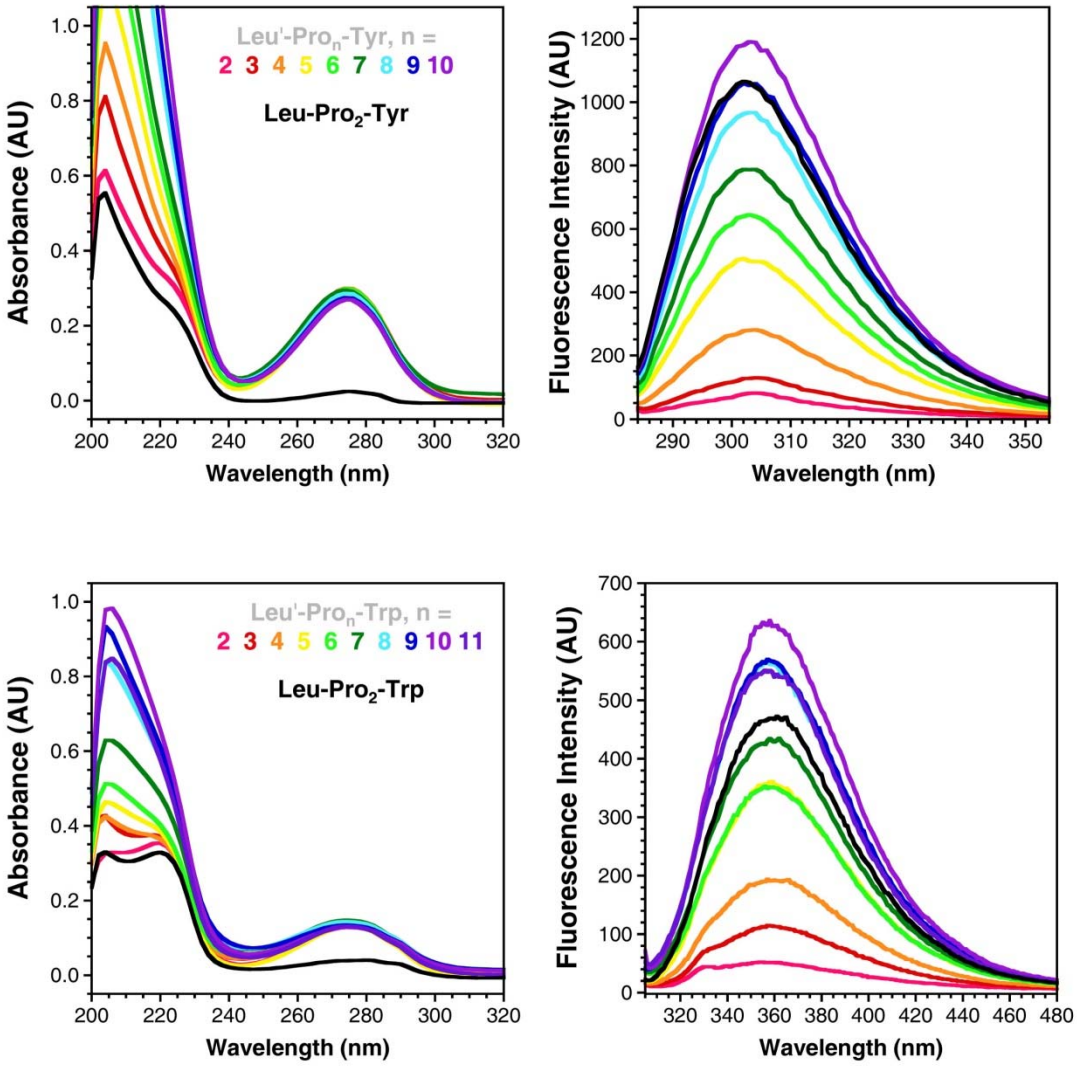

Figure 2.14 UV Absorption and Fluorescence Emission Spectra for Pro Series Peptides. Top Left: Tyr UV spectra collected between 200 and 320 nm and background corrected against average absorption from 450 to 500 nm. Top Right: Tyr fluorescence emission spectra collected between 285 and 400 nm with excitation at 275 nm. Bottom Left: Trp UV spectra collected between 200 and 320 nm and background corrected against average absorption from 550 to 650 nm. Bottom Right: Trp fluorescence emission spectra collected between 305 and 480 nm with excitation at 295 nm. Emission spectra colored according to corresponding absorption spectrum.

Proline Series Fluorescence. Fluorescence intensities at the wavelength of maximum emission (λ_{max} = 305 nm for Tyr and λ_{max} = 355 nm for Trp) were averaged from three separate trials for each proline peptide, except for L'P_{10}W, which represents six trials. These values were then normalized to the fluorescence intensity of Leu-Pro_2-Trp or Leu-Pro_2-Tyr, as appropriate, and plotted against the number of prolines separating the

chromophores and that distance in angstroms (Figure 2.15). The resulting data sets were fit to a sigmoidal function $F(x)$:

$$F(x) = c_1 + \frac{c_2 - c_1}{1 + \left(\frac{x}{c_3}\right)^{c_4}} \qquad \text{(Eq. 2.5)}$$

Here, c_1 through c_4 are constants and the variable x is defined with respect to the choice of the abscissa. In Figure 2.15, the data were fit to equation 2.5 with the independent variable in terms of n, the number of prolines. Using this fit, the distance in prolines at which 50% quenching is observed is 5.17 ± 0.59 prolines for Tyr ($R^2 = 0.98$) and 4.58 ± 0.79 prolines for Trp ($R^2 = 0.98$). Figure 2.16 shows the fits to the same equation with the independent variable explicitly cast in angstroms. The distance in angstroms at which 50% quenching is observed is 16.2 ± 2.0 Å for Tyr ($R^2 = 0.98$) and 15.1 ± 3.0 Å for Trp ($R^2 = 0.97$).

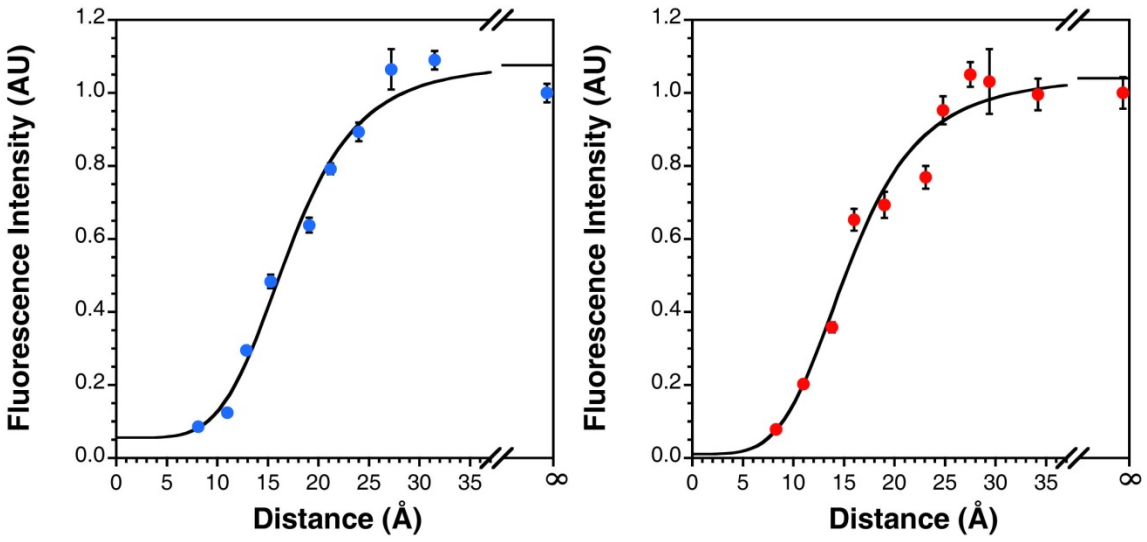

Figure 2.15 Fluorescence Intensity as a Function of Interchomophore Distance.
Explicit fits of Tyr (left) and Trp (right) fluorescence data were fit to equation 2.5 with the abscissa in angstroms. Distance calculations described below.

Quenching efficiency (E_Q) as a function of distance in angstroms was calculated using the formula:

$$E_Q = 1 - (F/F_0) \qquad \text{(Eq. 2.6)}$$

where F is the fluorescence at each data point and F_0 is taken to be the upper asymptotic limit, c_1. The resulting data sets were plotted against distance in angstroms and fit to equation 2.5. Both curves fit the data reasonably well, with goodness-of-fit values of $R^2 = 0.98$ for Tyr and $R^2 = 0.97$ for Trp. The distances corresponding to 50% quenching are 16.7 ± 2.2 Å for Tyr and 15.5 ± 3.0 Å, which agree with those values found fitting the raw fluorescence data as described above.

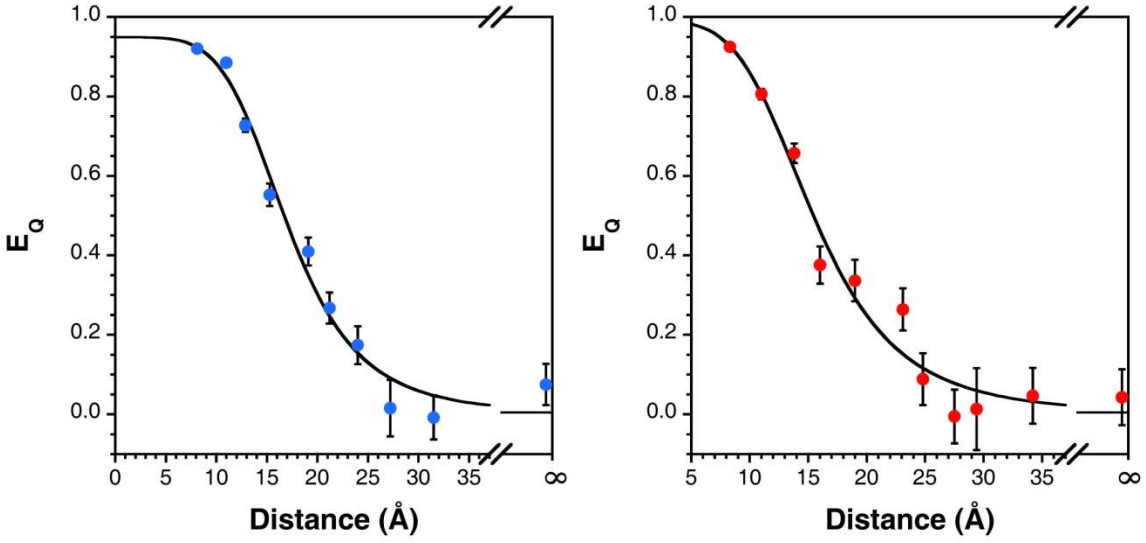

Figure 2.16 Quenching Efficiency as a Function of Interchromophore Distance.

Normalized fluorescence data were also fit to FRET, Dexter, and electron transfer models to evaluate their suitability for describing the observed quenching. For the FRET model, the data were fit to equation 2.7 using KaleidaGraph.

$$F = F_0\left\{1 - \frac{1}{1+\left(\frac{R}{R_0}\right)^6}\right\} \qquad \text{(Eq. 2.7)}$$

Here, F_0, the fluorescence of the donor in the absence of the acceptor, and R_0, the Förster distance, are adjustable parameters, and R is the separation between the chromophores in either the number of prolines or the distance in angstroms as calculated from molecular dynamics simulations (described below). Plots of these fits are shown in Figure 2.17 as solid traces. Although the curves generally follow the observed pattern of quenching, the values for R_0 that were found to do not agree with those calculated from theory. For example, R_0 was calculated to be 13.4 Å for the Tyr-thioamide pair and was found to be 16.6 ± 0.6 Å experimentally, assuming a FRET model. For the case of a Trp-thioamide FRET pair, we predicted R_0 to be 4.0 Å, but found a value of 15.3 ± 0.6 Å experimentally.

The data sets were also fit to a Dexter model (Eq. 2.8) using KaleidaGraph.

$$F = F_0\left\{1 - \frac{1}{1+k\exp\left(\frac{2R}{L_{Dex}}\right)}\right\} \qquad \text{(Eq. 2.8)}$$

Here F_0, k, and L_{Dex} (taken to be the sum of the chromophore radii) are adjustable parameters, and R is again the interchromophore separation in number of prolines or distance in angstroms.[10] Values of 7.64 ± 0.84 Å (R^2 = 0.99) and 7.38 ± 1.02 Å (R^2 = 0.98) were found for L_{Dex} for Tyr and Trp, respectively. These values are in relatively close agreement with those predicted by quantum mechanical calculations: 7.16 Å for

Tyr/thioamide and 7.44 Å for Trp/thioamide (see below for description of distance determination).

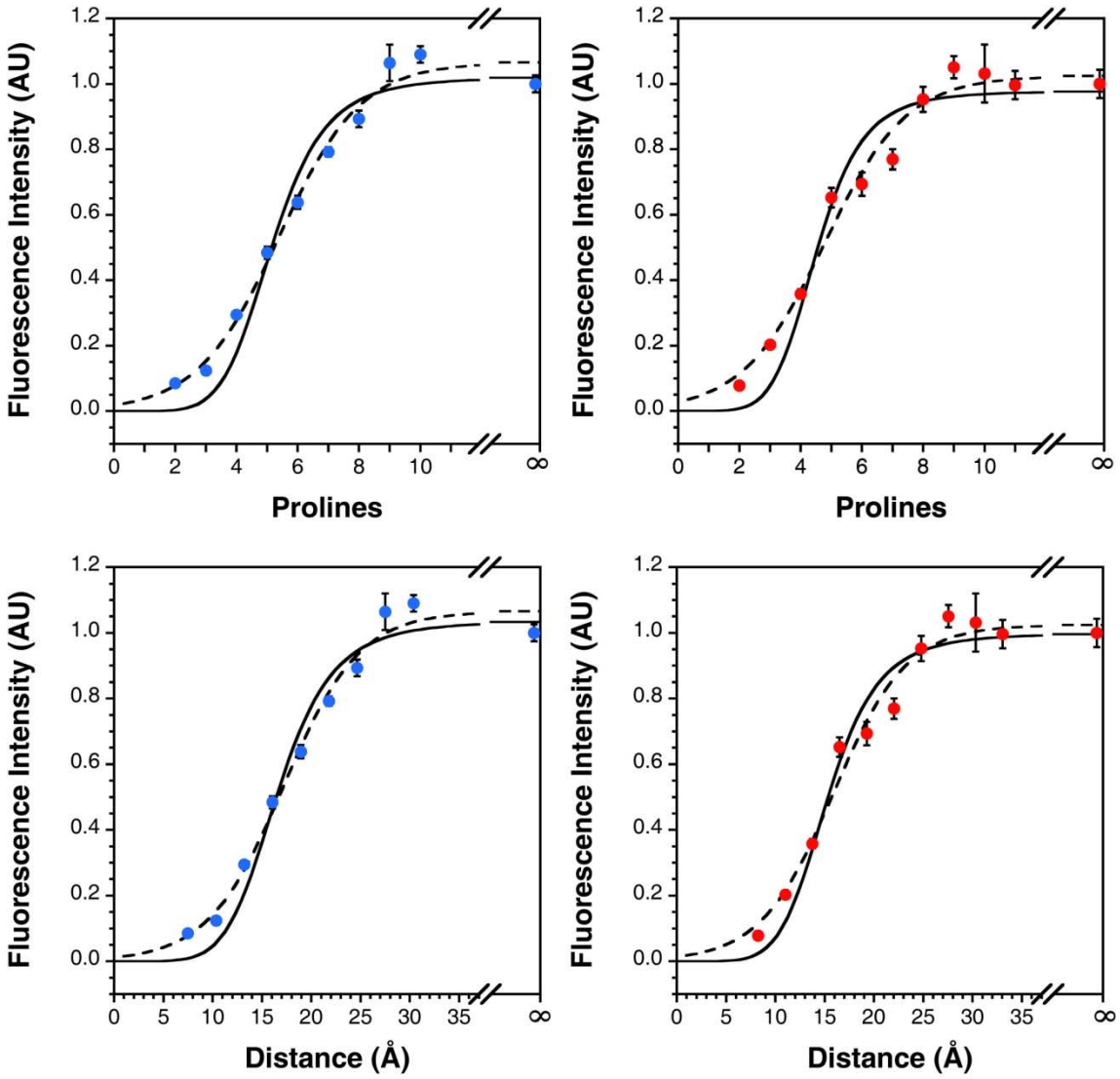

Figure 2.17 Fluorescence Intensity Fit to FRET and Dexter Models.
Tyrosine (left) and tryptophan (right) fluorescence data from fit to a FRET model (solid trace) or a Dexter model (dashed trace) with respect to chromophore separation in number of prolines (top) or distance in angstroms (bottom).

Fluorescence intensities were then converted to E_Q according to equation 2.6. The quenching efficiencies were fit to each model using equation 2.9 for FRET and equation 2.10 for Dexter.

$$E_Q = \frac{1}{1+\left(\frac{R}{R_0}\right)^6} \qquad \text{(Eq. 2.9)}$$

$$E_Q = \frac{1}{1+k\exp\left(\frac{2R}{L_{Dex}}\right)} \qquad \text{(Eq. 2.10)}$$

In each case, the values for F_0 obtained from the corresponding distance-in-angstroms plots were used to compute quenching efficiency. Using this method, R_0 was found to be 16.2 ± 0.5 Å (R^2 = 0.96) for the Tyr-thioamide pair and 15.1 ± 0.5 Å (R^2 = 0.95) for the Trp-thioamide pair. Again, these values are larger than those predicted by Förster theory. For the Dexter model, L_{Dex} was found to be 7.6 ± 0.7 Å (R^2 = 0.98) for the Tyr-thioamide pair and 7.5 ± 1.0 Å (R^2 = 0.96) for the Trp-thioamide pair.

Finally, the data were also fit to a general distance-dependent quenching (DDQ) model that has previously been used to describe electron transfer.

$$F = F_0\left\{1 - \frac{1}{1+k_a\exp\left(\frac{2R-a}{r_e}\right)}\right\} \qquad \text{(Eq. 2.11)}$$

In this equation, a is the distance of closest approach of the chromophores, r_e is the so-called characteristic distance, and k_a is the rate of the reaction at the distance a. As before, R is defined as the distance between the chromophores in prolines or angstroms. We note that the form of this equation is similar to that used to describe Dexter transfer, which also has a $1/e^R$ distance dependence. Both data sets were fit to equation 2.11 . The data were then converted to quenching efficiencies using equation 2.6 and fit to equation 2.12 with KaleidaGraph (Figure 2.18).

$$E_Q = \frac{1}{1 + k_a \exp\left(\frac{2R - a}{r_e}\right)} \qquad \text{(Eq. 2.12)}$$

Using this equation, r_e was found to be 3.79 ± 2.51 Å for Tyr and 3.75 ± 3.59 Å for Trp in the polyproline ruler system.

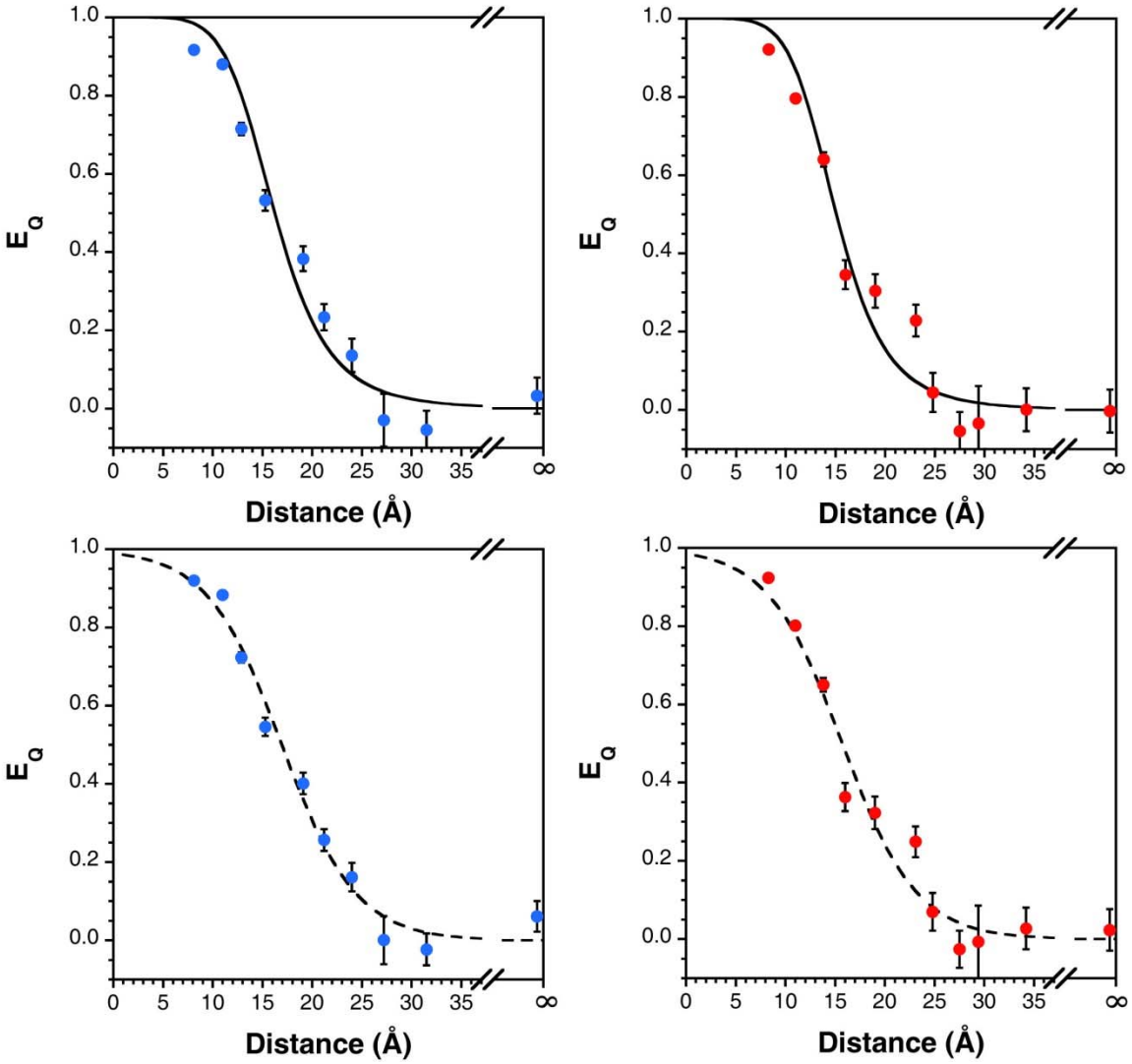

Figure 2.18 Proline Series Quenching Efficiency Described by FRET and Dexter Models. Thioamide quenching Tyr (left) and Trp (right) fluorescence efficiencies calculated with equation 2.6 and fit to $1/R^6$ distance dependence as described by FRET (Top) or $1/e^R$ distance dependence as described by Dexter transfer (Bottom). Error bars indicate standard error.

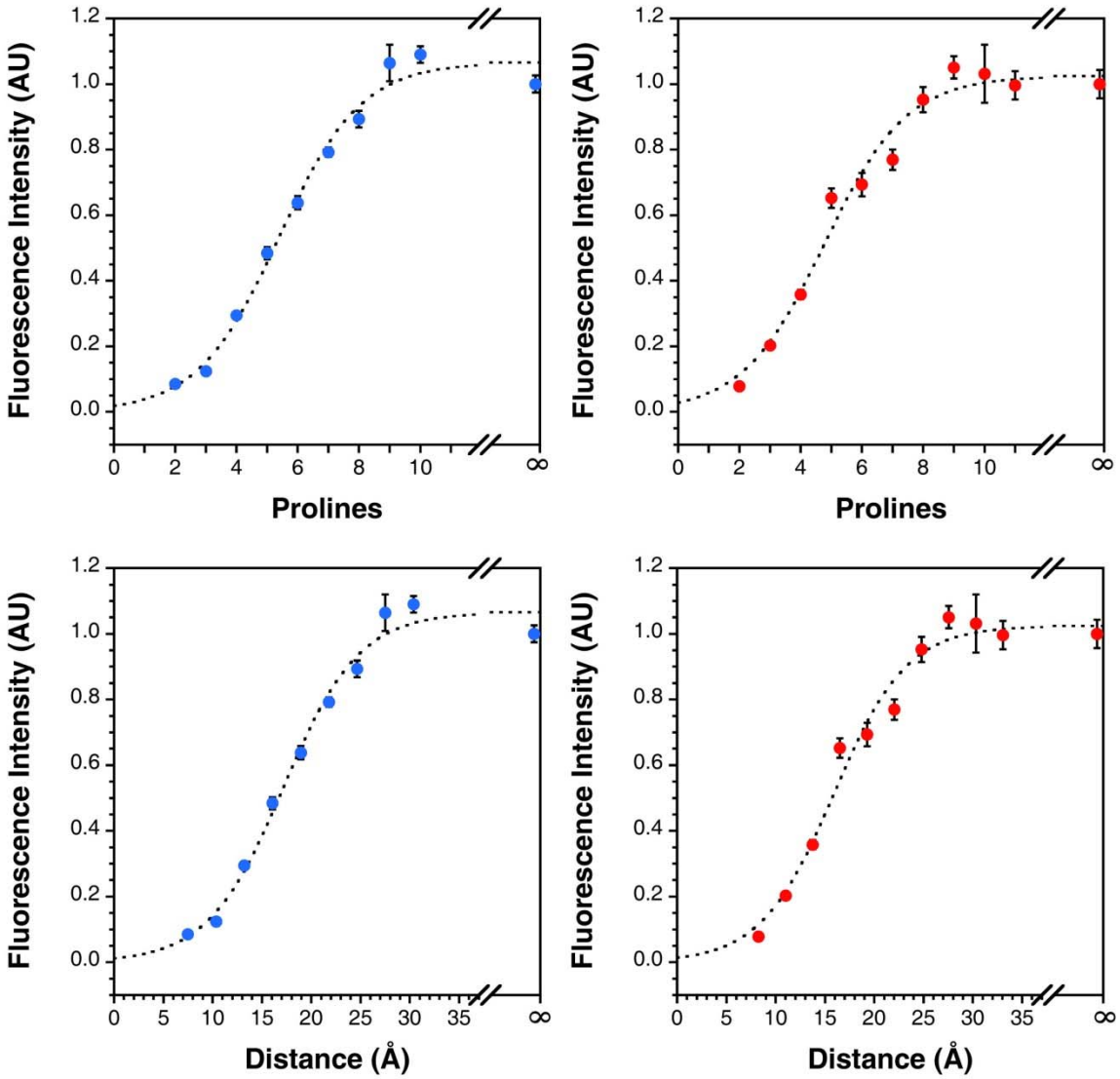

Figure 2.19 Proline Series Fluorescence Fit to a DDQ Model.
Tyr (left) and Trp (right) fluorescence data from Figure 2.15 fit to equation 2.11 with the chromophore separation in either number of prolines (Top) or distance in angstroms as determined by molecular dynamics simulations (bottom), described below. Error bars represent standard error.

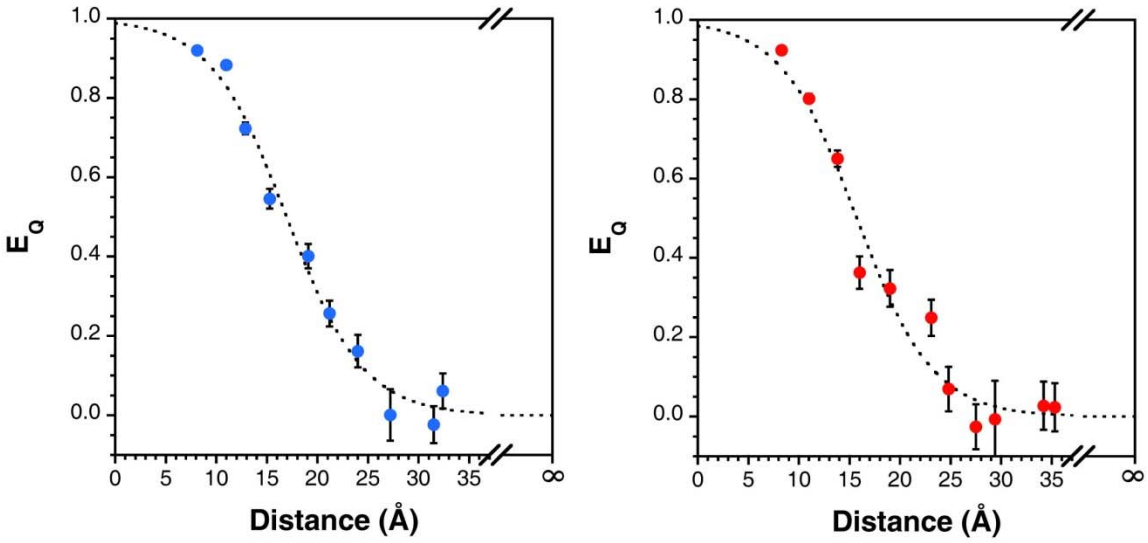

Figure 2.20 Proline Series Quenching Efficiency Described by a DDQ Model.
Thioamide quenching Tyr (left) and Trp (right) fluorescence efficiencies calculated with equation 2.6 and fit to $1/e^R$ distance dependence as described by the DDQ model (Equation 2.11). Error bars indicate standard error.

Molecular Dynamics Calculations: Polyprolines. Starting geometries for 22 proline series peptides (with N-terminal Leu, *C*-terminal Tyr or Trp and n intervening prolines, with n = 2-12) were generated using polyproline type II helix ϕ and ψ angles of - 78° and 149°, respectively. All peptide bonds were initially *trans* (i.e. ω = 180°). *N*- and *C*-termini were left as charged amines and carboxylates, respectively. These were inserted into a TIP3P water box with 10 Å periodic boundaries and no ions. Simulations were performed using the NAMD 2.7b2 package and the CHARMM27 all-atom force field.[147] Covalent bonds involving hydrogen were held rigid using the SHAKE algorithm, allowing a 2 fs time step. A cutoff distance of 12 Å was maintained in calculating nonbonded interactions. Bonded and nonbonded forces were evaluated at every time step, and full electrostatic forces were evaluated at every other time step through application of the particlemesh Ewald (PME) method.[148] The simulation was initiated at 100 K, equilibrated for 25 ps, and warmed to 300 K at a rate of 0.025 K/fs. After warmup, 5 x

10^6 step trajectories (10 ns) were run in the constant temperature and pressure (NPT) ensemble at 300 K.

Computational Studies were performed by E. James Petersson.

Chomophore Geometries: Polyprolines. For each 10 ns trajectory, a Tcl script was run in VMD (http://www.ks.uiuc.edu/) to cull distance and orientation information for the leucyl carbonyl and aromatic ring at every tenth timestep over the last 9 ns of the trajectory. The script identified the midpoint between CE1 and CE2 (CEmp) of Tyr or the midpoint between CD2 and CE2 (CDEmp) of Trp. Likewise, the script identified the Leu_1 C as the midpoint of the amide. (Figure 2.21) These midpoints serve as convenient anchors that are less sensitive to conformational fluctuations than a single atom position. The interchromophore distance (R) was determined as the CEmp-C distance for Leu-Pro_n-Tyr peptides and the CDEmp-Pro_n-C distance for Trp peptides). These distances are shown in Figure S10, plotted as a function of the number of Pro for the Tyr and Trp series peptides. Data for analogous calculations done on Leu-Pro_n-Phe were reported previously and are also shown in Figure 2.22.[115]

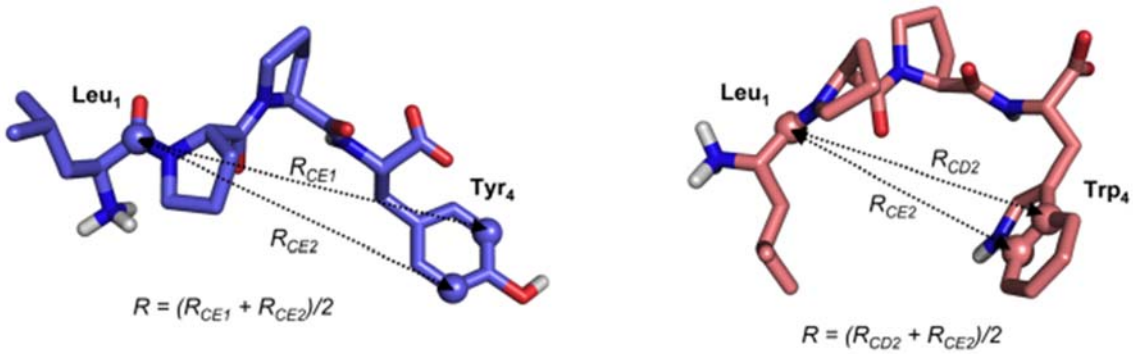

Figure 2.21 Interchomophore Distances.
Left: Timestep 1 from Leu-Pro_2-Tyr simulation shown with atoms used in determining distance (R) illustrated. Right: Timestep 278 from Leu-Pro_2-Trp simulation shown with atoms used in determining distance (R) illustrated. Images rendered using PyMOL.

The relationship between chromophore spacing and the number of prolines is relatively linear in all cases, thus a simple linear fit to the data was used to generate the distances in Å used on the abscissa of Figure 2.22. These equations are listed in Figure 2.21.

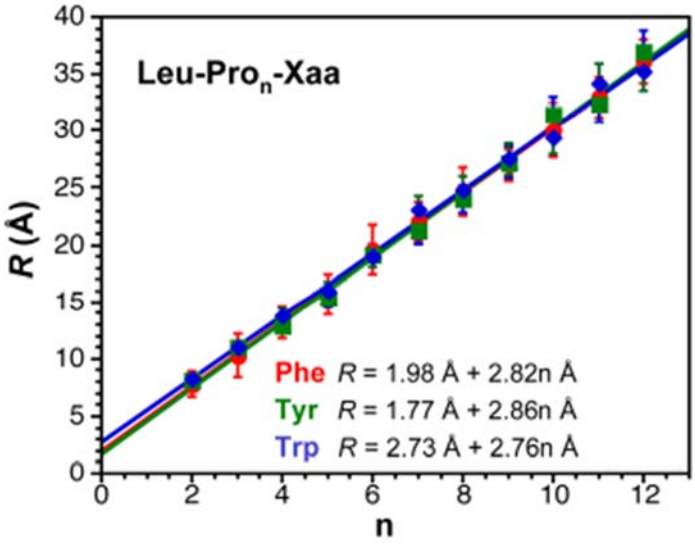

Figure 2.22 Interchromophore Distances from MD Simulations.
Average distances from last 9 ns of 10 ns simulations of polyproline peptides. Bars indicate RMSD. Equations indicate linear fits to data.

Table 2.3 Interchomophore Distances Determined from MD Simulations[a]

Pro_2	Tyr	Trp
Average	8.1	8.3
RMSD	1.8	1.7
Pro_3		
Average	11.0	11.0
RMSD	2.4	1.7
Pro_4		
Average	12.9	13.8
RMSD	1.2	1.7
Pro_5		
Average	15.3	16.0
RMSD	1.6	1.4
Pro_6		
Average	19.1	19.0
RMSD	2.1	1.8
Pro_7		
Average	21.2	23.1
RMSD	2.0	2.1
Pro_8		
Average	24.0	24.8
RMSD	1.9	1.3
Pro_9		
Average	27.2	27.5
RMSD	2.0	1.9
Pro_{10}		
Average	31.5	29.4
RMSD	2.6	1.9
Pro_{11}		
Average	32.4	34.2
RMSD	2.0	1.9
Pro_{12}		
Average	37.0	35.3
RMSD	2.1	2.6

[a] Distances (in Å) calculated as described in text.

Quantum Mechanical Calculations. Phenol and indole structures were optimized at the B3LYP level of theory with a 6-311G+ basis set using GAUSSIAN 03 (Gaussian, Inc.; Wallingford, CT). For Dexter fitting, the phenol (Tyr) diameter (d_{Tyr} = 7.38 Å) was computed as the sum of the O12 to C6 distance (4.16 Å) plus the van der Waals radii of oxygen (1.52 Å) and carbon (1.70 Å). The van der Waals radius (R_{Tyr} = 3.69 Å) was computed as one half of this value. The indole (Trp) diameter (d_{Trp} = 7.94 Å) was computed as the sum of the N1 to C8 distance (4.54 Å) plus twice the van der Waals radius of carbon (1.70 Å). The van der Waals radius (R_{Trp} = 3.97 Å) was computed as one half of this value. The thioamide van der Waals radius (R_{SCN} = 3.47 Å) has been reported previously.

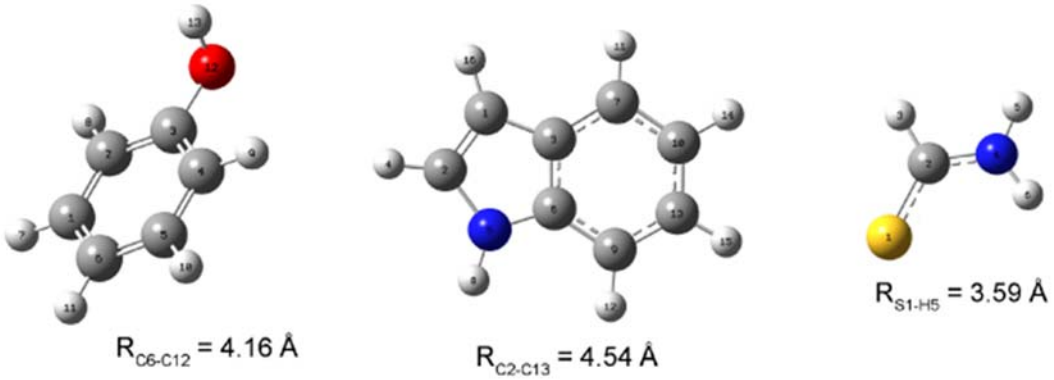

Figure 2.23 Chomophore Lengths.
B3LYP/6-311G+ minimized geometries for phenol (left), indole (center), and thioformamide (right). Interatomic distances used in calculating chomophore radii are indicated. Figure rendered using Gaussview 4.1.

Cloning of Calmodulin Expression Constructs. A plasmid containing the wild-type chicken calmodulin (CaM) gene was provided by Joshua Wand from the University of Pennsylvania School of Medicine. An insert containing the CaM gene (458 base pairs from start to stop codon) was cloned into a pET15b vector (Novagen, Gibbstown, NJ) between the NcoI and XhoI cut sites. Quikchange® site-directed mutagenesis was used to

generate the following CaM mutants: $Y_{100}F$, $Y_{139}F$, $Y_{100}W$, and $Y_{139}W$. The mutant plasmids were verified by DNA sequencing analysis with a T7 promoter primer.

A. CaM Gene

```
5'-atggctgatcaactgacagaagagcagattgcagaattcaaagaagctttttcactattt
gacaaggatggtgatggtactataactacaaaggagttggggactgtgatgagatcacttggt
cagaaccccacagaagcagaattacaggacatgatcaatgaagtagacgctgatggcaatggc
acaattgacttcccagagtttctgacaatgatggcaagaaaaatgaaagatacagatagcgaa
gaagaaattagagaagcgttccgtgtgtttgacaaggatggtaatggttacattagtgctgca
gaacttcgtcatgtgatgacaaatcttggggagaagctaacagatgaagaagttgatgaaatg
attagggaagcagacattgatggtgatggtcaagtaaactatgaagagtttgtacagatgatg
acagcgaagtga-3'
```

B. Chicken CaM Amino Acid Sequence

```
MADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLT
MMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEF
VQMMTAK
```

C. DNA Oligomers used for Quikchange® Mutagenesis

Y100F_For	5'-gacaaggatggtaatggtttcattagtgctgcagaactt-3'
Y100F_Rev	5'-aagttctgcagcactaatgaaaccattaccatccttgtc-3'
Y139F_For	5'-agacattgatggtgatggtcaagtaaactttgaagagtttgtaca-3'
Y139F_Rev	5'-tgtacaaactcttcaaagtttacttgaccatcaccatcaatgtct-3'
Y100W_For	5'-tgacaaggatggtaatggttggattagtgctgcagaacttcg-3'
Y100W_Rev	5'-cgaagttctgcagcactaatccaaccattaccatccttgtca-3'
Y139W_For	5'-tgatggtgatggtcaagtaaactgggaagagtttgtacagatgatgac-3'
Y139W_Rev	5'-gtcatcatctgtacaaactcttcccagtttacttgaccatcaccatca-3'

DNA Sequence of Chicken CaM Gene from Start to stop codon. B: Amino acid sequence of chicken CaM. Y_{100} and Y_{139} residues highlighted in red. C: Forward and reverse DNA oligomers used for site-directed mutagenesis.

Protein Expression and Purification. The plasmid containing the chicken CaM gene was used to transform *Escherichia coli* BL21(DE3) cells. Transformed cells were selected on the basis of ampicillin resistance. Single colonies were used to inoculate 50 mL of M9 minimal media supplemented with ampicillin (100 μg/mL). To an autoclaved solution containing 42.3 mM Na_2HPO_4, 22.0 mM KH_2PO_4, and 8.5 mM NaCl, the

following autoclaved solutions were added per liter of M9 salts: 10 mL of 10% NH_4Cl, 1 mL of 2 M $MgSO_4$, 1 mL of 15 mg/mL $FeCl_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL $ZnCl_2$ (in acidified H_2O), and 2 mL of 10% Bacto™ Yeast Extract. The primary 50 mL culture was incubated at 37 °C with shaking at 250 rpm overnight. The cells were harvested at 5000 g for 15 min and the resulting pellet was re-suspended in 1 L of M9 minimal media supplemented with ampicillin. The 1 L culture was incubated at 37 °C with shaking at 250 rpm until the absorbance at 600 nm reached 0.9 AU. Protein expression was induced with isopropyl D-galactoside (IPTG), and the culture was incubated at 25 °C for an additional 12 h. The cells were again harvested at 5000 g for 15 minutes and the resulting pellet was suspended in 15 mL of MOPS resuspension buffer: 50 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), 100 mM KCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol, pH 7.5. Cell lysis was achieved by incubating the resuspended cells with lysozyme (150 μg/mL) for 1 h at room temperature followed by brief sonication. Following sonication, the cell lysate was allowed to cool on ice for 5 min. $CaCl_2$ was added to the sonicated lysate to a final concentration of 5 mM prior to centrifugation for 20 minutes at 30,000 g, 4 °C. CaM was purified from the cleared cell lysate using a phenyl-sepharose CL-4B column with EDTA as eluant. Using a total resin bed volume of 20 mL, the column was first equilibrated with 4 column volumes of Buffer A (50 mM Tris base, 1 mM $CaCl_2$, pH 7.5). After the cleared cell lysate was loaded and allowed to pass through the resin, the column was washed with 4 column volumes of Buffer A, 4 column volumes of high-salt Buffer B (50 mM Tris base, 0.5 M NaCl, 0.1 mM $CaCl_2$, pH 7.5), and an additional 2 column volume washes of Buffer A to restore low-salt conditions. CaM was eluted with Buffer C (10 mM Tris base,

10 mM EDTA, pH 7.5) and collected in 4 mL fractions until absorbance at 280 nm was no longer detected. A second column purification was performed on the first batch of eluted fractions (re-saturated with $CaCl_2$ to a concentration of 20 mM) to obtain CaM with high levels of purity. Column fractions were dialyzed against 10 mM ammonium bicarbonate (pH 8.0) and stored as a lyophilized powder at -20 °C. SDS-PAGE analysis was performed to analyze dialyzed CaM elution fractions.

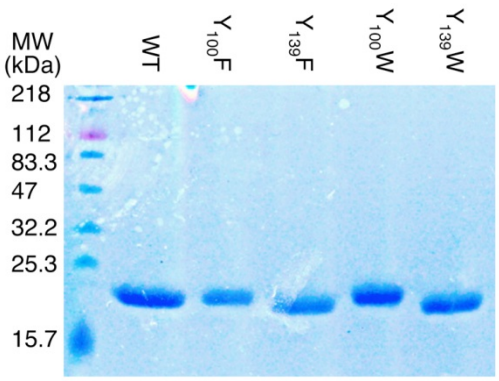

Figure 2.24 SDS-PAGE gel of purified CaM mutants.
Approximately 10 μM samples denatured at 95 °C in the presence of SDS.

Calmodulin Circular Dichroism Measurements. Mutant stability was evaluated using temperature-dependent circular dichroism (CD) spectroscopy. Since calcium-containing CaM (holo CaM) is thermostable (T_m > 90 °C), we specifically examined the thermal unfolding of the apo protein, for which the reported T_m is 55 °C.[149] To compare the apo and holo forms, purified protein was dialyzed against 2 mM EDTA in 50 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 6.70 or 10 mM $CaCl_2$ in 50 mM HEPES pH 6.70. CD data were obtained from approximately 20 μM protein samples diluted in the appropriate buffer monitoring at 222 nm, between 5 and 95 °C, using the variable temperature module provided with the Aviv 410 CD spectrometer. Data were collected with a 1 °C/min slope, 30 s averaging time, 2 min temperature equilibration, 5 s

response, and 1 nm band width. The resulting ellipticity (θ_D) measurements were transformed to molar residue ellipticity values (θ) using

$$\theta = \theta_D/(c\ell n_R) \qquad \text{(Eq. 2.13)}$$

where c is concentration (M), ℓ is the path length (cm), and n_R is the number of residues. To determine fraction folded (f_f) for the apo protein, linear baselines were fit to the low temperature ($\theta_F = m_F T + b_F$) or high temperature ($\theta_U = m_U T + b_U$) data. The full data range was then fit to equation 2.14 where $K = K = e^{-(\Delta H - T\Delta S)/RT}$, ΔH and ΔS are adjustable parameters and R = 8.3145 J•mol^{-1}•K^{-1}. Plots are shown for each mutant in Figure 2.25.

$$\theta = \theta_F(T)f_f(T) + \theta_u(T)(1 - f_f(T)) \quad f_f = K/(1+K) \qquad \text{(Eq. 2.14)}$$

Since the holo protein did not melt, only the molar residue ellipticity as a function of temperature is shown for each mutant (Figure 2.25).

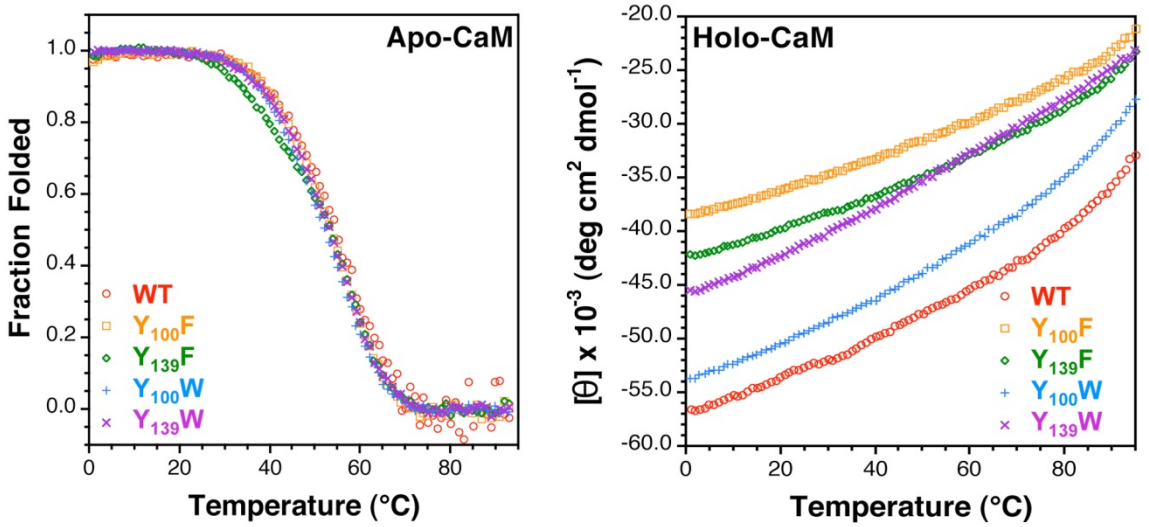

Figure 2.25 Temperature-Dependent Circular Dichroism Spectroscopy of Wildtype CaM and Mutants.
Left: Fraction folded as a function of temperature as determined by molar residue ellipticity (□) at 222 nm measured for 20 μM solutions of each protein in the absence of Ca^{2+}. Right: Molar residue ellipticity (θ_D) at 222 nm measured for 20 μM solutions of each protein in the presence of Ca^{2+}.

Peptide Binding Experiments. All peptide binding experiments were conducted in 15 mM HEPES buffer, 140 mM KCl, and 6 mM $CaCl_2$, pH 6.70. Dry peptides and lyophilized protein were brought up in a minimal amount of buffer to make fresh concentrated stock solutions for each experiment. The absorbance at 280 nm of each protein solution was used to determine its concentration ($\varepsilon_{280,WT}$ = 3000 M^{-1}•cm^{-1}; $\varepsilon_{280,Y100F}$ = 1500 M^{-1}•cm^{-1}; $\varepsilon_{280,Y139F}$ = 1500 M^{-1}•cm^{-1}; $\varepsilon_{280,Y100W}$ = 7400 M^{-1}•cm^{-1}; and $\varepsilon_{280,Y139W}$ = 5400 M^{-1}•cm^{-1}). For each mutant, solutions were prepared that contained approximately 10 μM protein and a variable amount of peptide, ranging from 0 to approximately 20 μM and representing at least seven distinct concentrations. Each sample was prepared in triplicate using one of two methods. For at least one trial of each mutant, concentrated stocks of protein and peptide were combined with buffer to prepare solutions at each peptide-concentration step. The fluorescence of each solution was measured as described below. For the second trial of each mutant, a sample of protein was diluted to approximately 10 μM in buffer. The fluorescence of this solution was measured as described below and then a concentrated peptide solution of either pOCNC or pOCNC-F_1' was added to obtain a sample that was approximately 20 μM in peptide. The fluorescence of this solution was then recorded and the solution was diluted with 10 μM protein to prepare samples of varying peptide concentration, while maintaining constant protein concentrations. The fluorescence of each sample was measured between dilutions. No variation was observed in the fluorescence of samples prepared by either method. Fluorescence spectra of pOCNC and pOCNC-F_1' were acquired in the absence of protein at each concentration with excitation wavelengths of 275 nm and 295 nm. Essentially no fluorescence was observed exciting the peptides at 295 nm. To confirm

that the observed quenching was coming from a peptide binding event, we incubated 10 μM CaM in buffer in the presence and absence of 20 μM thioacetamide and measured the fluorescence by exciting at 275 nm and 295 nm; no change in fluorescence was observed.

Corrected fluorescence measurements as the average of three scans were taken of each sample at 25 °C using quartz fluorometer cells with path lengths of 1.00 cm. For WT, $Y_{100}F$, and $Y_{139}F$ experiments, the excitation wavelength was 275 nm and emission data were collected from 280 -400 nm. For $Y_{100}W$ and $Y_{139}W$ experiments, the excitation wavelength was 295 nm and emission data were collected from 300 - 500 nm. For all experiments, the excitation and emission slit widths were 5 nm, the scan rate was 120 nm/min, the averaging time 0.5 s, and the data interval 1.0 nm. Representative fluorescence spectra are shown in Figure 2.26.

Since pOCNC and pOCNC-F_1' are slightly fluorescent when excited 275 nm (Figure 2.26), all binding experiment data was background corrected for contributions from peptide fluorescence. The raw signal of the peptide in buffer was subtracted from the raw signal of the protein•peptide complex at each concentration step. These background-corrected values were normalized to the fluorescence of the protein in the absence of peptide (i.e. to the 0 μM sample). The quenching efficiencies listed in Table 2.1 were estimated from the normalized fluorescence at 1:1 binding stoichiometry.

Predictions of the quenching efficiency for each protein were calculated using our empirical ruler and using Förster theory. The average distance between the thiocarbonyl carbon atom of pOCNC-F_1' and either the ε^1 and ε^2 carbons of Tyr or the δ^2 and ε^2 carbons of Trp were calculated using PyMol from the NMR structure PDB 1SY9. Using this metric the distance between the thioamide and tyrosine was found to be 19.3 Å for

Y_{100} and 15.9 Å for Y_{139}. For Trp, the distance was found to be 19.0 Å for W100 and 17.0 Å for W139. Equation (2.6) was used directly to calculate quenching efficiencies at these distances for each of the mutants ($Y_{100}F$, $Y_{139}F$, $Y_{100}W$, and $Y_{139}W$). For the WT protein, we weighted the contribution from each Tyr using the fluorescence sensitivity, which we defined to be the product of the extinction coefficient and the quantum yield (Φ_{Y100} = 0.105; Φ_{Y139} = 0.061) at 275 nm. For Förster calculations, we used equation 2.1 and the values of *J* that we had previously determined from the spectral overlap of thioleucylalanine ester with Trp or Tyr to determine a value of R_0 for each fluorophore. In these calculations, we took κ^2 to be 2/3, approximated *n* to be 1.33, and substituted the appropriate quantum yield for each donor (Φ_{W100} = 0.15; Φ_{W139} = 0.29). We used these values of R_0 in equation 2.1 with the distances inferred from the NMR structure to compute the theoretical quenching efficiency for a FRET mechanism for each mutant. Again, we weighted the quenching efficiency contributions for each Tyr in the wildtype enzyme with their fluorescence sensitivity factors as described above. The background-corrected, normalized fluorescence data for the titration of pOCNC-F_1' was fit to equation 2.15 using KaleidaGraph.

$$y = 1 - R\frac{\left(\frac{1}{K_a} + [P]_0 + [P]_0[L]_0\right) - \sqrt{\left(\frac{1}{K_a} + [P]_0 + [P]_0[L]_0\right)^2 - 4[P]_0^2[L]_0}}{2[P]_0} \qquad \text{(Eq. 2.15)}$$

Here we define $[L]_0$ and $[P]_0$ to be the total concentration of peptide and protein, respectively. K_a is the equilibrium constant for the peptide binding to the protein to form a 1:1 protein-peptide complex and *R* is an instrumental response parameter. K_a and *R* are adjustable parameters in the fit.

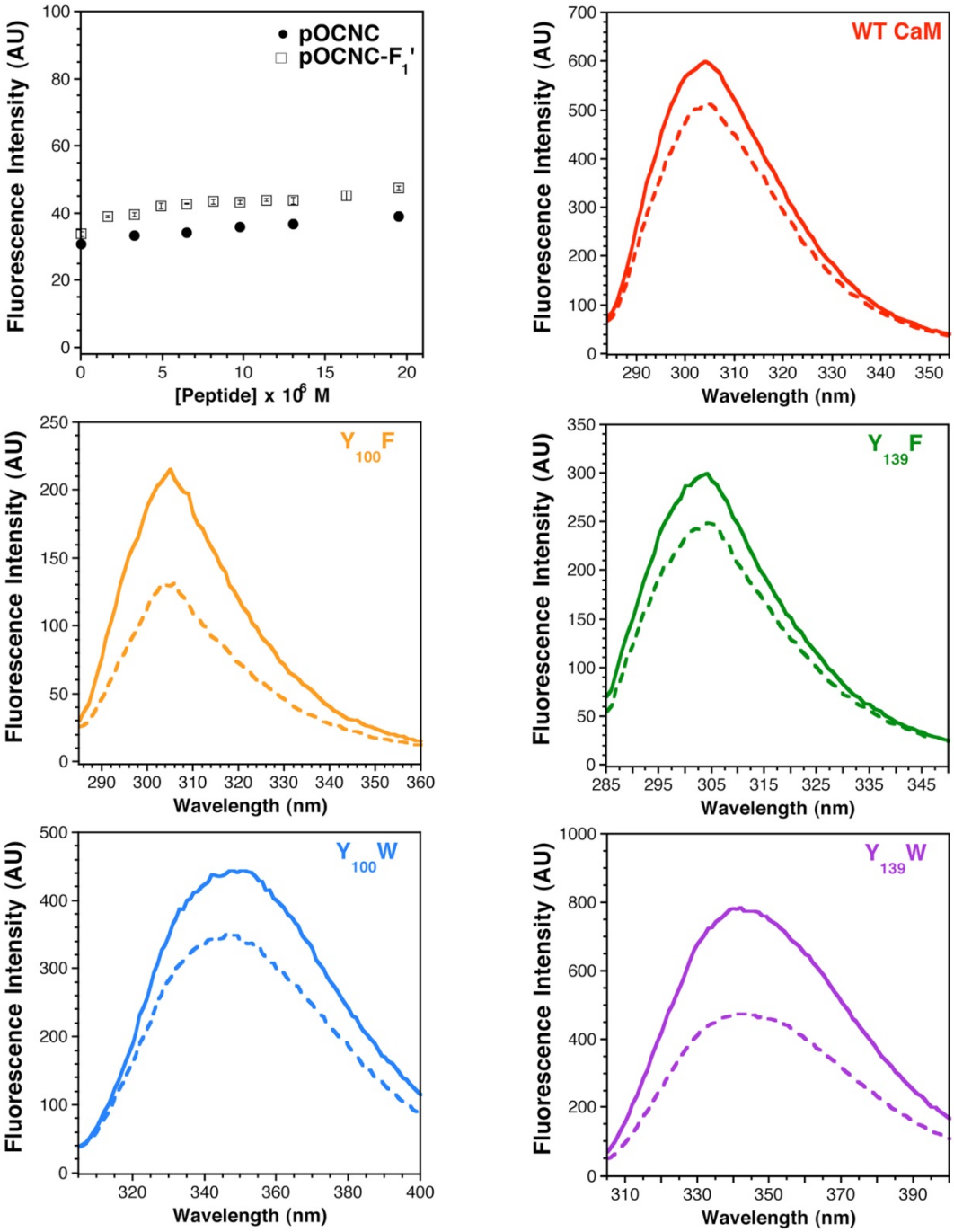

Figure 2.26 Representative Fluorescence Spectra of CaM Mutants.
Top Left: Fluorescence of pOCNC (circles) and pOCNC-F_1' (squares) as a function of peptide concentration in buffer (ex. 275 nm; em. 305 nm). Top Right: Emission spectra of 10 μM WT CaM in the presence (dashed trace) and absence (solid trace) of 13 μM pOCNC-F_1' (ex. 275 nm). Center: Emission spectra of 8.5 μM Phe mutants with (dashed) and without (solid) 13 μM pOCNC-F_1' (ex. 275 nm). Bottom Left: Emission spectra of 9.4 μM $Y_{100}W$ CaM with (dashed) and without (solid) 13 μM pOCNC-F_1' (ex. 295 nm). Bottom Right: Emission spectra of 13 μM $Y_{139}W$ CaM with (dashed) or without (solid) of 13 μM pOCNC-F_1' (ex. 295 nm).

Semi-Synthesis of αSV′$_3$C$_9$W$_{94}$ *was performed by Solongo Batjargal and Yanxin Wang.*[116]

Expression and Purification of WT αS and αSW$_{94}$. The plasmid containing the αS gene was used to transform *Escherichia coli* BL21(DE3) cells. Transformed cells were selected on the basis of ampicillin resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with ampicillin (100 μg/mL). The primary 5 mL culture was incubated at 37 °C with shaking at 250 rpm for 5 hours. 500 mL of LB media supplemented with ampicillin was inoculated with 1 mL of the primary culture. The 500 mL culture was incubated at 37 °C with shaking at 250 rpm for 20 hours. The cells were harvested at 5000 g for 15 min and the resulting pellet was resuspended in a high-salt buffer (750 mM NaCl, 10 mM Tris, pH 7.6, 1 mM EDTA, 1 mM PMSF) and sonicated. Following sonication, the cell lysate was boiled for 30 minutes prior to centrifugation for 20 minutes at 30,000xg, 4 °C. The cleared lysate was dialyzed overnight against purification buffer (50 mM Tris, 150 mM NaCl, pH 8.0) prior to loading on a Superdex 200 column (25 cm) connected to a BioCad Sprint (FPLC) system. FPLC fractions were dialyzed against phosphate buffered saline (pH 7.0) at 4 °C and stored at -80 °C until further use. SDS-PAGE analysis was performed to analyze the purity of αS elution fractions.

Aggregation Assays. Aggregation experiments were performed according to literature precedent. 350 μL samples for aggregation assays were prepared (97 μM WT αS with 3 μM αSW$_{94}$ or 3 μM αSV′$_3$C$_9$W$_{94}$) in phosphate buffered saline (pH 7.0) containing 1 mM β-mercaptoethanol. Aggregation was seeded by the addition of approximately 10% (wt/wt) pre-formed WT αS fibrils. The samples were incubated at 37 °C for 4 to 6 days with continuous shaking at 1100 rpm. 40 μL aliquots were periodically removed to

monitor changes in tryptophan fluorescence and ThT or Congo Red binding. Protocols and examples of primary fluorescence or absorbance data are given below.

Fluorescence Spectroscopy. Fluorescence experiments were performed following literature precedent. 40 μL aliquots were periodically removed from aggregation experiments to monitor changes in tryptophan fluorescence and ThT binding. The 40 μL aliquots were diluted to 120 μL with phosphate buffered saline prior to loading the sample into quartz cuvettes. To selectively monitor changes in Trp fluorescence in the presence of tyrosine, the excitation wavelength was 295 nm and emission data was collected from 315 to 385 nm. ThT binding was determined by directly adding 1 mL of ThT solution (20 μM in phosphate buffered saline) to the cuvette after monitoring Trp fluorescence. The samples were incubated within the cuvettes for two minutes prior to collecting ThT fluorescence data. The excitation wavelength was 446 nm and emission data was collected from 460 to 600 nm. For all experiments, the temperature was 37 °C, the excitation and emission slit widths were 5 nm, the scan rate was 120 nm/min with an averaging time of 0.5 s, and the data interval 1.0 nm. Examples of primary spectra are shown in Figure 2.27 and averaged data for Trp and ThT fluorescence from four aggregation trials are shown in Table 2.4.

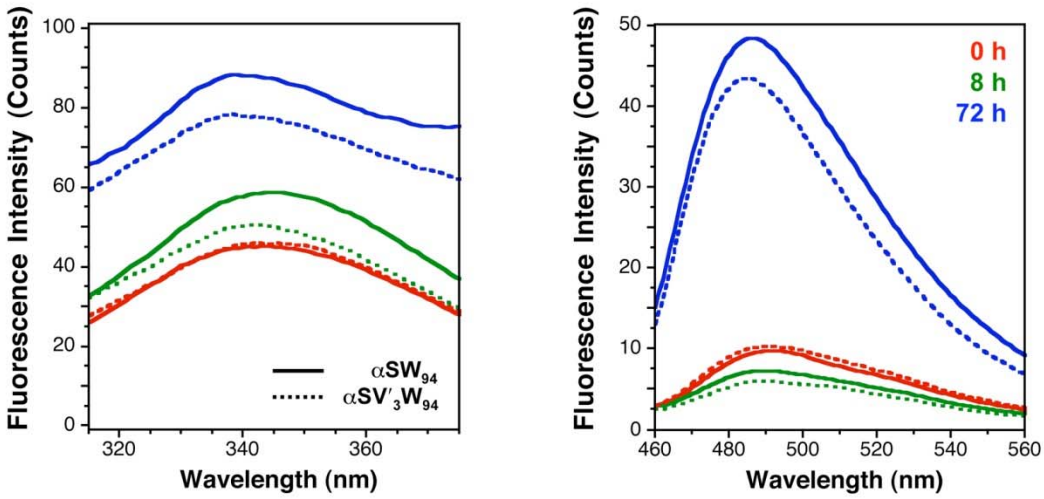

Figure 2.27 Examples of Fluorescence Data from αS Aggregation Experiments.
Left: Trp fluorescence, excited at 295 nm. Right: ThT fluorescence, excited at 436 nm. In both cases, red, green, and blue indicate data at 0, 8, and 72 h respectively. Solid lines: 1:30 αSW_{94}/WT mix. Broken lines 1:30 $\alpha SV'_3C_9W_{94}$/WT mix.

Table 2.4 Normalized Fluorescence Data from Aggregation Experiments

Time (h)	Oxo F_{Trp}	Oxo F_{Trp} SE	Thio F_{Trp}	Thio F_{Trp} SE	Eq	Eq SE	Oxo F_{ThT}	Oxo F_{ThT} SE	Thio F_{ThT}	Thio F_{ThT} SE
0	1		1		1		1		1	
4	1.155	0.052	1.141	0.015	0.988	0.046	1.117	0.045	1.040	0.096
8	1.115	0.012	1.079	0.020	0.968	0.021	1.034	0.106	1.040	0.131
16	1.255	0.079	1.182	0.059	0.942	0.076	1.564	0.169	1.263	0.155
32	1.379	0.081	1.214	0.077	0.881	0.076	2.372	0.357	1.485	0.221
72	2.134	0.188	1.809	0.090	0.848	0.086	3.323	1.034	2.871	1.005
96	2.201	0.201	1.851	0.192	0.841	0.116	2.480	0.576	2.260	0.837

F_{Trp} at 350 nm and F_{ThT} at 490 nm collected as described above for 1:30 αSW_{94}/WT mix (Oxo) or 1:30 $\alpha SV'_3C_9W_{94}$/WT (Thio). SE = Standard error

Congo Red Absorbance Assay for Aggregation Experiments. CR assays were performed following literature precedent. After 96 hours of aggregation, a 20 μL aliquot from the aggregation assay was removed and added to 110 μL of 20 μM CR in phosphate buffered saline (pH 7.0). A blank sample was prepared by adding 20 μL of buffer to the CR solution. All samples were incubated at room temperature for 30 minutes prior to collecting absorbance data.

PAGE Gel Analysis of Aggregation Experiments. Immediately following the assembly of the aggregation reaction, a 10 μL aliquot from each sample was removed for PAGE gel analysis and stored at - 80 °C. After 96 hours of aggregation, an additional 10 μL aliquot was removed from each sample. Each aliquot was diluted to 30 μL with MilliQ water and centrifuged for 45 min at 13,200 rpm to determine the loss of soluble protein post-aggregation. 15 μL of the resultant supernatant was removed from the centrifuged samples and run on an 18% SDS-PAGE gel. The gel was stained with Coomassie blue and scanned.

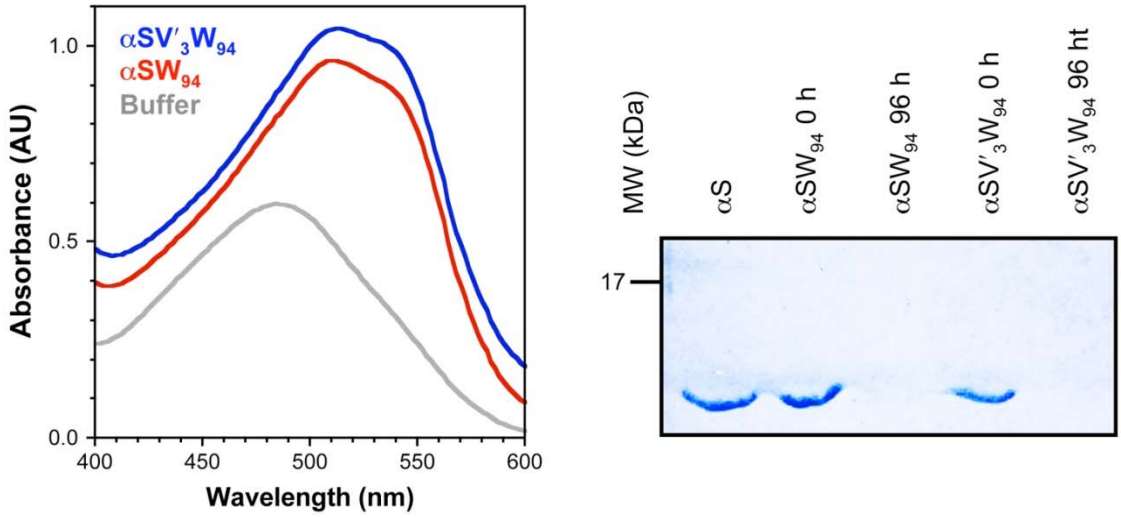

Figure 2.28 Congo Red and PAGE Gel Analysis of Aggregation Experiments.
Left: CR absorption assay shows a red shift and increase in absorbance, indicating fibrillization of αSV′$_3$C$_9$W$_{94}$/αS mixtures or αSW$_{94}$/αS mixtures after 4 days. Right: Gel band intensity of soluble fraction before aggregation (Pre) and after 4 days of aggregation (Post) shows a loss of soluble monomer due to fibrillization.

Chapter 3 : Incorporation of Minimally-Perturbing Unnatural Probe Pairs into Full-Length α-Synuclein for Folding Studies

§ 3.1 Labeling Proteins with Fluorophore/Thioamide Förster Resonant Energy Transfer Pairs by Combining Unnatural Amino Acid Mutagenesis with Native Chemical Ligation

Thioamide Quenching of Cnf in Proteins. Thioamides can quench fluorophores that have spectral emission overlap with thioamide absorption, such as cyanophenylalanine (Cnf) and tyrosine (Tyr), through a Förster resonance energy transfer FRET mechanism.[115,150] Other fluorophores are quenched by thioamides through a photoinduced electron transfer (PET) mechanism, like tryptophan, coumarin, and acridone.[113,151] Here we will focus on thioamide FRET partners, for which more reliable distance information can be obtained by applying Förster theory. To use thioamide FRET quenching to study protein conformation, we generate constructs in which we label the protein with a Cnf donor fluorophore and a thioamide analog of one of the natural amino acids. We denote these analogs using either the three or one letter amino acid code with a prime symbol indicating the thiocarbonyl (thioleucine is shown in Figure 3.1). Distance constraints are assigned by computing quenching efficiency ($E_Q = 1 - F_{Thio}/F_{Oxo}$) from the fluorescence of equimolar concentrations of Cnf/thioamide-labeled protein (F_{Thio}) and a Cnf-labeled control (F_{Oxo}). After appropriate corrections to the Förster radius (R_0) for each Cnf location, the Cnf/thioamide separation (R_{FRET}) is determined according to Förster theory. (Figure 3.1 and discussion below) These distance measurements are used to monitor conformational changes in the protein.

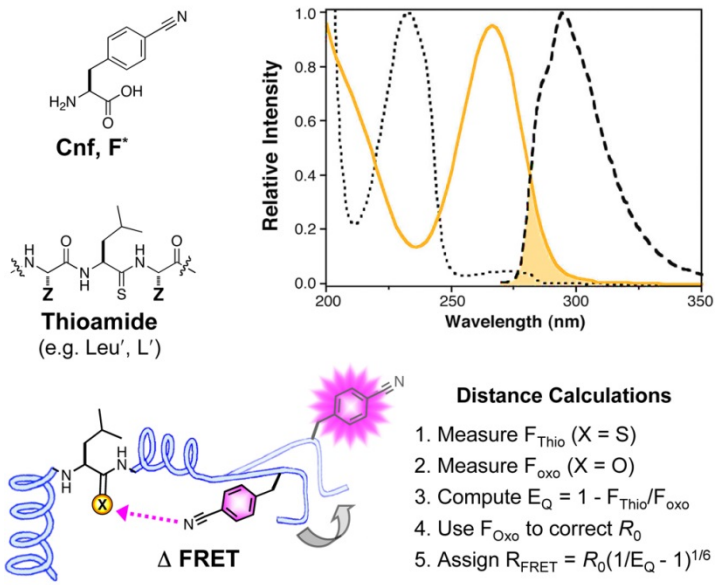

Figure 3.1 Cyanophenylalanine/Thioamide FRET Interactions.
Top Left: Cyanophenylalanine (Cnf, F*) and leucyl thioamide (Leu', L'). Top Right: Normalized absorption (dotted black line) and fluorescence emission (dashed black line) spectrum of Cnf and absorption spectrum of Leu'Ala (yellow line). The shaded area indicates the spectral overlap that contributes to FRET. Bottom: Distance (R_{FRET}) calculations are made using Förster theory as described in text.

All of our initial studies used small peptides and proteins in which the thioamide was incorporated by solid phase peptide synthesis (SPPS). In order to use fluorophore/thioamide pairs to monitor structural changes in full-sized proteins, we needed to identify methods for incorporating these two modifications with high efficiency. Happily, we have found that thioamides are compatible with native chemical ligation (NCL) reactions.[116] NCL is a valuable method for synthesizing proteins by fragment condensation of a C-terminal thioester peptide and an N-terminal Cys peptide in buffered aqueous media.[152,153] We found that thioamides could be placed at nearly any position in the thioester or Cys peptide fragments. Furthermore, we found that thioamides could be used in NCL reactions in which the N-terminal Cys fragment was expressed as a protein construct in *E. coli* (expressed protein ligations). In general, we wish to minimize the need for manual synthesis by maximizing the role of biosynthesis in preparing labeled

proteins. These protein ligations can limit the portion that must be prepared by SPPS to a few amino acids around the thioamide. SPPS is deemed necessary for the incorporation of backbone thioamide modifications, but sidechain fluorophores like Cnf can be incorporated during ribosomal biosynthesis.

To easily incorporate Cnf, we have employed unnatural amino acid (Uaa) mutagenesis methods pioneered by Schultz and coworkers.[22,23] These methods allow one to insert a Uaa at a specific site in the protein by using a so-called orthogonal aminoacyl tRNA synthetase (aaRS) that is selective for the Uaa. The UaaRS charges a tRNA that recognizes the UAG stop codon. The UAG stop codon is mutated into the protein coding sequence at the site of interest. Plasmids coding for the UaaRS, tRNA, and protein of interest are transformed into *E. coli*. After growth in the presence of the Uaa, protein containing the unnatural amino acid is purified for use in biochemical experiments or subsequent NCL reactions.

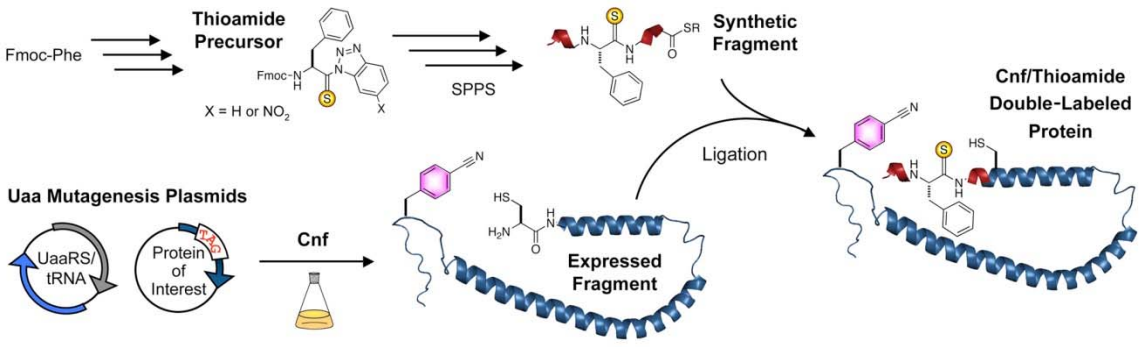

Figure 3.2 Combining Unnatural Amino Acid Mutagenesis and Native Chemical Ligation. Thioamide-bearing peptides are synthesized on solid phase starting from activated benzotriazole precursors and standard Fmoc-protected amino acids. For N-terminal thioamides, a thioester is installed as a masked group or formed by post-synthetic activation. To generate the Cnf-labeled protein fragment, *E. coli* cells are transformed with two DNA plasmids; one encoding the protein with a UAG stop codon at the site of interest, and one encoding an aminoacyl tRNA synthetase that is selective for an unnatural amino acid (UaaRS) with a tRNA that recognizes the UAG stop codon. Protein is expressed under standard conditions and the N-terminal Cys generated by proteolysis. Ligation of the two fragments gives the full length protein, labeled with both Cnf and the thioamide.

Here, we made use of Uaa mutagenesis to incorporate donor fluorophores for thioamide FRET experiments examining both protein/protein interactions and intramolecular measurements of protein folding. In intermolecular experiments, no manipulation of the Uaa-containing protein is required following purification. The protein is simply mixed with a thioamide-labeled peptide or protein and quenching observed. We use this strategy to examine a series of pairwise interactions in which Cnf is incorporated in the calcium sensor protein calmodulin (CaM) and the thioamide is incorporated in a CaM-binding peptide. The combinatorial nature of these experiments allows one to rapidly generate a matrix of *m* x *n* FRET measurements from *m* Cnf-labeled proteins and *n* thioamide-labeled peptides. To carry out intramolecular experiments, we combined Uaa mutagenesis with NCL. Cnf was incorporated by Uaa mutagenesis into protein constructs in which the termini could be used in ligation to thiopeptides. Strategies for ligating to both the C- and N-termini are described. The resulting double-labeled proteins can then be used in intramolecular folding experiments. Again, combinatorial assembly can be used to make *m* x *n* double-labeled proteins from *m* N-terminal fragments and *n* C-terminal fragments. We applied our double-labeling strategy to study conformational changes in α-synuclein (αS), an amyloidogenic protein found to aggregate in Parkinson's Disease (PD).[51,53,154]

Our methods should be applicable to a wide variety of proteins: essentially any proteins that express well in *E. coli* and are amenable to refolding after ligation. Moreover, since this combination of Uaa mutagenesis and NCL to label proteins on the backbone and sidechain is unprecedented (examples of sidechain ligation using Uaa mutagenesis exist),[31,155] the considerations here should be useful to those who wish to

combine these two methods for other reasons, such as incorporating olefin or D-amino acid modifications in combination with unnatural sidechains.[125,156-158]

§ 3.2 Results and Discussion

Note: CaM peptide binding assays were performed with Colin Fadzen. Synthesis of $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$ was performed by Solongo Batjargal.

To begin combining Uaa mutagenesis with thioamides, we returned to CaM as our model system. Since Cnf is excited at 240 nm and fluoresces with a maximum at 295 nm, we began by mutating the two endogenous Tyr residues in CaM to Phe, generating a construct hereafter referred to as CaM^{F}. This "all Phe background" ensures that Tyr fluorescence does not confound our ability to interpret Cnf data. Both the extinction coefficient and quantum yield of Phe (ε_{240} = 65 $cm^{-1}•M^{-1}$, Φ = 0.022) are substantially smaller than the values for Cnf (ε_{240} = 10,333 $cm^{-1}•M^{-1}$, Φ = 0.110), so in these proteins all fluorescence emission at 295 nm should derive from Cnf.

To incorporate Cnf, we used a plasmid containing a UaaRS/tRNA pair previously selected for Cnf incorporation by Mehl and coworkers.[159] This was paired with a plasmid encoding CaM under isopropyl β-D-1-thiogalactopyranoside isopropyl (β-IPTG) control. We expressed CaM^{F} in M9 media supplemented with Cnf and purified the full-length protein by two rounds of phenyl sepharose chromatography. Four positions were chosen for incorporation of Cnf: F_{13}, F_{17}, F_{93}, and Y_{100}. Four pOCNC peptides were synthesized: an all Phe oxopeptide (pOCNC), and three thiopeptides – pOCNC-F'_{1}, pOCNC-L'_{11}, and pOCNC-F'_{16}. Before performing fluorescence titration experiments, native PAGE gel analysis was used to ensure that each CaM^{F}/pOCNC combination formed stable complexes. This allowed us to eliminate $CaM^{F}F^{*}_{17}$ from further binding experiments.

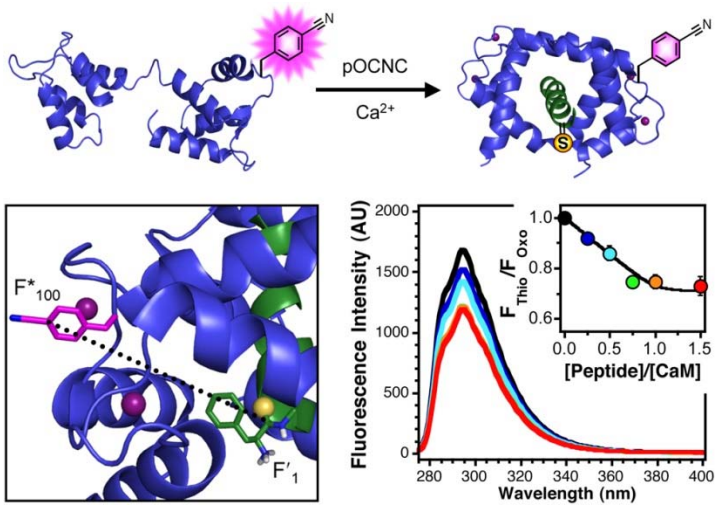

Figure 3.3 FRET in CaM/Peptide Complexes.
Top: Binding of thioamide-labeled peptides (green) to Cnf-labeled CaM (blue) in the presence of Ca^{2+}. Bottom Left: An image showing Cnf (pink) at position 100 in calmodulin ($CaM^{F}F^{*}_{100}$) and thiophenylalanine (yellow sulfur atom) at position 1 in the pOCNC peptide (pOCNC-F'$_1$). Ca^{2+} ions are shown as purple spheres. Adapted from PDB ID 1SYD using PyMol. The 18.9 Å separation shown represents an average of the separation of the center of the thioamide from the center of the α-carbons of the Cnf ring. Bottom Right: Fluorescence emission spectra of solutions of 10 μM $CaM^{F}F^{*}_{100}$ in the presence of increasing concentrations of pOCNC-F'$_1$. Inset: Relative fluorescence of complexes of $CaM^{F}F^{*}_{100}$/pOCNC-F'$_1$ (F_{Thio}) to an oxoamide control $CaM^{F}F^{*}_{100}$/pOCNC (F_{Oxo}).

For each of the other Cnf mutants, we performed the same set of experiments, which are described using $CaM^{F}F^{*}_{100}$ as an example. Fluorescence wavelength scans were acquired for mixtures of 10 μM $CaM^{F}F^{*}_{100}$ with between 0 and 15 μM pOCNC or pOCNC-F′$_1$. For each concentration, the fluorescence emission from the pOCNC-F′$_1$ complex (F_{Thio}) was normalized to the fluorescence emission from the pOCNC complex (F_{Oxo}). For F^{*}_{100} a concentration-dependent decrease was observed for F_{Thio}, while no change was seen F_{Oxo}. Plotting F_{Thio}/F_{Oxo} (or $1 - E_Q$) as a function of concentration shows that binding reaches saturation at 1:1 $CaM^{F}F^{*}_{100}$/pOCNC-F′$_1$ (Figure 3.3) The quenching (FRET) efficiency can then be used to calculate a separation according to $R_{FRET} = R_0(1/E_Q - 1)^{1/6}$. We observed 25% quenching at saturating concentrations of pOCNC-F′$_1$, which translates into a 20 Å separation using Förster theory. This is in excellent

agreement with the 19 Å average separation computed from the 20 lowest energy structures deposited from the NMR structure determination by Contessa *et al*.[133] These calculations were performed using $\Phi = 0.110$ and an orientation factor (κ^2) of 2/3.

We completed comparable experiments for each pair of Cnf mutant and thiopeptide. The separation in the seven pairs spanned a distance range from 8 to 31 Å, covering the range that is accessible with the Cnf/thioamide FRET pair. In the case of $CaM^{F}F^{*}_{100}$/pOCNC-F'_1, no change in fluorescence was observed in the "oxoamide control" experiment. For other mutants, some changes in fluorescence were observed for pOCNC. Regardless, these effects can be deconvoluted by calculating E_Q relative to the oxoamide control as we do. Such oxoamide control experiments are essential as many factors can influence fluorescence. [135,160,161] Therefore, we always compare any thioamide quenching measurement to the most relevant control, in this case the CaM/pOCNC oxopeptide complex.

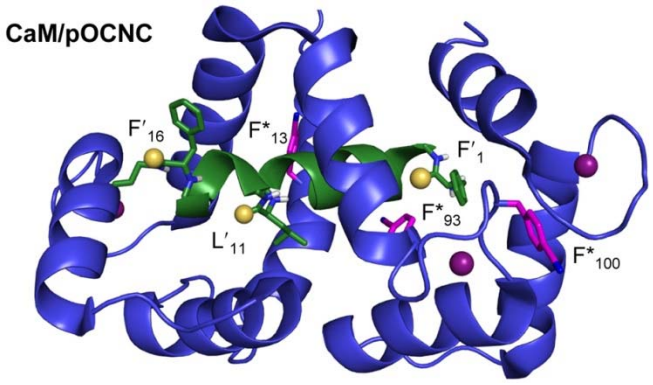

Figure 3.4 CaM/pOCNC Label Sites.
Cnf shown at thee positions in CaM with three possible pOCNC thiopeptides. Images based on one of the 20 low energy structures in PDB ID 1SYD, modified using PyMol.

The oxopeptide complex is also valuable in determining the appropriate value of Φ to be used in FRET distance calculations. The changes in fluorescence intensity upon pOCNC binding represent a change in Cnf quantum yield as it moves into a new

environment. For example, the emission of $CaM^{F}F^{*}_{93}$ starts out at 3.8% of the emission of an equimolar concentration of $CaM^{F}F^{*}_{100}$, but it increases 3.2-fold upon pOCNC binding. Therefore, the relevant quantum yield for the $CaM^{F}F^{*}_{93}$/pOCNC complex is 0.014. We use this value in our calculation of R_{FRET} for all thiopeptide complexes with $CaM^{F}F^{*}_{93}$. Similar corrections to the quantum yield were performed for $CaM^{F}F^{*}_{13}$. Correction of the quantum yield of $CaM^{F}F^{*}_{100}$ was not necessary, presumably because Cnf is solvent-exposed at this position.

The distances calculated from our FRET measurements are collected in Table 3.1. All measurements are within 3 Å of the average value observed in the NMR structure except for F^{*}_{100}/F'_{16}. Moreover, our measurements are always correct in predicting the relative order of thioamide proximity for a given Cnf mutant. While it may be surprising that we achieve reasonable agreement between theory and experiment using a value of 2/3 for κ^2, deriving κ^2 values from the 20 lowest energy NMR structures has only a small effect on this agreement (see discussion below). The R_{FRET} values obtained using $\kappa^2 = 2/3$ are more representative of the application of thioamide FRET in a structure determination experiment (i.e. one in which the target structure is unknown and κ^2 would be difficult to fit accurately).

Table 3.1 Distances in CaM/pOCNC Complexes

CaM/pOCNC Mutants (Φ)[a]	% E_Q Obsvd[b]	R_{FRET} (Å)[c]	R_{NMR} (Å)[d]
All Phe control	0	--	--
F^*_{13}/F'_1 (0.003)	11 ± 1	13	14
F^*_{13}/L'_{11} (0.003)	37 ± 2	11	14
F^*_{17}	NB[e]	--	--
F^*_{93}/F'_1 (0.014)	79 ± 2	12	9
F^*_{93}/L'_{11} (0.014)	43 ± 4	13	12
F^*_{100}/F'_1 (0.110)	25 ± 3	20	19
F^*_{100}/L'_{11} (0.110)	17 ± 2	21	24
F^*_{100}/F'_{16} (0.110)	10 ± 3[f]	23	31

[a]Φ determined by comparison to fluorescence emission of 10 μM Cnf, Φ = 0.110. [b]E_Q determined for 1:1 CaM/pOCNC complex. [c]R_{FRET} calculated from E_Q as described in text. [d]R_{NMR} is an average value calculated from the twenty lowest energy structures in PDB ID 1SYD. [e]NB indicates no binding by native PAGE gel analysis. [f]Mixture of R_{17} D- and L- epimers of pOCNC-F'_{16} peptide. Both bind with high affinity by native PAGE gel analysis.

These CaM experiments involved intermolecular quenching, but our method, like FRET in general, is equally applicable to intramolecular conformational changes. However, obtaining Cnf/thioamide double-labeled proteins represents a biosynthetic challenge, as one ideally wishes to have complete freedom in placing the chromophores within the protein sequence. While this is easily accomplished in small peptides made by SPPS, and conceivably could be done by NCL, we wished to use a combination of NCL and Uaa mutagenesis to minimize unnecessary peptide synthesis. Therefore, we assessed the compatibility of these two methods. We chose αS as a model system because we had previously shown that thioamides could be incorporated into αS, and we wished to compare PET quenching of Trp in αS to FRET quenching of Cnf at similar locations.[116]

First, we examined thioamide incorporation at the αS N-terminus. For protein expression, we used the UaaRS/tRNA plasmid pDULE2-Cnf and a plasmid encoding αS with an N-terminal truncation that could be revealed after proteolysis of a His_{10}

purification sequence using Factor Xa (Figure 3.5, **a**).[162] The proteolyzed protein contains an N-terminal Cys for use in NCL with the synthetic thioester. The Factor Xa method provides a general method for generating NCL constructs containing unnatural amino acids, but in the specific context of αS, prolonged incubation with Factor Xa can lead to protein loss *via* aggregation. Therefore, we have also used constructs directly encoding αS with an N-terminal truncation plus Met in a fashion similar to that described by Hejjaoui *et al* (Figure 3.5, **b**).[163] The N-terminal Met residue is removed *in vivo* by Met aminopeptidase to reveal the requisite Cys.[164] Protein harvested from *E. coli* contains an N-terminal thiazolidine ring formed by the condensation of Cys with either pyruvate or acetaldehyde. This is treated with methoxylamine hydrochloride to reveal free Cys for NCL reactions. Regardless of the method used in producing the expressed fragment, once purified, usage in subsequent NCL reactions is identical.

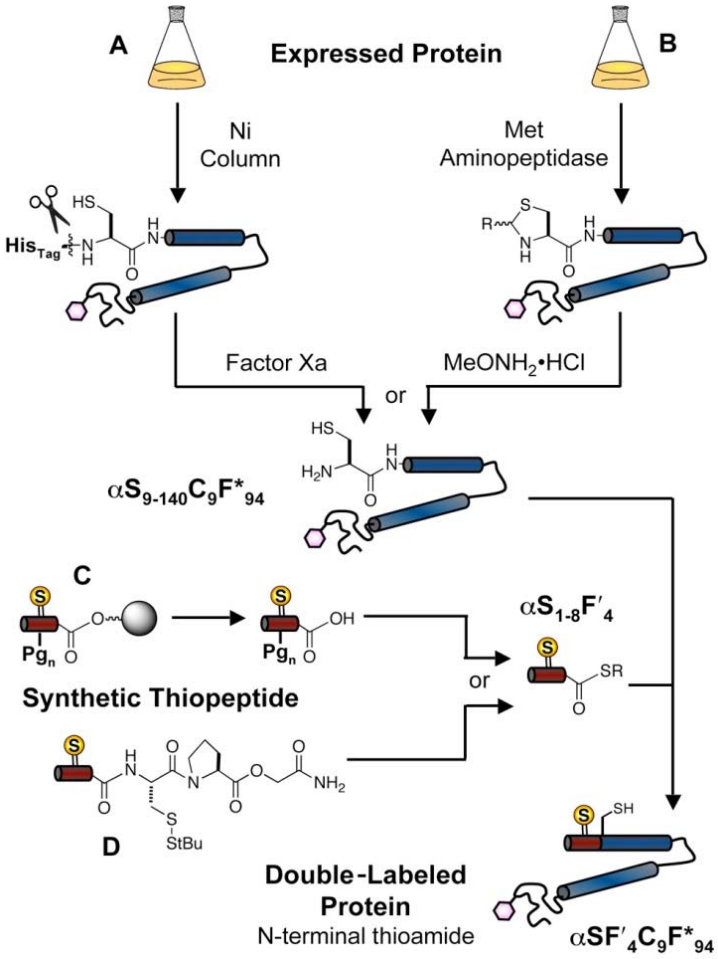

Figure 3.5 Semi-synthesis of Double-Labeled Proteins with N-terminal Thioamides. Expressed C-terminal protein fragment: A) Using a plasmid encoding an N-terminal His_{Tag} with a Factor Xa site; B) Using a plasmid encoding the truncated protein where N-terminal Met is removed by Met aminopeptidase. $MeONH_2$•HCl used to cleave the thiazolidine which forms with intracellular aldehydes. (R = Me or CH_2CO_2H) Synthetic N-terminal thiopeptide thioester: C) Synthesis on Cl-Trt resin and activation to form thioester with PyBOP, followed by deprotection with TFA; D) Synthesis on Rink amide resin with C^bPG_o masked thioester. The expressed protein fragment and synthetic thiopeptide are combined in the presence of thiophenol and the ligated, full-length protein isolated.

Synthesis of a thioamide-containing thioester peptide using Fmoc chemistry could be accomplished by one of two methods. The thiopeptide can be synthesized on Cl-Trt resin, cleaved under mild conditions (10% AcOH) to retain the sidechain protecting groups, and then activated using PyBOP to form the thioester (Figure 3.5, **c**).[165,166] This method can lead to epimerization of the α-carbon of the C-terminal amino acid, but this

can be avoided by using short (< 30 min) reaction times during PyBOP activation.[167] A more significant inherent limitation is the insolubility of protected peptides > 15 amino acids in most solvents.[165] This limits use of the PyBOP method to short peptides.

The other method makes use of a thioester that is masked during synthesis as a disulfide. Several variations of masked thioester exist (Figure 3.5, **d**). In our previous work, we used a thee residue C^bPG_o sequence where C^b is a *t*-Bu-protected Cys and G_o is glycolic acid.[116,168,169] Deprotection of the Cys thiol initiates a cascade reaction generating a diketopiperazine thioester that is competent in NCL reactions. Synthesis of peptides with the C^bPG_o linker is inefficient because of ester linkage instability. Therefore we reserve its use for longer peptides where it is necessary.

All αS protein constructs tested contain Tyr to Phe mutations at positions 39, 125, 133, and 136 to generate an "all Phe" background construct, αS^F. As was the case for CaM^F, these mutations allow us to more easily interpret Cnf fluorescence data. Combinations of methods **a** or **b** with **c** or **d** were used to synthesize double-labeled αS. While all combinations were viable, for $\alpha S^F F'_4 C_9 F^*_{94}$, our highest yields were obtained with a combination of methods 3.5b and 3.5c. Protein expression yields were typically 3 mg/L of culture after methoxylamine treatment. Ligations were typically carried out on the 0.1 μmol scale at pH 7.2 with 1% thiophenol. After 24 hours, 60 - 70% yields of ligated product were typically seen, and the products purified by C4 reverse phase HPLC. PAGE gel, spectroscopic, and MALDI MS analyses of the ligation confirm production of homogeneous, double-labeled αS. Data for other reactions are shown in Materials and Methods.

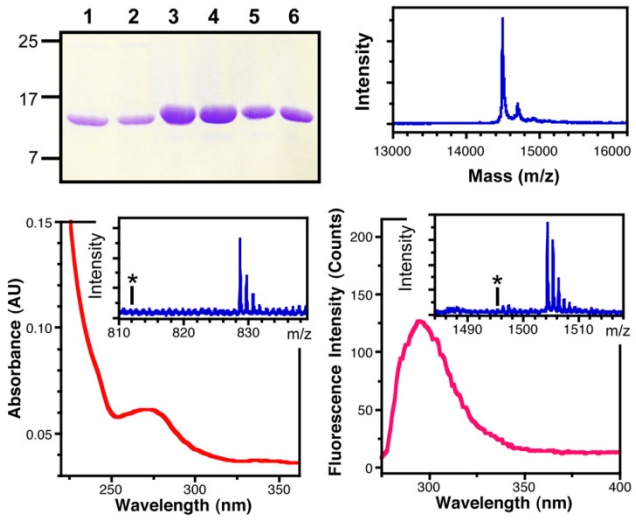

Figure 3.6 Characterization of αS^{F}F'$_{4}$C$_{9}$F*$_{94}$ Ligation Product.
Top Left: PAGE gel showing: αS$^{F}_{9\text{-}140}$C$_{9}$F*$_{94}$ before (1) and after (2) MeONH$_2$•HCl treatment, ligation of αS$^{F}_{9\text{-}140}$C$_{9}$F*$_{94}$ to αS$^{F}_{1\text{-}8}$F'$_{4}$-SR thioester (3), purified αS^{F}-F'$_{4}$C$_{9}$F*$_{94}$ (4), αS^{F}C$_{9}$F*$_{94}$ (5), and αS^{WT} (6). See HPLC chomatogram in Materials and Methods for separation of ligated product. Top Right: MALDI MS analysis of full-length αS^{F}F'$_{4}$C$_{9}$F*$_{94}$. Bottom Left: UV/Vis absorption spectrum of αS^{F}F'$_{4}$C$_{9}$F*$_{94}$ showing thioamide absorption at 266 nm. Inset: MALDI MS analysis of trypsinized 1-6 fragment, confirming the presence of the thioamide at F'$_{4}$; Calcd m/z (M+H): 828.35, Obsvd: 828.65. The asterisk indicates the absence of peaks corresponding an oxoamide at Phe$_{4}$ in the 1-6 fragment; Calcd m/z (M+H): 812.38. Bottom Right: Fluorescence emission spectra of αS^{F}F'$_{4}$C$_{9}$F*$_{94}$ showing Cnf emission at 295 nm. Inset: MALDI MS analysis of trypsinized 81-96 fragment, confirming the presence of the Cnf at position 94; Calcd m/z (M+H): 1503.78, Obsvd: 1504.30. The asterisk indicates the absence of peaks corresponding Tyr at position 94 in the 81-96 fragment; Calcd m/z (M+H): 1494.78.

The methods shown in Figure 3.5 permit protein double-labeling with thioamides located at the N-terminus. To insert thioamides at the C-terminus, we made use of intein fusion constructs popularized by Muir and others.[153] These self-splicing domains have been engineered to stall at a thioester intermediate and can be intercepted by treatment with a small molecule thiol to give a thioester suitable for NCL. We have found that the expression of intein fusion constructs is compatible with Uaa mutagenesis. In fact, yields of intein fusions of αS Cnf mutants (e.g. αS$^{F}_{1\text{-}122}$F*$_{94}$-Int) are higher (20 mg/L culture) than comparable expressions of the full-length protein using either T7 or PRK vectors previously used to express αS Cnf mutants (3 - 5 mg/L). The synthetic thiopeptide with N-terminal Cys (αS$^{F}_{123\text{-}140}$C$_{123}$A′$_{124}$) was prepared by standard SPPS.

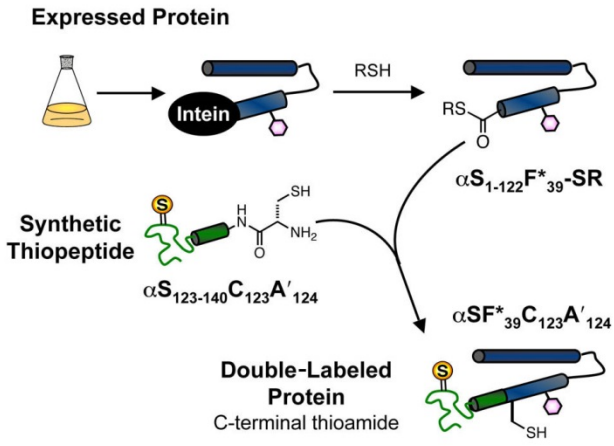

Figure 3.7 Semi-synthesis of Double-Labeled Proteins with C-terminal Thioamides.
An expressed N-terminal protein fragment is generated using a plasmid encoding a truncated version of the protein with a C-terminal intein. Thiolysis of the stalled intein generates a thioester suitable for NCL. The synthetic thiopeptide with an N-terminal Cys is made by standard SPPS. The expressed protein fragment and synthetic thiopeptide are combined in the presence of thiophenol and ligated, full-length protein isolated.

The thioester $\alpha S^{F}_{1\text{-}122}F^{*}_{39}$-SR was prepared from $\alpha S^{F}_{1\text{-}122}F^{*}_{39}$-Int using the thiol mercaptoethane sulfonate (MESNA). This was combined with $\alpha S^{F}_{123\text{-}140}C_{123}A'_{124}$ in a buffer containing thiophenol. Although some C-terminal Asn cyclization (~ 25% based on crude MALDI MS estimates, see Supporting Information) of the intein thioester occurred in competition with the desired MESNA thiolysis, this was not a major problem since the resulting cyclic succinimide was not competent for ligation. Ligations were typically carried out on the 0.3 μmol scale. After 16 hours, 50% yields were obtained, and the products purified by FPLC using an ion-exchange column. Once again PAGE gels, spectroscopy, and MALDI MS were used to analyze the ligation reaction. These data are shown in Figure 3.8. The data confirm that we have again produced homogeneous, full-length, double-labeled αS.

Using these two methods, we should be able to label a variety of proteins with thioamides at the N- or C-terminus, with complete freedom in placement of the donor fluorophore. Since we have previously shown that thioamides are compatible with the

$MeONH_2$•HCl conditions used to deprotect N-terminal thiazolidines, the methods used here should also be compatible with internal thioamide-labeling through three component ligations. In addition, we note that the use of C-terminal inteins may be generally useful to those employing Uaa mutagenesis. One major limitation of the method is that the inherent competition with release factors at the UAG stop codon can lead to protein truncation.[170] Purifying the full-length protein containing the Uaa can be difficult for sites near the C-terminus, where the truncated form is only slightly shorter. The His-tagged intein provides a convenient handle that will only be present in the untruncated protein. Since the intein can be cleaved by thiolysis or hydrolysis in a traceless manner, the Uaa-labeled fragment or protein can be obtained in pure form in an efficient manner.

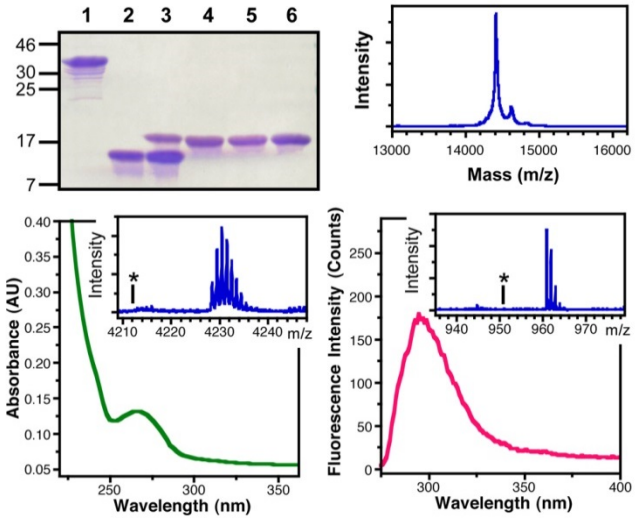

Figure 3.8 Characterization of $\alpha S^F F^*_{39} C_{123} A'_{124}$ Ligation Product.
Top Left: PAGE gel showing: $\alpha S^F{}_{1\text{-}122} F^*_{39}$-Int before (1) and after thiolysis of the intein to generate $\alpha S^F{}_{1\text{-}122} F^*_{39}$-SR thioester (2), ligation of $\alpha S^F{}_{123\text{-}140} C_{123} A'_{124}$ to $\alpha S^F{}_{1\text{-}122} F^*_{39}$-SR thioester (3), purified $\alpha S^F F^*_{39} C_{123} A'_{124}$ (4), $\alpha S^F F^*_{39} C_{123}$ (5), and αS^F (6). Top Right: MALDI MS analysis of full-length $\alpha S^F F^*_{39} C_{123} A'_{124}$. Bottom Left: UV/Vis absorption spectrum of $\alpha S^F F^*_{39} C_{123} A'_{124}$ showing thioamide absorption at 266 nm. Inset: MALDI MS analysis of trypsinized 103-140 fragment, confirming the presence of the thioamide at A'_{124}; Calcd m/z (M+H): 4228.69, Obsvd: 4228.47. The asterisk indicates the absence of peaks corresponding an oxoamide at Ala_{124} in the 103-140 fragment; Calcd m/z (M+H): 4212.72. Bottom Right: Fluorescence emission spectrum of $\alpha S^F F^*_{39} C_{123} A'_{124}$ showing Cnf emission at 295 nm. Inset: MALDI MS analysis of trypsinized 35-43 fragment, confirming the presence of Cnf at position 39; Calcd m/z (M+H): 960.52, Obsvd: 960.77. The asterisk indicates the absence of peaks corresponding Tyr at position 39 in the 35-43 fragment; Calcd m/z (M+H): 951.51.

In order to demonstrate the value of double-labeling a protein such as αS, we carried out experiments in which intramolecular conformational changes were monitored. αS aggregates by first forming soluble oligomers which eventually convert into larger, insoluble fibrils.[53,60,85,154] Both the oligomers and fibrils have been shown to be cytotoxic, and recent evidence has shown that extracellular αS can enter neurons and catalyze misfolding of αS within healthy neurons.[171-174] Therefore, there is substantial current interest in understanding the structural dynamics of the meta-stable αS monomers and the rearrangements that occur on formation of oligomers and subsequent fibrillization. Several researchers have recently shown the value of being able to construct semi-synthetic αS to study the effects of post-translational modification on its folding, our modification should provide a uniquely non-perturbing method for tracking its dynamics.[91,116,163,175,176]

We envision thioamide quenching as a method for site-selectively monitoring conformational changes during αS refolding. While a number of researchers have studied αS conformation using other techniques such as NMR, IR, and EPR, we find fluorescence to be particularly useful because the sensitivity of the technique permits one to use low concentrations of double-labeled αS which should not undergo aggregation on the timescale of our experiments.[60,85,174,177-181] This strategy allows us to monitor long-range (25 - 30 Å) interactions that are difficult to observe with other techniques. Moreover, FRET is easily applied to interactions involving residues in dynamic regions of proteins, such as the termini of αS.

We have previously shown that conformational dynamics during aggregation can be monitored by intramolecular quenching of Trp by a thioamide.[116] Here, we

demonstrated that we can monitor conformational changes in monomeric αS, using urea or trimethylamine oxide (TMAO) to denature or compact αS, respectively. Performing oxoamide control experiments allowed us to correct for any changes in fluorescence not due to the thioamide, and correct the Cnf quantum yield in order to properly interpret the observed quenching in terms of interchromophore distance.

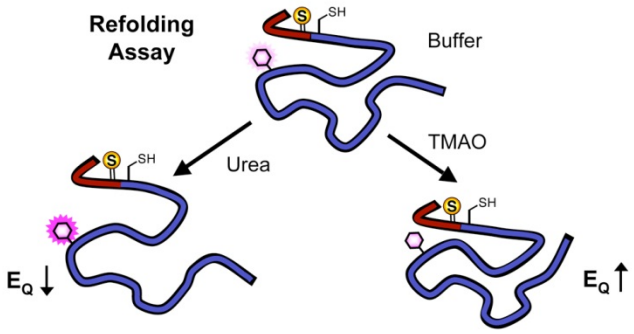

Figure 3.9 αS Refolding Assay.
Monomeric double-labeled αS (e.g. $\alpha S^{F}F'_{4}C_{9}F^{*}_{94}$) is diluted to 1 μM in either Tris buffer with 1 mM BME, or the same buffer with urea or TMAO added. Cnf fluorescence is measured at 295 nm (F_{Thio}). Identical experiments are carried out with donor-only labeled αS (e.g. $\alpha S^{F}C_{9}F^{*}_{94}$) to determine F_{Oxo}. $E_Q = 1 - F_{Thio}/F_{Oxo}$.

The NCL procedures that we use to synthesize our thioamide-labeled proteins introduce a Cys into the αS sequence at position 9 or 123. However, we have previously shown that the presence of non-native Cys in the sequence does not cause aberrant aggregation when incubated with β-mercaptoethanol (BME) to prevent disulfide formation.[116] While Cys can have an effect on Cnf emission, since we include Cys mutations in our oxoamide control experiments, we can account for these changes as well.[182,183]

We carried out refolding experiments with the constructs described above, $\alpha S^{F}F'_{4}C_{9}F^{*}_{94}$ and $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$, as well as $\alpha S^{F}F'_{4}C_{9}F^{*}_{39}$. In the "native" state, when our constructs were incubated in buffer alone, we observed varying levels of quenching relative to the equivalent oxoprotein control. (Figure 3.10) For F^{*}_{94}/F'_{4} and F^{*}_{94}/A'_{124} we

observe no quenching. For F^*_{39}/F'_4, we observe some quenching, which we interpret as resulting from local compaction of the N-terminus. All of these observations are consistent with previous FRET experiments performed on αS. The observed N-terminal compaction is similar to the roughly 30 Å separation observed by Rhoades and coworkers using Alexa Fluor 488 and Alexa Fluor 594 dyes in single molecule FRET experiments.[70] Ferreon *et al*. observed FRET efficiencies consistent with separations of ~ 70 Å in single molecule FRET experiments with αS labeled at positions 7 and 84 with the same dye pair.[184,185] This is well outside the useful range of the Cnf/thioamide FRET pair, so it is not surprising that we see no quenching for F^*_{94}/F'_4. Our results for F^*_{39}/F'_4 are consistent with experiments carried out by Lee, Gray, and coworkers using Trp and nitrotyrosine as a FRET pair.[68,186] Our F^*_{39}/A'_{124} data are consistent with reports by Nath *et al*. showing a ~ 45 Å interaction between probes at positions 33 and 130, which is beyond the range of the Cnf/thioamide pair.[70]

When 6 M urea was used as an additive, an increase in F_{Thio}/F_{Oxo} was observed for F^*_{39}/F'_4, consistent with greater unfolding of the protein. For the other two mutants, no change in F_{Thio}/F_{Oxo} was observed, indicating that these regions were indeed less compacted. Control experiments using Cnf itself indicate that although an inherent change in Cnf quantum yield plays a role in the observed change in fluorescence, there is nonetheless an observable expansion of αS once one corrects for inherent effects on Cnf fluorescence. In contrast, an increase in quenching (i.e. FRET) was observed as a function of TMAO concentration for all three mutants. The changes in E_Q as a function of TMAO concentration can be seen in Figure 3.10. This indicates increased compaction of both N- and C- termini, as well as folding of the N-terminal region on itself.

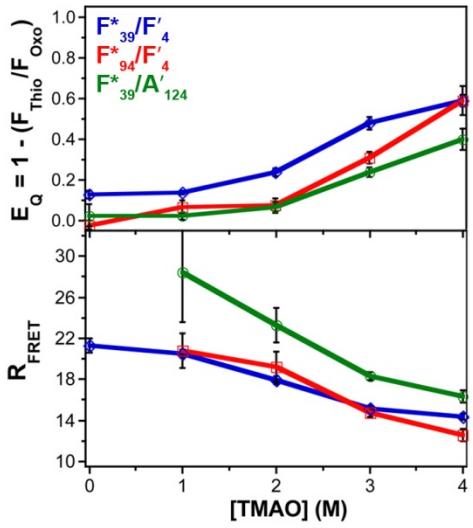

Figure 3.10 Refolding of αS in TMAO.
Top: Quenching efficiency of $\alpha S^{F}F'_{4}C_{9}F^{*}_{94}$ (red), $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$ (green), or $\alpha S^{F}F'_{4}C_{9}F^{*}_{39}$ (blue) determined in phosphate buffer at 25 °C at varying concentrations of TMAO. Quenching efficiency determined by comparison to the equivalent oxoprotein: $\alpha S^{F}C_{9}F^{*}_{94}$, $\alpha S^{F}F^{*}_{39}C_{123}$, or $\alpha S^{F}C_{9}F^{*}_{39}$, respectively. Bottom: Interchromophore distances computed using Förster theory as described in Supporting Information.

To assign distance constraints based on these measurements, we converted the measured E_Q values to distances using Förster theory. As with CaM, we computed a specific R_0 value for each construct using the quantum yields of Cnf for the oxoamide control proteins $\alpha S^{F}C_{9}F^{*}_{39}$ (R_0 = 15.6 Å in 0 M TMAO), $\alpha S^{F}C_{9}F^{*}_{94}$ (R_0 = 13.6 in Å in 0 M TMAO), or $\alpha S^{F}F^{*}_{39}C_{123}$ (R_0 = 13.5 Å in 0 M TMAO). Förster radii for each mutant in other TMAO concentrations are given in Supporting Information. We use these rather than simply using R_0 values calculated from free Cnf because the variation among the different mutants demonstrates that the protein context influences quantum yield. The R_{FRET} values that result from our calculations indicate that Cnf and the thioamide come into very close proximity in TMAO. While these results are consistent with the single molecule studies of Deniz and coworkers, it is surprising that our R_{FRET} values are so small.[185] This may be the result of a dominant contribution from very close contact. Ferreon *et al*. observe subpopulations with near unity FRET efficiencies, but the Alexa

Fluor FRET pair that they use has high FRET efficiency for any separation less than 30 Å.[184,187] Our probes have a shorter effective range (8 – 31 Å) than theirs, and therefore provide a more precise report on the close contact induced by TMAO. We will continue to study this phenomenon, not only with thioamides, but also with other short range FRET pairs such as Cnf and Trp.[48,188] Finally, it may be more appropriate to use polymer models rather than simple Förster theory to interpret our FRET data. These models have clear value when interpreting single molecule FRET data, which may have broad and asymmetric FRET efficiency distributions. Although the value is less clear for ensemble measurements, we will explore the interpretation of our FRET data using models such as those used by Lee, Gray, Rhoades, and Plaxco.[68,70,71,186,189]

The application of our recently developed Cnf/thioamide FRET probe pair to studies of the folding of full-sized proteins requires the ability to synthesize those labeled proteins in an efficient manner. Although we could, in theory, construct Cnf- and thioamide-labeled proteins purely through ligation, we have developed a far more expedient approach by melding the methods of unnatural amino acid mutagenesis and protein ligation. Intermolecular studies simply require that Cnf be incorporated into proteins expressed in *E. coli* through established unnatural amino acid mutagenesis techniques. Intramolecular studies require more complex semi-synthesis to obtain doubly labeled proteins, wherein a Cnf-labeled protein fragment can be expressed in *E. coli* and then ligated to a synthetic thiopeptide. This combination of methods minimizes SPPS of natural amino acid segments. The same methods can be used to synthesize proteins labeled with other donor fluorophores such as coumarin, or any fluorophore conjugated to a bioorthogonal handle, such as *p*-azidophenylalanine.[190,191] In fact, we have identified

a synthetase for the incorporation of the fluorescent amino acid acridonylalanine (Acd), which we have previously shown to be quenched by thioamides.[113,192] We are currently working to optimize yields of Acd/thioamide labeled proteins using methods analogous to those described here.

The doubly labeled proteins produced using these methods can be applied to monitor protein folding and dynamics. For example, here we have studied the structure of a CaM/peptide complex and determined the average separation of residues in the disordered protein αS in various buffer conditions. The benchmarking results using CaM help to validate our method and provide a foundation for the application of thioamide FRET to other systems. These experiments show clearly that meaningful interpretations of the interchromopore distances require determining changes in the donor quantum yield. This is easily accomplished using the oxoamide control experiments that we always include in order to determine quenching efficiency. Moreover, the dependence of quenching efficiency on Cnf quantum yield supports a FRET mechanism for quenching as opposed to Dexter or PET mechanisms, which were difficult to discount based on previous data.[10,115] The confirmation of a FRET mechanism for Cnf/thioamide interactions allows one to confidently interpret thioamide quenching using Förster's equations for future applications. Nonetheless, it is worth noting that there was an imperfect match of observed distances (R_{FRET}) and predicted distances (R_{NMR}) that was not resolved by considering the orientation parameter κ^2. This may be because simply using the 20 structures developed in solving the NMR structure does not fully account for intermediate geometries between these valleys on the potential energy surface. Calculating R_{FRET} by sampling κ^2 over extended molecular dynamics trajectories may

give improved agreement with the FRET data. Additionally, fluorescence lifetime measurements (which are concentration-independent) would remove the influence of small deviations in concentration on our determination of E_Q. We will explore both of these aspects in subsequent work.

The αS experiments show how one might use our method in an intramolecular sense. The addition of TMAO results in compaction of monomeric αS, consistent with previous independent observations by Deniz as well as small angle X-ray scattering experiments by Fink.[80,185] The short effective range of our probes allows us to obtain more precise information about close contact of regions of the protein than similar experiments with larger fluorophores. Moreover, the small size of our chomophores leaves less uncertainty in their location relative to the protein structure. Thus, the thioamide probes provide new short-range information on TMAO-induced conformations of αS, complementing data obtained in single molecule experiments. We continue to apply fluorophore/thioamide pairs to the study of folding transitions in αS monomers as well as to monitoring conformational changes during aggregation.

Of course, our labeling strategy is not restricted to CaM or αS. Any protein which can be expressed at high levels in *E. coli* and reversibly unfolded should be amenable to double-labeling using a combination of Uaa mutagenesis and NCL. This strategy should be particularly appealing when one of the modifications cannot be incorporated ribosomally (i.e. backbone modifications such as olefin isosteres). For larger dyes that are attached post-translationally, Uaa mutagenesis at two sites, or used in combination with Cys modification is probably more efficient.[28,32,38,193] However, even in those cases, it can be difficult to separate labeled from unlabeled protein, whereas the ligated products of

our reactions differ substantially in size and can easily be separated by HPLC or FPLC. The major current limitations of our labeling methods are that Uaa mutagenesis expression yields can be low (3 - 5 mg/L), and if these are coupled to inconsistent NCL yields, it can be difficult to obtain sufficient quantities of protein for biophysical experiments. We will continue methodological studies to improve both aspects to make these methods as accessible as possible to the general biochemical community.

§ 3.3 Efficient, Traceless Semi-Synthesis of α-Synuclein Labeled with a Fluorophore/Thioamide FRET Pair

Despite the agreement of our Cnf/thioamide data with other studies of αS in TMAO, we were somewhat concerned about the presence of the non-native Cys residues in our labeled αS constructs.[185] Cys mutants of αS display enhanced aggregation kinetics and altered fibril morphology as a result of intermolecular disulfide bond formation.[194,195] Although we performed the previous FRET studies with our Cys-containing double-labeled constructs in the presence of a reducing agent, it is nonetheless possible that some amount of disulfide formation occurred during our experiments.

While many laboratories carry out protein synthesis with subsequent radical desulfurization of Cys to form Ala, this strategy is not viable for us as desulfurization of the thioamide would also occur.[196] Methods developed in our laboratory enable us to generate double-labeled αS for misfolding studies that does not contain cysteine at the ligation site. It has been shown that NCL can be performed with an N-terminal homocysteine (Hcs) in place of Cys.[197-200] Hcs can then be selectively methylated to yield methionine at the ligation site. Hcs-mediated ligation was previously restricted to short C-terminal peptides in which Hcs could be installed using SPPS. We have recently established that we can use *E. coli* aminoacyl transferase (AaT) to deliver disulfide-

protected Hcs (*S*-(thiomethyl)homocysteine, Hcm) to the N-terminus of a large expressed protein fragment.[201] In *E. coli* cells, AaT transfers Phe, Leu, or Met from an aminoacyl tRNA to the N-terminus of a protein.[202,203] By using a modified methionine aminoacyl tRNA synthetase enzyme (Met*RS), we can generate Hcm-tRNA *in situ*, which serves as a viable substrate for AaT.[204] Once Hcm is transferred to the N-terminus of a protein, it is easily deprotected by a reducing agent to form Hcs for ligation. By combining unnatural amino acid mutagenesis with EPL at homocysteine (which is later converted to methionine), we can incorporate our minimal FRET pair into αS in a traceless manner that minimizes unnecessary synthesis of peptide fragments composed of natural amino acids. A retrosynthetic analysis of our target protein, Ac-αS^{F}Asp'$_2$Cnf$_{39}$, is shown in Figure 3.11.

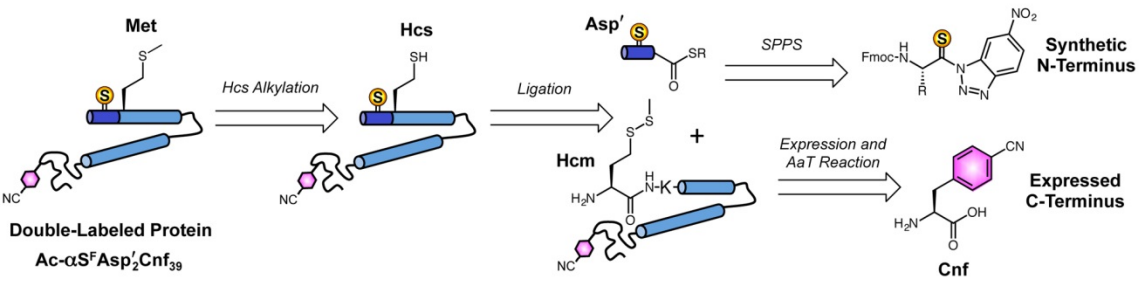

Figure 3.11 Retrosynthetic Analysis of Ac-αS^{F}Asp'$_2$Cnf$_{39}$.

§ 3.4 Results and Discussion

Previously, we have shown that fluorenylmethoxycarbonyl (Fmoc)-protected thiocarboxybenzotriazoles can be used with conventional Fmoc-protected amino acids to synthesize thioamide-containing peptides.[205,206] The procedure for the preparation of Fmoc-Asp′(O*t*-Bu)-nitrobenzotriazole is illustrated in Scheme 3.1.[114] Coupling of 4-nitro-1,2-phenylenediamine to Fmoc-Asp(O*t*-Bu)-OH was achieved using isobutyl chloroformate to generate the coupled product in 89% yield. The amide was then added

to a solution of P_4S_{10} and anhydrous Na_2CO_3 at 0 ºC. Selective thionation of the Asp backbone carbonyl was completed within 1 h to afford the thioamide in 79% yield. Intramolecular diazonium cyclization of the thioamide gave the activated benzotriazole in 67% yield. The benzotriazole can be isolated from the final reaction by precipitation with ice-cold water in sufficient purity for peptide coupling reactions. Since the activated benzotriazoles are subject to degradation though hydrolysis and intramolecular nucleophilic attack by the carbonyl oxygens, it is best to minimize handling of compound the benzotriazole prior to peptide coupling.

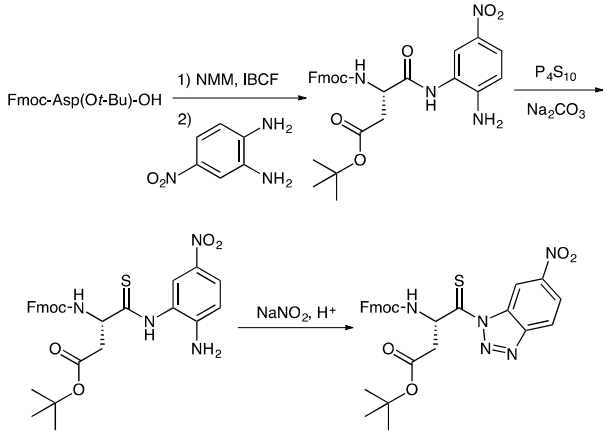

Figure 3.12 Synthesis of Aspartyl Thioamide Precursor

The N-terminal thioamide-containing fragment of αS was synthesized using standard Fmoc-based SPPS procedures with the exception of direct acylation of Val_3 by the pre-activated thioaspartate precursor compound (i.e., no activating agents were added). The thiopeptide was synthesized on 2-chlorotrityl resin, cleaved under mild conditions (10% AcOH) to retain the Asp side-chain protecting group, and then activated using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) to form the mercaptopropionate thioester (Figure 3.13). Epimerization of the α-carbon of the C-terminal phenylalanine residue was eliminated by reducing the PyBOP activation time

to 25 minutes. Following acidic deprotection of Asp′(O*t*-Bu)$_2$, the thioester peptide was isolated by HPLC (2.5 μmol, 2.5% isolated yield) and its identity was confirmed by MALDI MS analysis.

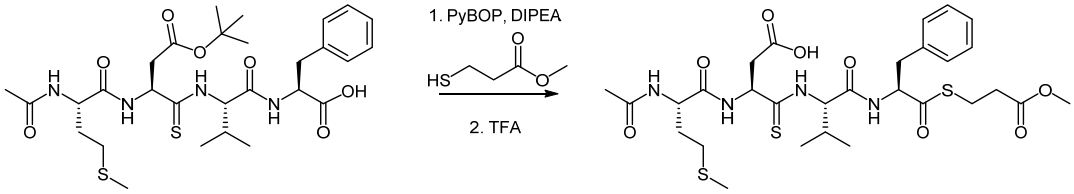

Figure 3.13 Synthesis of Thioamide-Containing Thioester Peptide for Native Chemical Ligation

The expressed C-terminal Cnf-containing αS fragment for NCL was generated as depicted in Figure 3.14. AaT selectively modifies the α-amine of proteins bearing a lysine or arginine as the N-terminal residue. Prior to this work, we demonstrated that Met$_5$Lys$_6$ could be used as a point of disconnection in the synthesis of αS. To minimize perturbation of native αS dynamics resulting from Cnf incorporation, we chose to replace the similarly sized Tyr or Phe at position 39. A plasmid containing the C-terminal protein fragment ($\alpha S^F_{6\text{-}140}$) with a TAG codon installed at position 39 was prepared. An N-terminal His$_{11}$ tag and a Factor Xa proteolysis site precede the αS gene. *E. coli* cells were transformed with two plasmids, one containing the Cnf-selective mutant synthetase (CnfRS) with tRNA and the other containing His_{Tag}-$\alpha S^F_{6\text{-}140}TAG_{39}$. Following an initial growth period, protein expression was induced by adding Cnf and isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture medium. The expressed protein, His_{Tag}-$\alpha S^F_{6\text{-}140}Cnf_{39}$, was purified by Ni-affinity chromatography and subsequently cleaved by Factor Xa to yield $\alpha S^F_{6\text{-}140}Cnf_{39}$. Following purification by ion-exchange chromatography, the truncated protein was incubated with AaT, Met*RS, *E. coli* total tRNA, Hcm, and ATP for 2 h resulting in complete formation of $\alpha S^F_{5\text{-}140}Hcm_5Cnf_{39}$. $\alpha S^F_{5\text{-}140}Hcm_5Cnf_{39}$ was

subjected to an additional round of ion-exchange chromatography prior to use in the NCL reaction. Typically, about 1 μmol of pure $\alpha S^{F}_{5\text{-}140}Hcm_5Cnf_{39}$ is obtained from a 1 L protein expression.

The ligation reaction was initiated by incubation of $\alpha S^{F}_{5\text{-}140}Hcm_5Cnf_{39}$ with 1.5 equivalents of Ac-MetAsp′ValPhe-SR in degassed ligation buffer (6 M guanidinium hydrochloride, 0.2 M sodium phosphate, 20 mM tris(2-carboxyethyl)phosphine, 2% thiophenol, at pH 7.5). Disulfide deprotection of the Hcm residue was observed by MALDI MS within five min. The ligation reaction was allowed to proceed for 24 h prior to buffer exchange into methylation reaction buffer (20 mM Tris, pH 8.6). Ac-$\alpha S^{F}Asp'_2Hcs_5Cnf_{39}$ was converted to Ac-$\alpha S^{F}Asp'_2Cnf_{39}$, with native Met_5, by a short treatment (10 min) with 3000 equivalents of MeI. Allowing the reaction to proceed for longer periods of time led to *S*-alkylation of the thioamide (data not shown). Ac-$\alpha S^{F}Asp'_2Cnf_{39}$ was isolated from the NCL reaction by HPLC and characterized by MALDI MS, as well as UV-Vis and fluorescence spectroscopy (Figure 3.14). The yield over the two-step ligation and methylation sequence was 41%. MALDI MS analysis of trypsin-digested Ac-$\alpha S^{F}Asp'_2Cnf_{39}$ confirmed selective alkylation of Hcs to form Met in the purified product.

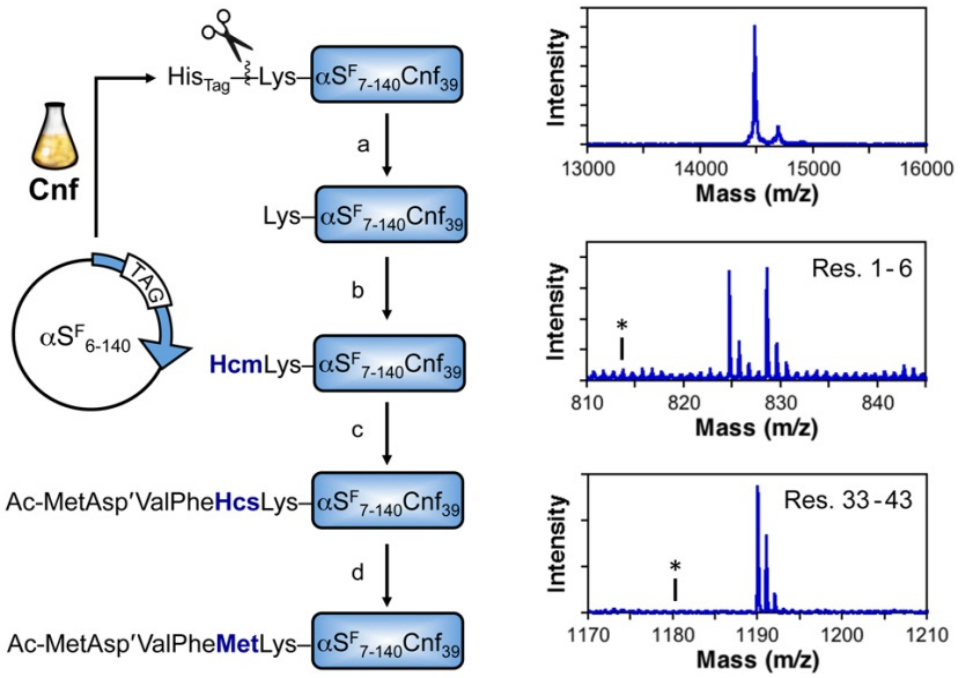

Figure 3.14 Ac-αS^{F}Asp′$_2$Cnf$_{39}$ Synthesis.
Left: Functionalization of αS$^{F}_{6\text{-}140}$Cnf$_{39}$ by a) cleavage of His$_{11}$ tag at Factor Xa site, b) attachment of Hcm by AaT-catalyzed modification, c) ligation of an N-terminal thioester peptide (3.5), and d) conversion of Hcs to Met by methylation to form Ac-αS^{F}Asp′$_2$Cnf$_{39}$. Top Right: MALDI MS analysis of full length -αS^{F}Asp′$_2$Cnf$_{39}$. Calcd m/z (M+H)$^+$ 14480.1, found 14479.6. Middle Right: Trypsinized fragment corresponding to residues 1-6 of Ac-αS^{F}Asp′$_2$Cnf$_{39}$ confirms methylation. Calcd m/z (M+H)$^+$ 828.34, found 828.46. The asterisk indicates the expected mass of the 1-6 fragment with unmethylated Hcs at position 5. Calcd m/z (M+H)$^+$ 814.39. Bottom Right: Trypsinized fragment corresponding to residues 33-43 of Ac-αS^{F}Asp′$_2$Cnf$_{39}$ confirms Cnf incorporation. Calcd m/z (M+H)$^+$ 1189.66, found 1189.85. The asterisk indicates the expected mass of the 33-43 fragment with Tyr at position 39. Calcd m/z (M+H)$^+$ 1180.65.

Using a similar double-labeled construct containing cysteine at the ligation site (Ac-αS^{F}Phe′$_4$Cys$_9$Cnf$_{39}$), we observed thioamide quenching of Cnf after compaction of the protein in high concentrations of TMAO.[206] Since the Ser-to-Cys mutation in this protein constituted a small deviation from the native αS protein sequence, we sought to investigate this phenomenon using our newly synthesized Ac-αS^{F}Asp′$_2$Cnf$_{39}$ construct. To correct for any changes in Cnf fluorescence that are not due to the presence of the thioamide, an oxoamide control protein (αS^{F}Cnf$_{39}$) was also prepared for fluorescence experiments. In buffer without TMAO, thioamide quenching of Cnf in Ac-αS^{F}Asp′$_2$Cnf$_{39}$

(i.e., FRET) was minimal. In accordance with our previous observations, thioamide quenching of Cnf increased as a function of TMAO concentration (Figure 3.15, left axis). To assign distance constraints based on these measurements, the measured quenching efficiency (E_Q) values were converted to distances using Förster theory. In 2 M and 4 M TMAO, R_{FRET} was calculated as 16.4 Å and 13.7 Å, respectively (Figure 3.15, right axis). Taken together with our previous results, these data indicate that the N-terminal region of αS undergoes significant compaction in high TMAO concentrations. Furthermore, our new results conclusively demonstrate that the observed quenching is not due to Cys-mediated dimerization of αS.

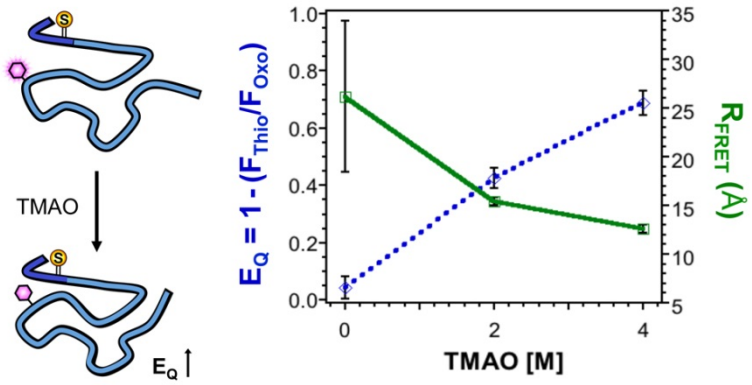

Figure 3.15 Ac-αS^{F}Asp′$_2$Cnf$_{39}$ Refolding Assay.
Left: Monomeric double-labeled αS (Ac-αS^{F}Asp′$_2$Cnf$_{39}$) is mixed with TMAO to induce compaction and an increase in quenching efficiency (E_Q). Right: E_Q (blue dashed line) of Ac-αS^{F}Asp′$_2$Cnf$_{39}$ determined at varying concentrations of TMAO. Interchomophore distance (R_{FRET}, green solid line) computed using Förster theory as described below.

By combining unnatural amino acid mutagenesis with EPL at Met, we installed our minimal Cnf/thioamide FRET probe pair into αS without introducing a mutation at the ligation site. We demonstrated that selective alkylation of Hcs can be performed in the presence of a thioamide, rendering our new ligation strategy compatible with the minimal probe pairs that we have developed. The αS refolding studies presented here demonstrate how this methodology can be used to study protein folding dynamics,

though these techniques are by no means limited to αS. Over 20,000 proteins in the PDB contain MetArg or MetLys motifs that can be used as points of disconnection in their retrosynthetic analyses.[201] We are currently working to expand these methods even further by combining EPL at Met with multiple ligation strategies. This methodology will allow us to install thioamide-containing synthetic fragments in the middle of a large protein target in a traceless and efficient manner.

§ 3.5 Combining Uaa Mutagenesis with Expressed Protein Ligation to Incorporate Acd/Thioamide Pairs into Full-Length Proteins.

Recently, our laboratory has developed methods for the efficient synthesis and *in vivo* incorporation of acridon-2-yl alanine (Acd), a blue-wavelength fluorescent amino acid.[47] Acd is a useful fluorophore because of its small size (222 $Å^3$), near unity quantum yield in water (F = 0.95), unusually long lifetime (t ~ 15 ns) and high photostability (< 5% degradation after 3 hs irradiation).[113,207,208] To incorporate Acd ribosomally, we screened a library of synthetase mutants for the ability to incorporate Acd in response to a TAG codon (Figure 3.16). An *Mj*TyrRS previously selected for 4-(2′-bromoisobutyramido)phenylalanine (Brb) was identified as capable of utilizing Acd for protein incorporation.[209] With a synthetase mutant identified, we were able to show that Acd can be incorporated into several other proteins by expressing them in *E. coli*, including α-synuclein, triose phosphate isomerase, and calmodulin.

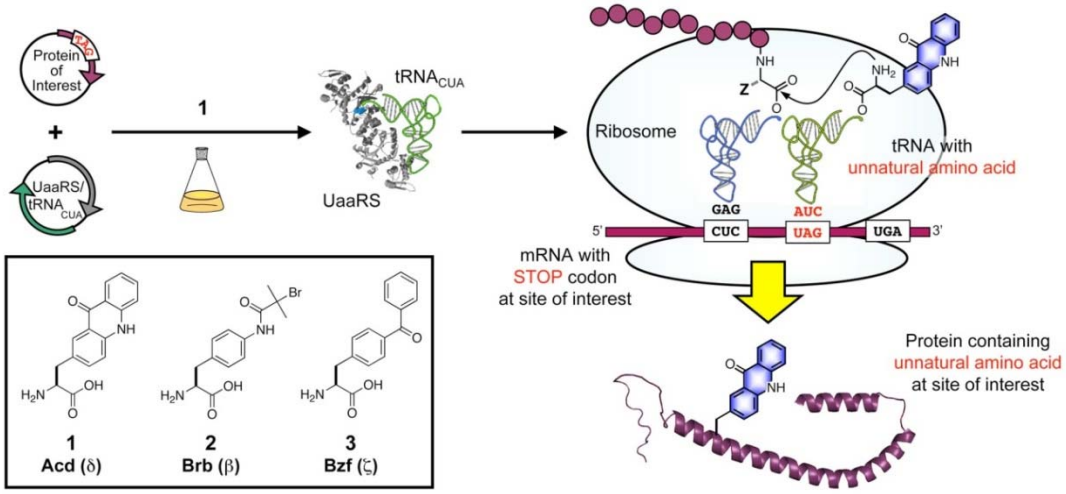

Figure 3.16 Ribosomal Incorporation of Acd.
*E. co*li cells are transformed with two DNA plasmids: one encoding an mRNA for the protein of interest with a UAG stop codon at the site of interest, and one encoding an aminoacyl tRNA synthetase that is selective for an unnatural amino acid (UaaRS) with a cognate tRNA that recognizes the UAG stop codon ($tRNA_{CUA}$). The UaaRS charges $tRNA_{CUA}$ with the unnatural amino acid and the aminoacyl $tRNA_{CUA}$ delivers the Uaa to the site of interest in the protein during ribosomal biosynthesis. Inset: UaaRSs were screened for successful incorporation of acridon-2-ylalanine (1, Acd, δ) from a pool previously selected for incorporation of 4-(2′-bromoisobutyrlamido)phenylalanine (2, Brb, ζ) and p-benzoylphenylalanine (3, Bzf, β).

Previously, we have shown that Acd can be efficiently quenched by a thioamide though a PET mechanism.[113] This makes the Acd/thioamide pair an extremely small, non-perturbing set of labels since Acd is only 47% larger than Trp and the thiomide is essentially isosteric with an amide. As an initial proof of principle, we have shown that thioamide/Acd PET quenching can be used to monitor the thermally-induced unfolding of a small well-folded polypeptide known as HP35.

Prior to this work, we combined unnatural amino acid mutagenesis with expressed protein ligation at Hcs to install a FRET pair, Cnf and a thioamide, into full-length αS (described in more detail in Section 3.5).[117] Thioamide FRET partners, such as Cnf, provide reliable distance information. However, Cnf cannot be selectively excited in the presence of Tyr. Therefore, our studies were limited to mutants of αS in which each native Tyr residue was replaced by Phe. Here, we have shown that we can combine Acd

incorporation with expressed protein ligation at Hcs to incorporate the thioamide/Acd pair into full-length αS in a traceless manner.

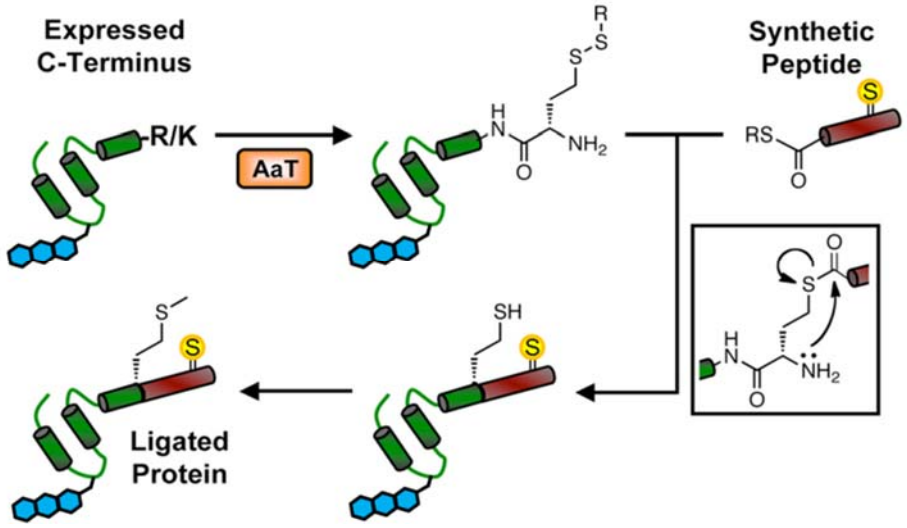

Figure 3.17 Combining Uaa Mutagenesis with NCL at Hcs to Generate Full-Length aS Containing a Thioamide/Acd Pair.

§3.6 Results and Discussion

The expressed C-terminal Acd-containing αS fragment for NCL was generated as shown in Figure 3.17. Previously, we have shown that we can use Met_5Lys_6 as the point of disconnection.[210] To diminish perturbation of native αS dynamics resulting from Acd incorporation, we chose to replace the aromatic residue phenylalanine (F_{94}). Our laboratory and others have shown that mutation of F94 to a bulkier aromatic amino acid, tryptophan, does not abrogate amyloid formation.[116] A plasmid containing $\alpha S_{6\text{-}140}$ with a TAG codon installed at position 94 was prepared ($\alpha S_{6\text{-}140}TAG_{94}$). In this construct, the αS gene is preceded by an N-terminal His_{11} tag and a Factor Xa proteolysis site. *E. coli* cells were dually transformed with a plasmid containing a UaaRS/tRNA pair selected for Acd incorporation as well as the plasmid containing $His_{11}\text{-}\alpha S_{6\text{-}140}TAG_{94}$. Following the initial growth period, protein expression was induced by adding acridonyl-2-alanine and IPTG

to the culture medium. His_{11}-$\alpha S_{6\text{-}140}Acd_{94}$ was purified by Ni-affinity chromatography and subsequently cleaved by Factor Xa to yield $\alpha S_{6\text{-}140}Acd_{94}$ bearing the requisite N-terminal lysine for the AaT ligation reaction. Following purification by ion-exchange chromatography, $\alpha S_{6\text{-}140}Acd_{94}$ was incubated with AaT, Met*RS, tRNA, Hcm, and ATP for 2 h resulting in complete formation of $\alpha S_{5\text{-}140}Hcm_5Acd_{94}$. $\alpha S_{5\text{-}140}Hcm_5Acd_{94}$ was subjected to an additional round of ion-exchange chromatography prior to carrying out the native chemical ligation reaction. Despite performing several rounds of ion-exchange chromatography, a significant amount of Tyr-containing protein was observed by MALDI-MS.

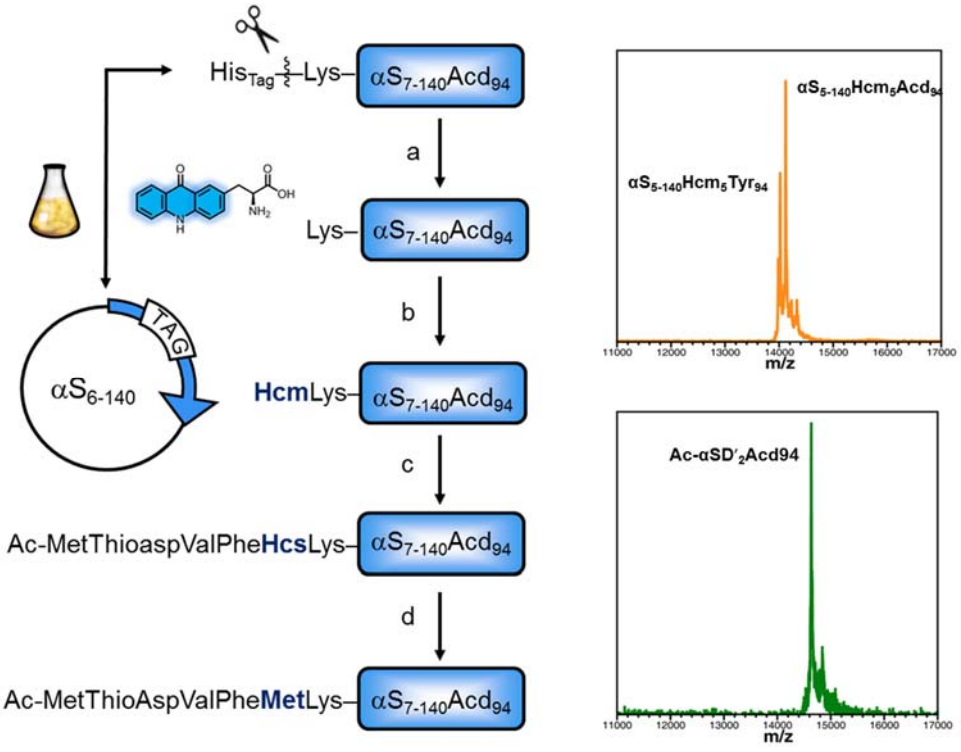

Figure 3.18 Ac-αSD′$_2$Acd$_{94}$ Synthesis.
Left: Functionalization of $\alpha S_{6\text{-}140}Acd_{94}$ by a) cleavage of His_{11} tag at Factor Xa site, b) attachment of Hcm by AaT-catalyzed modification, c) ligation of an N-terminal thioester peptide, and d) conversion of Hcs to Met by methylation to form Ac-$\alpha S^F D'_2 Cnf_{39}$. MALDI-MS analysis of $\alpha S_{6\text{-}140}Hcm_5Acd_{94}$. Calcd m/z $(M + H)^+$ 14121, found 14122 (corresponding to $\alpha S_{5\text{-}140}Hcm_5Acd_{94}$) and 14102 (corresponding to $\alpha S_{5\text{-}140}Hcm_5Tyr_{94}$).

Native chemical ligation was performed by incubating $\alpha S_{5\text{-}140}Hcm_5Acd_{94}$ (in the presence of the Tyr-containing protein contaminant) with 2 equivalents of AcMD'VF-SR in degassed ligation buffer. Complete disulfide deprotection of Hcm was observed by MALDI MS within five minutes (data not shown). The ligation reaction was allowed to proceed for 6 h, at which point unligated starting materials were not observed within the limits of detection by MALDI MS. As expected, two ligation products were formed, due to misincorporation of Tyr by the Acd synthetase. Fortuitously, these products were separable by HPLC, and Ac-$\alpha SD'_2Hcs_5Acd_{94}$ was cleanly isolated and confirmed by MALDI MS analysis. Finally, purified $\alpha SD'_2Hcs_5Acd_{94}$ was converted to $\alpha SD'_2Acd_{94}$ by a short treatment (10 min) with 3,000 equivalents of aqueous MeI.

Although we were able to demonstrate that expressed protein ligation at Hcs could be combined with Acd incorporation to yield full-length αS containing a thioamide/Acd pair, we encountered several unexpected experimental challenges during this process. First, the yields of protein expression were extremely low (.5 mg – 1 mg/L). Additionally, we observed significant incorporation of Tyr during protein expression. Since we did not observe Tyr incorporation during the expression of other proteins, we suspect at this stage that this issue may be vector and/or plasmid sequence dependent.

Upon repeating these initial studies, we observed an additional experimental issue that required immediate attention. By MALDI-MS, we found evidence of misincorporation of N-phenyl-p-aminophenylalanine (Npf) resulting from decarboxylation during the global deprotection/cyclization step of Acd synthesis. This was true even when Npf was present as only a trace (<1% by HPLC) impurity in the Acd batch.[47] Since this byproduct, which is difficult to remove by silica chomatography, is

more likely to be obtained using sulfuric acid, a two-step protocol was adopted with saponification using LiOH followed by polyphosphoric acid (PPA) cyclization (Figure 3.19). Protein expressed using Acd batches from this PPA route consistently yielded pure protein containing no Npf at the site of interest.

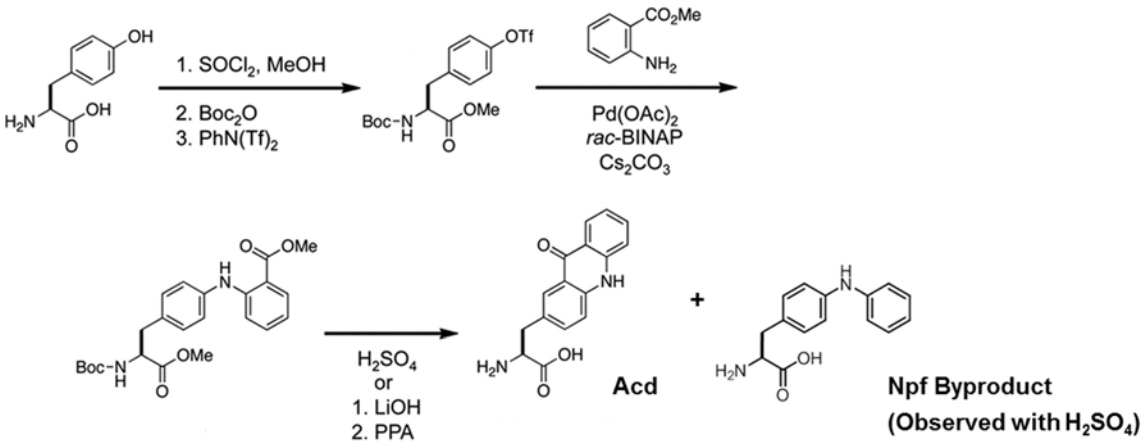

Figure 3.19 Synthesis of Acridon-2-ylalanine.
In the final step, global deprotection using H_2SO_4 yields Acd as well as trace amounts of the Npf contaminant. Performing basic hydrolysis of OMe followed by PPA-mediated Boc deprotection/Friedel crafts cyclization yields pure Acd for genetic incorporation.

Recently, the Mehl laboratory has performed negative selections to identify a mutant Acd synthetase that can effectively discriminate against Npf and Tyr misincorporation. With this improved Acd synthetase in hand, we should be able to dramatically improve the yields of Acd-containing truncated αS mutants for EPL.

§ 3.7 Conclusions

The application of our minimally perturbing thioamide/Cnf FRET pair to studies of full-sized proteins requires the ability to synthesize the labeled constructs in an efficient manner. Here, we have shown that can combine methods of unnatural amino acid mutagenesis with protein ligation to produce homogeneous doubly-labeled αS for FRET studies. We were able to show that Cnf and thioamides can be installed at either terminus of αS. Additionally, we demonstrated that we can monitor conformational

changes in monomeric αS using urea or TMAO to denature or compact the protein, respectively. Although thioamide/Cnf FRET is useful for studying the monomeric conformation of αS, Cnf cannot be selectively excited in the presence of αS oligomers. We have therefore shown that we can use our traceless double-labeling strategy to incorporate the red-shifted thioamide/Acd PET pair into full-length αS for misfolding studies. Importantly, these methods are not limited to studies of αS; virtually any protein that is compatible with NCL can be doubly-labeled with thioamide/fluorophore pairs in a similar fashion.

§ 3.8 Materials and Methods

General Information. Fmoc-L-4-cyanophenylalanine (Fmoc-Cnf-OH) was purchased from Peptech (Burlington, MA). Boc-L-thionophenylalanine-1-(6-nitro)benzotriazolide, Fmoc-Gln(Trt)-OH, Fmoc-Asn(Trt)-OHwere purchased from Bachem (Torrance, CA) or EMD Chemicals (Philadelphia, PA). Benzotriazol-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP), Fmoc-Ala-OH, Fmoc-Leu-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Th(tBu)-OH, Fmoc-Met-OH, Fmoc-Gly-OH, Fmoc-Val-OH, Fmoc-Phe-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Cys(Trt)-OH A QuickChange® site-directed mutagenesis kit was purchased from Stratagene (La Jolla, CA). DNA oligomers were purchased from Integrated DNA Technologies, Inc (Coralville, IA). DNA sequencing was performed at the University of Pennsylvania DNA sequencing facility. Protein purification was conducted on a BioCad Sprint fast protein liquid chromatography (FPLC) system (GMI Inc.; Ramsey, MN; originally from Perseptive Biosystems). HPLC purification was carried out on a Varian Prostar system (currently Agilent Technologies;

Santa Clara, CA). Intein plasmid pTXB1 and all restriction enzymes were purchased from New England Biolabs (Ipswich, MA). 2-chlorotrityl chloride resin, and Rink amide resin were purchased from Novabiochem (San Diego, CA). Piperidine and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from American Bioanalytical (Natick, MA). Sigmacote, *N*,*N*-diisopropyl ethylamine (DIPEA), thiophenol, trifluoroacetic acid (TFA), tris(2-carboxyethyl)phosphine hydrochloride (TCEP), were purchased from Sigma-Aldrich (St. Louis, MO). Ni-NTA resin was from Qiagen (Valencia, CA). *E. coli* BL21(DE3) cells were purchased from Stratagene (La Jolla, CA). Sequencing-grade trypsin was purchased from Promega (Madison, WI). Restriction Grade Factor Xa protease was purchased from Novagen (San Diego, CA). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA). Milli-Q filtered (18 MΩ) water was used for all solutions (Millipore; Billerica, MA). Matrix-assisted laser desorption ionization (MALDI) mass spectra were collected with a Bruker Ultraflex III MALDI-TOF-TOF mass spectrometer (Billerica, MA). UV/Vis absorbance spectra were obtained with a Hewlett-Packard 8452A diode array spectrophotometer (Agilent Technologies, Santa Clara, CA). Fluorescence spectra were collected with a Varian Cary Eclipse fluorescence spectrophotometer fitted with a Peltier multicell holder (currently Agilent Technologies).

The following specifications are specific to work performed described in section 3.3

Protein purification was conducted on an ÄKTA fast protein liquid chromatography (FPLC) system (GE Healthcare Life Sciences; Piscataway Township, NJ, USA). Fluorescence spectra were collected using a Photon Technology International (PTI) Quantamaster 40 fluorescence spectrometer (Birmingham, NJ, USA). Proton (^{1}H) and

carbon (^{13}C) NMR spectra were collected with a Bruker DRX 500 MHz spectrometer (Billerica, MA, USA). DEPT- 135 peaks are indicated parenthetically after ^{13}C chemical shifts. High resolution ESI mass spectrometry (HRMS) was conducted with a Waters LCT Premier XE LC/MS (Milford, MA, USA).

CaM Peptide Synthesis and Purification. CaM binding peptides were synthesized on Rink amide resin. Peptides were purified by reverse-phase HPLC using a binary system of aqueous (Buffer A: water + 0.1% TFA) and organic (Buffer B: CH_3CN + 0.1% TFA) phases. Solvent gradients, column descriptions, and retention times are listed in Tables 3.11 and 3.12. Purified peptides were lyophilized or dried in a vacuum centrifuge.

Table 3.2 Summary of Purified Peptides.

Peptide	Gradient Description†	Retention Time (min)	Column (Size: Brand)
FRRIARLVGVLREFAFR (pOCNC)	S1	19.6	Semi-prep: Vydac 218 TP C18
F'RRIARLVGVLREFAFR (pOCNC-F′$_1$)	S1	20.6	Semi-prep: Vydac 218 TP C18
FRRIARLVGVL'REFAFR (pOCNC-L′$_{11}$)	S1	22.1	Semi-prep: Vydac 218 TP C18
FRRIARLVGVLREFAF'R‡ (pOCNC-F′$_{16}$)	S2	13.3, 14.2	Prep: Waters C18

†Described in Table 3.12 ‡Separable epimers were observed during HPLC purification.

Table 3.3 Solvent Gradients Used for Peptide Purification and Analysis.

Gradient	Time (min)	Buffer A (%)	Gradient	Time (min)	Buffer A (%)
S1	0:00	98	S2	0:00	98
	5:00	98		5:00	98
	8:00	71		9:00	69
	29:00	59		25:00	67
	33:00	0		27:00	0
	38:00	0		30:00	0
	43:00	98		35:00	98

Table 3.4 Calculated and Observed Peptide and Protein Masses.

Peptide	Calculated m/z $[M+H]^+$	Observed m/z $[M+H]^+$	Calculated m/z $[M+Na]^+$	Observed m/z $[M+Na]^+$
FRRIARLVGVLREFAFR (pOCNC)	2106.26	2106.75	2128.24	--
F'RRIARLVGVLREFAFR (pOCNC-F'_1)	2122.23	2122.43	2128.24	--
FRRIARLVGVL'REFAFR (pOCNC-L'_{11})	2122.23	2122.16	2128.24	--
FRRIARLVGVLREFAF'R‡ (pOCNC-F'_{16})	2122.23	2122.62	2128.24	--
CaM $F^*_{13}F_{100}F_{139}$ ($CaM^FF^*_{13}$)	16700.3	16699.8	16722.3	--
CaM $F^*_{93}F_{100}F_{139}$ ($CaM^FF^*_{93}$)	16700.3	16699.9	16722.3	--
CaM $F^*_{100}F_{139}$ ($CaM^FF^*_{100}$)	16700.3	16700.6	16722.3	--

DNA Oligomers used for CaM Quikchange® Mutagenesis

a. Mutation F_{13} to TAG
Forward: 5' – ACTGACAGAAGAGCAGATTGCAGAATAGAAAGAAGCTTTTTCACTATTTGAC – 3'
Reverse: 5' – GTCAAATAGTGAAAAAGCTTCTTTCTATTCTGCAATCTGCTCTTCTGTCAGT – 3'

b. Mutation F_{17} to TAG
Forward: 5' – GATTGCAGAATTCAAAGAAGCTTAGTCACTATTTGACAAGGATGGT – 3'
Reverse: 5' – ACCATCCTTGTCAAATAGTGACTAAGCTTCTTTGAATTCTGCAATC – 3'

c. Mutation F_{93} to TAG
Forward: 5' – AATTAGAGAAGCGTTCCGTGTGTAGGACAAGGATGG – 3'
Reverse: 5' – CCATCCTTGTCCTACACACGGAACGCTTCTCTAATT – 3'

d. Mutation Y_{100} to F
Forward: 5' – GACAAGGATGGTAATGGTTTCATTAGTGCTGCAGAACTT – 3'
Reverse: 5' – AAGTTCTGCAGCACTAATGAAACCATTACCATCCTTGTC – 3'

d. Mutation Y_{100} to TAG
Forward: 5' – GACAAGGATGGTAATGGTTAGATTAGTGCTGCAGAACTTCG – 3'
Reverse: 5' – CGAAGTTCTGCAGCACTAATCTAACCATTACCATCCTTGTC – 3'

e. Mutation Y_{139} to F
Forward: 5' – AGACATTGATGGTGATGGTCAAGTAAACTTTGAAGAGTTTGTACA – 3'
Reverse: 5' – TGTACAAACTCTTCAAAGTTTACTTGACCATCACCATCAATGTCT – 3'

Figure 3.20 DNA Oligomers for CaM Quikchange® Mutagenesis.

Cloning of Calmodulin Expression Constructs. A plasmid containing the wild-type chicken calmodulin (CaM) gene was provided by Joshua Wand from the University of Pennsylvania School of Medicine. An insert containing the CaM gene was cloned into a pET15b vector (Novagen, Gibbstown, NJ) between the NcoI and XhoI cut sites. Quikchange® mutagenesis was used to generate the following mutant plasmids: pET15b-

CaM-TAG$_{13}$F$_{100}$F$_{139}$, pET15b-CaM-TAG$_{17}$F$_{100}$F$_{139}$, pET15b-CaM-TAG$_{93}$F$_{100}$F$_{139}$, and pET15b-CaM-TAG$_{100}$F$_{139}$.

CaM Expression and Purification. Each plasmid containing a CaM mutant (pCaM-TAG$_{13}$, pCaM-TAG$_{17}$, pCaM-TAG$_{93}$, or pCaM-TAG$_{100}$, see Supporting information for plasmid construction) was transformed into pDULE2-Cnf-containing *E. coli* BL21(DE3) cells. These cells were cultured and made competent through the Hanahan method.[211] First, the pDULE2-Cnf plasmid was transformed into BL21(DE3) cells. Then, a colony was picked based on streptomycin (Strep) resistance to grow cultures of competent cells. Once pCaM was transformed into these cells, selection was based on both ampicillin (Amp) and Strep resistance. Single colonies were used to inoculate 5 mL of LB media containing Amp (100 μg/mL) and Strep (100 μg/mL). This primary 5 mL culture was incubated at 37 °C with shaking at 250 RPM overnight, and then added to 1 L of M9 minimal media containing Amp and Strep at the same concentrations. M9 minimal media was prepared by adding the following autoclaved solutions to 1 L of autoclaved water containing 6 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl and 1 g NH_4Cl: 1 mL of 2 M $MgSO_4$, 1 mL of 15 mg/mL $FeCl_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL $ZnCl_2$ (in acidified H_2O), 2 mL of 10% Bacto™ Yeast Extract, 20 mL of 10% glycerol (v/v) and 1 μL of 1 M $CaCl_2$. The 1 L secondary culture was incubated at 37 °C with shaking at 250 RPM until the OD_{600} reached 0.7 - 0.9 AU. Protein expression was induced with isopropyl D-galactoside (IPTG) and 95 mg (0.5 mM) Cnf, and the culture was incubated at 37 °C for an additional 12 h. The cells were recovered from media at 6000 x g for 10 minutes. The resulting pellet was suspended in 15 mL of CaM resuspension buffer (50 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS), 100 mM KCl, 1 mM

ethylenediaminetetraacetic acid (EDTA), pH 7.5). Cell were lysed by sonication, alternating between one minute of sonication. $CaCl_2$ was added to the sonicated lysate to a final concentration of 20 mM prior to centrifugation for 10 minutes at 6,000 x g, 4 °C. Following centrifugation, the supernatant was collected for purification. CaM was purified from the cleared cell lysate using a phenyl-sepharose (PS) CL-4B column with EDTA as the eluent. A column with a total resin bed volume of approximately 8 mL was first equilibrated with four 10 mL washes of PS Buffer A (50 mM Tris base, 1 mM $CaCl_2$). Then the cleared cell lysate was added to the column and collected as it ran though to save any protein that did not bind to the column. Next, the column was washed four times with 10 mL PS Buffer A. The column was then washed with four 10 mL portions of high-salt PS Buffer B (50 mM Tris base, 0.5 M NaCl, 0.1 mM $CaCl_2$). Finally, an additional 10 mL wash of PS Buffer A was performed to restore the column to a low-salt condition. CaM was eluted with 18 mL of PS Buffer C (10 mM Tris base, 10 mM EDTA, pH 7.5. This elution solution was then re-saturated with $CaCl_2$ to a concentration of 20 mM. This elution solution was repurified on a second PS column using the same procedure to reduce truncation product. For the second PS column, rather than adding 18 mL of PS Buffer C directly to the column, 1 mL elutions were collected and analyzed by SDS-PAGE gel. Collection stopped when the protein was no longer visible by gel. The elution fractions were dialyzed against CaM Buffer (140 mM KCl, 15 mM HEPES, 6 mM $CaCl_2$, pH 6.7) and stored at -20 °C. Proteins were analyzed by SPS PAGE gel and MALDI MS for purity.

Native PAGE Gel Analysis of CaM Mutants. CaM peptide binding can be detected by native (non-denaturing) PAGE gel analysis.[212] In order to determine whether mutant CaM

was capable of binding peptide, 25 μL samples were prepared containing 10 μM of each CaM mutant, with or without a stoichiometric equivalent of pOCNC-F′$_1$. Prepared samples were then incubated for one hour at 4 °C and loaded into a non-denaturing gel with 2.5 μL of 60% glycerol and 2.5 μL of .01% bromophenol blue to assist loading. The gel was stained with Coomassie blue and scanned.

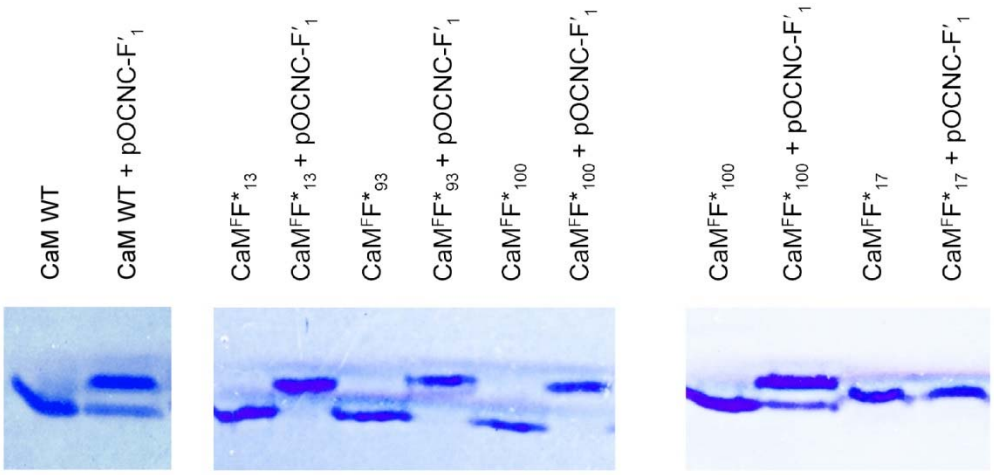

Figure 3.22 Native PAGE analysis of CaM mutants.
Left: PAGE analysis demonstrates upward shift of the CaM protein band upon addition of pOCNC-F′$_1$. No shift is observed upon addition of pOCNC-F′$_1$ to CaM^{F}F*$_{17}$, indicating this CaM mutant does not bind pOCNC-F′$_1$.

Calmodulin Circular Dichoism Measurements. Mutant stability was evaluated though both wavelength-dependent and temperature-dependent circular dichoism (CD) spectroscopy. The wavelength-dependent signature for both the calcium-containing CaM (holo CaM) and calcium-free CaM (apo CaM) was examined. Since holo CaM is thermostable, only the thermal unfolding of apo CaM was determined. The apo form of the protein was prepared by dialyzing the protein into 2 mM EDTA in 50 mM HEPES, pH 6.70. The holo form of the protein was prepared by dialyzing into 6 mM $CaCl_2$, 15 mM HEPES, 140 mM KCl, pH 6.70. Prepared samples containing 20 μM CaM were monitored at 222 nm between 5 and 95 °C using the variable temperature module for

temperature-dependent scans and between 200 and 300 nm at room temperature for wavelength-dependent scans. Data were collected with a 1 °C/min slope, 30 s averaging time, 2 min temperature equilibration, 5 s response, and 1 nm band width. The data were processed as previously described.

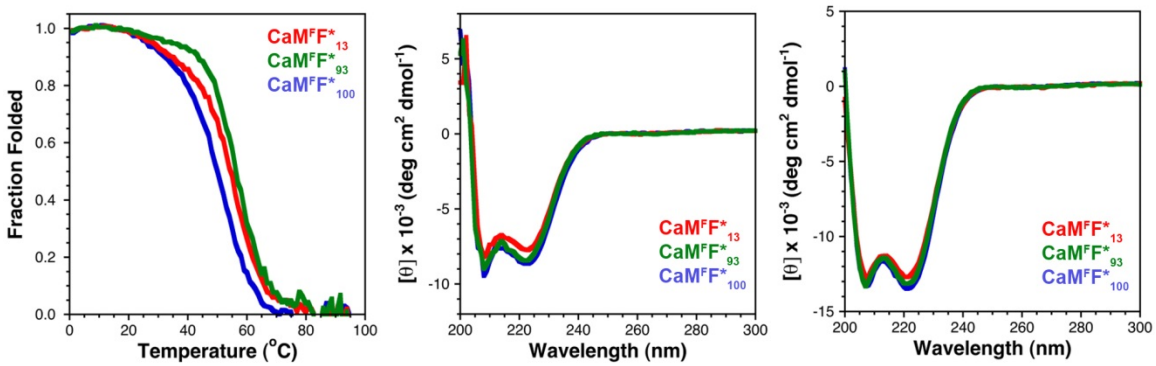

Figure 3.21 Temperature and Wavelength-Dependent Circular Dichoism Measurements of CaM Mutants.
Left: Temperature-dependent trace shown for apo CaM mutants. Middle: Wavelength-dependent trace (25 °C) shown for apo CaM mutants. Right: Wavelength-dependent trace (25 °C) shown for holo CaM mutants. Coloring scheme for all thee plots $CaM^{F}F^{*}_{13}$ (red), $CaM^{F}F^{*}_{93}$ (green), $CaM^{F}F^{*}_{100}$ (blue).

Fluorescence Spectroscopy. Corrected fluorescence spectra were collected in triplicate using quartz fluorometer cells with path lengths of 1.00 cm. For all Cnf experiments, the excitation wavelength was 240 nm and emission data was collected from 275 - 400 nm as the average of thee scans. The excitation and emission slit widths were 5 nm, the scan rate was 120 nm/min, the averaging time 0.5 s, and the data interval 1.0 nm.

CaM/pOCNC Binding Assays. After synthesis, HPLC purification, and lyophilization, the various pOCNC peptides were redissolved in CaM buffer (140 mM KCl, 10 mM HEPES, 6 mM $CaCl_2$, pH 6.7). The concentrations of the peptides were determined by UV/Vis spectroscopy, using an extinction coefficient of 10, 270 $M^{-1}cm^{-1}$ at 273 nm for the thioamide-containing peptides and 600 $M^{-1}cm^{-1}$ at 257 nm for the oxoamide-containing peptide. The concentration of the CaM proteins was determined though the

bicinchioninic acid (BCA) assay. For each peptide/protein combination, six solutions were prepared that contained 10 μM CaM protein and 0 to 15 μM pOCNC peptide. In addition, a solution of 10 μM peptide in the absence of protein was made to confirm that no background fluorescence resulted from the peptide. The solutions were prepared three times in order to obtain data in triplicate. E_Q was computed from the raw fluorescence data for CaM with thioamide-labeled pOCNC (F_{Thio}) or pOCNC (F_{Oxo}).

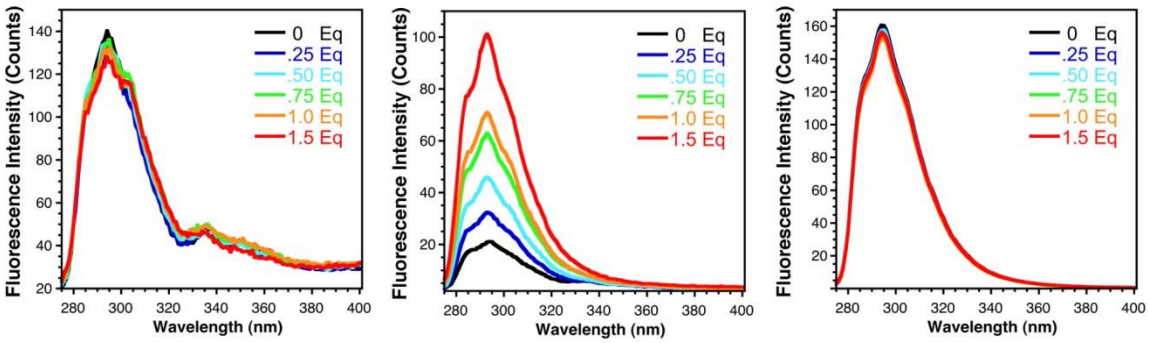

Figure 3.22 Fluorescence scans of pOCNC binding with CaM mutants.
Left: $CaM^{F}F^{*}_{13}$, Middle: $CaM^{F}F^{*}_{93}$, Right: $CaM^{F}F^{*}_{100}$.

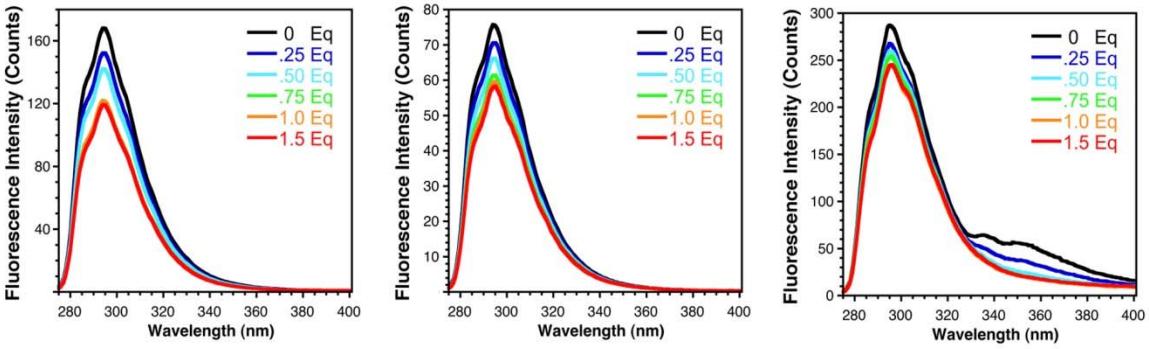

Figure 3.23 Fluorescence scans of thioamide pOCNC mutants binding with $CaM^{F}F^{*}_{100}$.
Left: pOCNC-F′$_1$, Middle: pOCNC-L′$_{11}$, Right: pOCNC-F′$_{16}$.

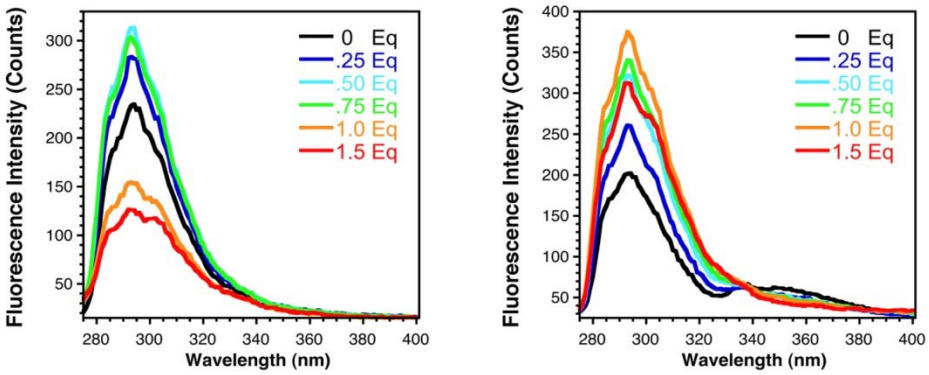

Figure 3.24 Fluorescence scans of pOCNC mutants binding with CaM^{F}F*$_{93}$.
Left: pOCNC-F′$_{1}$, Right: pOCNC-L′$_{11}$.

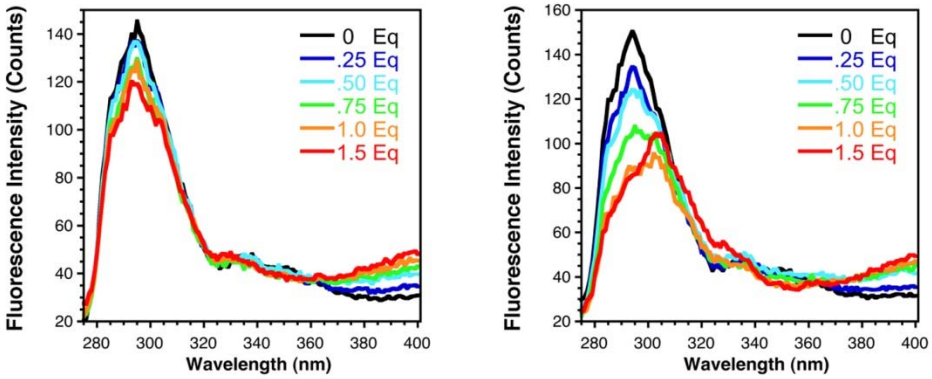

Figure 3.25 Fluorescence Scans of pOCNC Thioamide Tutants Binding with CaM^{F}F*$_{13}$.
Left: pOCNC-F′$_{1}$, Right: pOCNC-L′$_{11}$.

Titration Analysis of Binding Studies with CaM Mutants. Normalized titration curves were obtained by plotting the ratio of the fluorescence of thioamide pOCNC complexes to the corresponding oxoamide pOCNC complexes (both measured at 295 nm) as a function of CaM/pOCNC ratio.

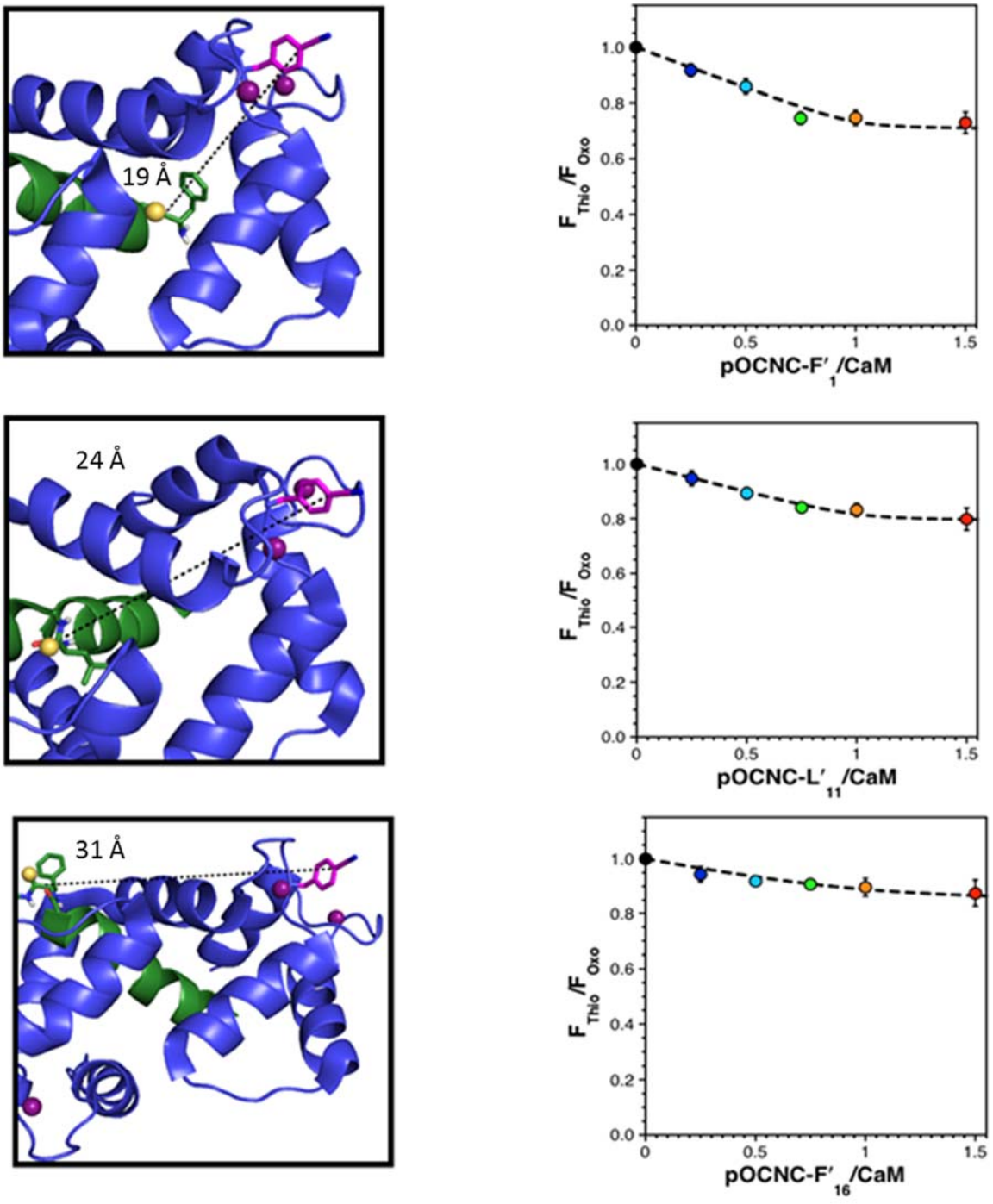

Figure 3.26 Binding of Thioamide-Containing-pOCNC with CaM^{F}F*$_{100}$.
In each set, Left: Pymol figure illustrating the distance between the chromophores. Right: titration data for the binding experiment. Top: pOCNC-F′$_1$, Middle: pOCNC-L′$_{11}$, Bottom: pOCNC-F′$_{16}$.

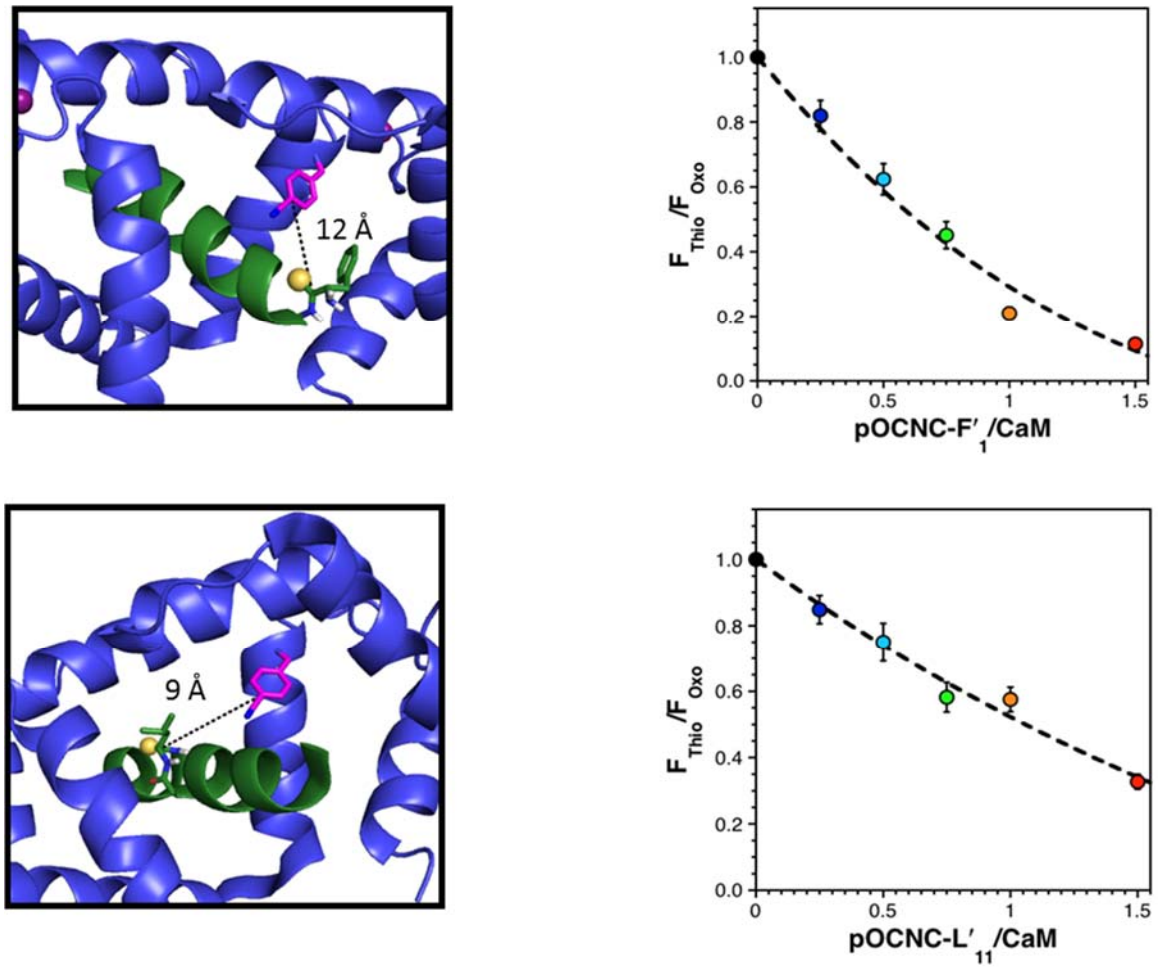

Figure 3.27 Binding of Thioamide-Containing-pOCNC with CaM^{F}F*$_{93}$.
In each set, Left: Pymol Figure illustrating the distance between the chromophores. Right: titration data for the binding experiment. Top: pOCNC-F′$_1$, Bottom: pOCNC-L′$_{11}$.

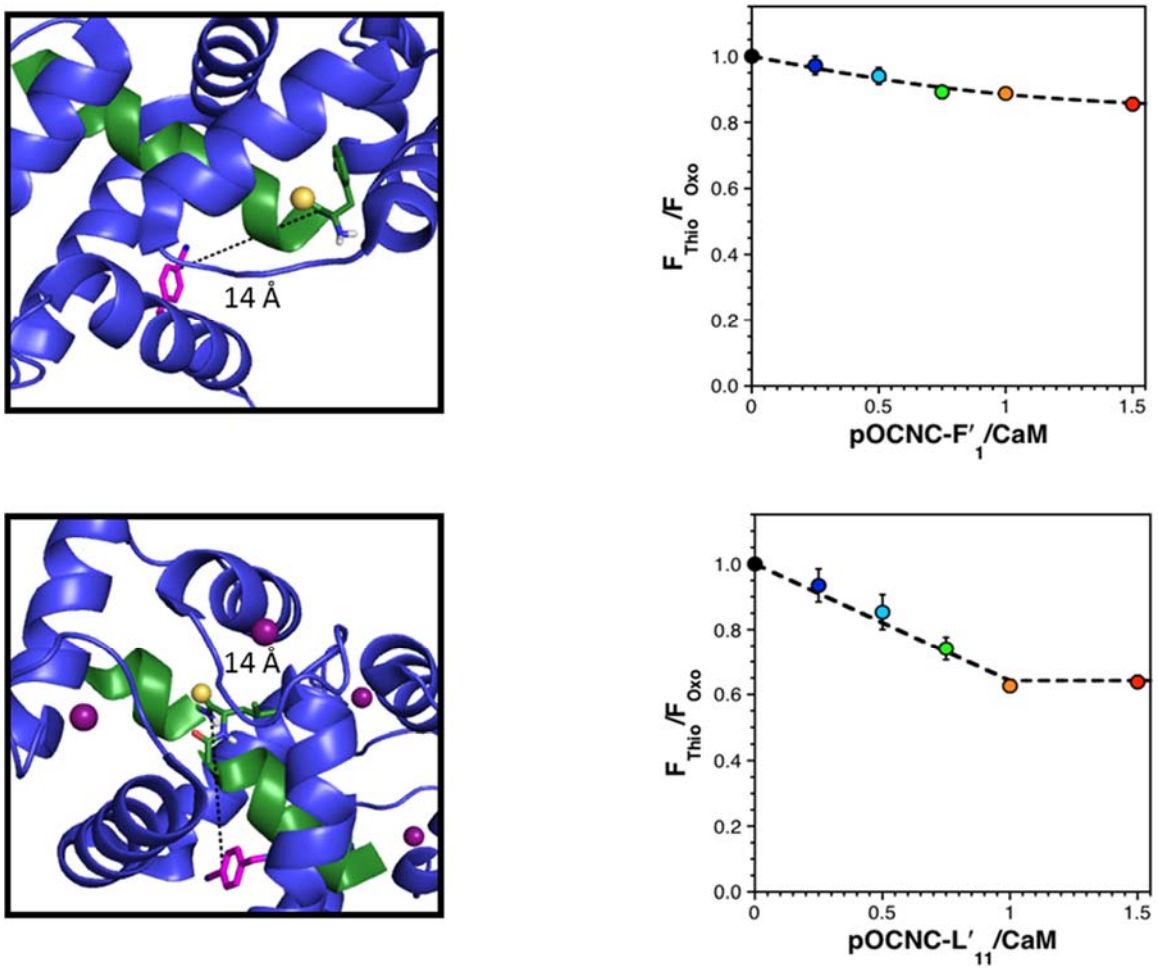

Figure 3.28 Binding of Thioamide-Containing-pOCNC with CaM^{F}F*$_{13}$.
In each set, Left: Pymol Figure illustrating the distance between the chromophores. Right: titration data for the binding experiment. Top: pOCNC-F′$_1$, Bottom: pOCNC-L′$_{11}$.

Förster Distance Calculation. The Förster distance is given in Å by equation 3.1

$$R_0^6 = \frac{9000(\ln 10)\kappa^2 Q_D J}{128\pi^5 n^4 N_A} \qquad \text{(Eq. 3.1)}$$

where κ^2 is a geometrical factor that relates the orientation of the donor and acceptor transition moments, Q_D is the quantum yield of the donor, n is the index of refraction of the solvent, N_A is Avogadro's number, and J is the spectral overlap integral defined in units of $M^{-1}\cdot cm^{-1}\cdot nm^4$. Combining constants and rearranging gives R_0 as

$$R_0 = 0.211\{Q_D \kappa^2 n^{-4} J\}^{1/6} \qquad \text{(Eq. 3.2)}$$

J is formally defined as

$$J = \int_0^\infty f_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \qquad \text{(Eq. 3.3)}$$

where $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor at each wavelength λ and $f_D(\lambda)$ is the normalized donor emission spectrum given by

$$f_D(\lambda) = \frac{F_{D\lambda}(\lambda)}{\int_0^\infty F_{D\lambda}(\lambda)d\lambda} \qquad \text{(Eq. 3.4)}$$

where $F_{D\lambda}(\lambda)$ is the fluorescence of the donor at each wavelength λ. J was calculated to be $7.0\times10^{12} \pm 2\times10^{11}$ $M^{-1}\bullet cm^{-1}\bullet nm^4$ as previously described.[115] Substituting this result into equation 3.2 , as well as 0.11 for the quantum yield of Cnf, 1.33 for the index of refraction of water, and 2/3 for κ^2 gives R_0 = 15.6 Å for Cnf/thioamide FRET pairs with Cnf at position 100 in CaM.[213] For Cnf/thioamide FRET pairs with Cnf at position 13 in CaM, we use 0.003 as a quantum yield for Cnf, and calculate R_0 = 8.6 Å. For Cnf/thioamide FRET pairs with Cnf at position 93 in CaM, we use 0.014 as a quantum yield for Cnf, and calculate R_0 = 11.1 Å. Quantum yields for each mutant were determined by comparing the fluorescence output of the mutant relative to the fluorescence of $CaM^FF^*_{100}$ under oxopeptide (pOCNC) bound conditions using identical concentrations and fluorometer settings. Comparison of $CaM^FF^*_{100}$ fluorescence to free Cnf showed that no substantial quenching occurred in this mutant in the absence of thiopeptide (data not shown). These values of R_0 were then used in determining experimental chomophore separations (R_{FRET}) for each set of Cnf/thioamide probe locations according to equation (3.5)

$$R_{FRET} = R_0\left(\frac{1}{E_Q} - 1\right)^{1/6} \qquad \text{(Eq. 3.5)}$$

where E_Q is defined by equation (3.6)

$$E_Q = 1 - F_{Thio} / F_{Oxo} \qquad \text{(Eq. 3.6)}$$

The resulting R_{FRET} values are shown in Table 3.1.

Chromophore Geometries were determined by E. James Petersson and Colin Fadzen.

Chromophore Geometries from CaM/pOCNC NMR Structure. For each pair of Cnf donor and thioamide acceptor locations, the corresponding residues were identified in the 1SYD PDB file, and a Tcl script was run in VMD (http://www.ks.uiuc.edu/) to cull distance and orientation information for the Cnf ring (using the corresponding positions in the natural Phe or Tyr in 1SYD) and the thioamide Xxx′ carbonyl (using the corresponding positions in the given amide backbone unit in 1SYD). The following positions were determined in accordance with the vectors identified in previous quantum mechanical calculations: the midpoint between CE1 and CE2 (CEmp) of Cnf as a proxy for the center of the Cnf transition dipole; the midpoint between Xxx′ C and N of the next amino acid (CNmp) as a proxy for the midpoint of the thioamide dipole.[115] The interchomophore distance (R_{NMR}) was determined as the distance between these two midpoints. The culled data were also used to calculate theoretical FRET efficiencies for each set of probe locations. The donor-acceptor dihedral angle (ϕ_{DA}) was determined as the Xxx′ O, CNmp, CEmp, Phe CE2 dihedral angle. The donor angle (θ_D) was determined as the Xxx′ O, CNmp, CEmp angle; and acceptor angle (θ_A) was determined as the CNmp, CEmp, CE2 angle. These parameters are illustrated in Figure 3.30.

$R_{CNmp/CEmp}$ = 19.3 Å
θ_D = 114.7°
θ_A = 67.5°
ϕ_{DA} = 91.4°

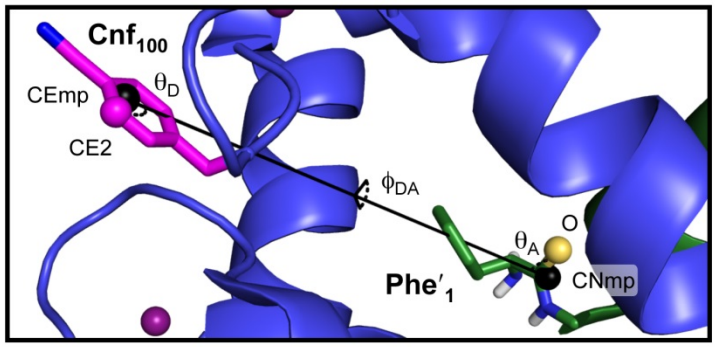

Figure 3.29 FRET Orientational Parameters.
Structure 1 from PDB 1SYD shown with atoms used in determining κ^2 and relevant distances and angles illustrated. CNmp and CEmp indicated by black spheres. Shown with nitrile and thioamide modifications for clarity, although geometry calculations are identical for unmodified 1SYD file.

The orientational parameter from Förster theory, κ^2, was calculated as:

$$\kappa^2 = (\sin\theta_D \sin\theta_A \cos\phi_{DA} - 2\cos\theta_D \cos\theta_A)^2 \qquad \text{(Eq. 3.7)}$$

where ϕ_{DA}, θ_D, and θ_A are defined as above. Geometry-specific R_0 values were calculated for each of the twenty structures in PDB 1SYD using equations 3.2 and 3.3 where the donor quantum yield (Q_D) is taken as 0.11, the spectral overlap integral (J) is calculated as above, and the index of refraction (n) is taken as 1.33.[213] For each timestep, this value of R_0 was used to calculate a geometry-specific FRET quenching efficiency (E_Q) using the corresponding donor-acceptor separation, R_{NMR} (the CNmp/CEmp distance). E_Q was calculated using equation (3.6). Values averaged over all 20 structures in PDB 1SYD for R_{NMR}, κ^2, R_0, and E_Q are collected in Table 3.5 These R_{NMR} values are listed in Table 3.1.

Table 3.5 Orientation Parameters and Theoretical E_{FRET} from CaM/pOCNC NMR Structures[a]

CaM_F_{13}/pOCNC_F_1 (Φ = 0.003)	R_{NMR}	κ^2	R_0	E_Q
Average	14.1	1.01	8.0	0.07
RMSD	0.3	0.63	1.5	0.04
CaM_F_{13}/pOCNC_L_{11} (Φ = 0.003)	R_{NMR}	κ^2	R_0	E_Q
Average	14.1	0.92	8.1	0.07
RMSD	0.4	0.47	1.3	0.03
CaM_F_{93}/pOCNC_F_1 (0.014)	R_{NMR}	κ^2	R_0	E_Q
Average	8.9	1.12	11.1	0.73
RMSD	0.3	0.58	1.4	0.14
CaM_F_{93}/pOCNC_L_{11} (0.014)	R_{NMR}	κ^2	R_0	E_Q
Average	12.2	1.19	10.8	0.39
RMSD	0.5	0.68	1.8	0.15
CaM_F_{100}/pOCNC_F_1 (0.110)	R_{NMR}	κ^2	R_0	E_Q
Average	18.9	1.30	16.2	0.32
RMSD	0.3	0.57	2.1	0.11
CaM_F_{100}/pOCNC_L_{11} (0.110)	R_{NMR}	κ^2	R_0	E_Q
Average	24.4	1.02	15.2	0.09
RMSD	0.4	0.58	2.2	0.04
CaM_F_{100}/pOCNC_F_{16} (0.110)	R_{NMR}	κ^2	R_0	E_Q
Average	31.3	1.05	14.6	0.02
RMSD	0.7	0.68	2.8	0.01

[a] Parameters calculated as described in text. All distances in Å.

Epimer Analysis of pOCNC-F′$_{16}$. Upon purification of pOCNC-F′$_{16}$, two distinct peaks were observed with different retention times. We attribute this to epimerization resulting from a reversible cyclization reaction similar to the Edman degradation reaction. Both peaks contained the correct mass of pOCNC-F′$_{16}$ by MALDI MS. The fluorescence titration assay was performed with a combination of both epimers of the peptide. To confirm that both epimers bound CaM^{F}F*$_{100}$, a native PAGE gel was run with and without each epimer of pOCNC-F′$_{16}$ as described previously.

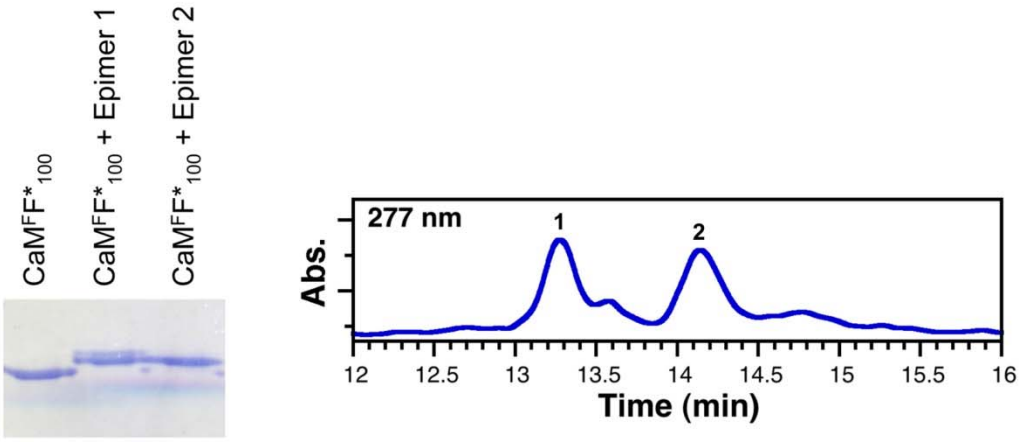

Figure 3.30 Native PAGE and HPLC analysis of pOCNC-F′$_{16}$.
Left: Gel analysis shows shift upon addition of each epimer of pOCNC-F′$_{16}$ to CaM^{F}F*$_{100}$, indicating binding. Right: HPLC purification trace at 277 nm. Peaks 1 and 2 correspond to the peptide epimers.

αS Peptide Synthesis and Purification. αS peptide fragments were synthesized on chlorotrityl resin. Peptides were purified by HPLC using a binary system of Buffer A and Buffer B. Thioesterification was performed as described in the main text. Solvent gradients, column descriptions, and retention times are listed in Tables **3.6** and **3.7**. Purified peptides were lyophilized or dried in a vacuum centrifuge. Purified proteins were exchanged into Tris buffer (20 mM Tris, 100 mM NaCl, pH 7.4).

Table 3.6 Summary of HPLC Purified Peptides and Proteins.

Peptide	Gradient Description[†]	Retention Time (min)	Column (Size: Brand)
Ac-MDVF'MKGL-S$(CH_2)_2CO_2$Me	S3	24.1	Semi-prep: YMCBasic C8
Ac-MDV'FMKGL-C^{b}PG$_o$	S4	31.9	Semi-prep: Vydac 218 TP C18
αS$_{123\text{-}140}$C$_{123}$A'$_{124}$	S5	20.5	Prep: Waters C18
Ac-αS^{F}F'$_4$C$_9$F*$_{39}$	S6	31.5	Semi-prep: Vydac 218 TP C4
Ac-αS^{F}F'$_4$C$_9$F*$_{94}$	S6	28.9	Semi-prep: Vydac 218 TP C4
αS^{F}F*$_{39}$C$_{123}$A'$_{124}$	S7	28.9	Semi-prep: Vydac 218 TP C4
Ac-αS^{F}V'$_3$C$_9$F*$_{39}$	S6	29.5	Semi-prep: Vydac 218 TP C4

Table 3.7 Calculated and Observed Peptide and Protein Masses.

Peptide	Calculated m/z $[M+H]^+$	Observed m/z $[M+H]^+$	Calculated m/z $[M+Na]^+$	Observed m/z $[M+Na]^+$
Ac-MDVF'MKGL-S$(CH_2)_2CO_2$Me	1101.38	1101.43	1123.36	---
Ac-MDV'FMKGL-C^bPG_o	1343.56	1343.91	1366.54	1365.89
$\alpha S_{123\text{-}140}C_{123}A'_{124}$	2078.86	2078.73	2100.66	2100.69
$\alpha S^FC^9F^*_{39}$	14438.1	14437.6	14461.1	---
$\alpha S^FC^9F^*_{94}$	14438.1	14439.6	14461.1	---
$\alpha S^FF^*_{39}C_{123}$	14386.1	---	14409.1	14411.3
$\alpha S^F_{\Delta 1\text{-}8}\ C_9F^*_{39}$	13516.0	13517.2	13534.0	---
Ac-$\alpha S^FF'_4C_9F^*_{39}$	14494.7	14495.7	14517.7	---
Ac-$\alpha S^FF'_4C_9F^*_{94}$	14494.7	14495.9	14517.7	---
Ac-$\alpha S^FV'_3C_9F^*_{39}$	14494.7	14495.4	14517.7	---
$\alpha S_{1\text{-}122}F^*_{39}SR$	12475.2	12475.5	12497.2	---
$\alpha S^FF^*_{39}C_{123}A'_{124}$	14412.2	14412.5	14434.2	---

Table 3.8 Solvent Gradients Used for Peptide Purification and Analysis.

Gradient	Time (min)	Buffer A (%)	Gradient	Time (min)	Buffer A (%)
S3	0:00	98	S4	0:00	98
	5:00	98		5:00	98
	10:00	85		10:00	75
	30:00	65		25:00	40
	35:00	0		30:00	0
	40:00	0		35:00	0
	45:00	98		40:00	98
S5	0:00	95	S6	0:00	95
	5:00	95		5:00	95
	10:00	77		10:00	70
	27:00	60		30:00	40
	30:00	0		34:00	0
	35:00	0		38:00	0
	40:00	98		43:00	95
S7	0:00	95			
	5:00	95			
	10:00	70			
	30:00	50			
	34:00	0			
	38:00	0			
	43:00	95			

Construction of Recombinant pET16b αS Expression Plasmids. The generation of pET16b-His_{11}IEGR-$\alpha S_{\Delta 1\text{-}8}C_9$ has been described previously. All tyrosine residues (Y_{39}, Y_{125}, Y_{133}, and Y_{136}) were mutated to phenylalanine by multiple rounds of QuikChange® mutagenesis to generate pET16b-His_{11}IEGR-$\alpha S^{F}_{\Delta 1\text{-}8}C_9$. The amber stop codon (TAG) was subsequently mutated into the pET16b-His_{11}IEGR-$\alpha S^{F}_{\Delta 1\text{-}8}C_9$ construct at nucleotides corresponding to position 39 or 94 to generate pET16b-His_{11}IEGR-$\alpha S^{F}_{\Delta 1\text{-}8}C_9TAG_{39}$ and pET16b-His_{11}IEGR-$\alpha S^{F}_{\Delta 1\text{-}8}C_9TAG_{94}$.

Construction of Recombinant pRK127 αS Expression Plasmids. A plasmid containing the human wild-type αS gene cloned between NdeI and HindIII in the

expression vector pRK172 was provided by Dr. Virginia Lee (Perelman School of Medicine, University of Pennsylvania). All tyrosine residues were mutated to phenylalanine by multiple rounds of Quikchange® mutagenesis to generate pRK172-αS^F. The amber stop codon (TAG) was subsequently mutated into the pRK172-αS^F construct at nucleotides corresponding to amino acid position 39 or 94 to generate pRK172-$\alpha S^F TAG_{39}$ and pRK172-$\alpha S^F TAG_{94}$. In order to create N-terminal cysteine constructs for native chemical ligation, a primer was designed to delete nucleotides corresponding to amino acids 2-8 while simultaneously mutating Ser_9 to Cys. This primer was applied to pRK172-$\alpha S^F TAG_{39}$ and pRK172-$\alpha S^F TAG_{94}$ to generate pRK172-$\alpha S^F{}_{\Delta 2\text{-}8} C_9 TAG_{39}$ and pRK172-$\alpha S^F{}_{\Delta 2\text{-}8} C_9 TAG_{94}$ (Figure 3.32). Ser_9 was mutated to Cys in all full-length oxoamide control constructs to yield pRK172-$\alpha S^F C_9 TAG_{39}$ and pRK172-$\alpha S^F C_9 TAG_{94}$.

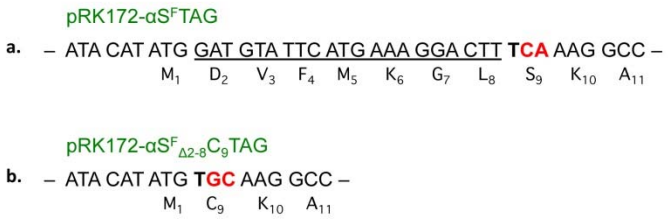

Figure 3.31 Construction of pRK172-$aS^F{}_{\Delta 2\text{-}8} C_9 TAG$ mutants.
A primer was designed to delete nucleotides corresponding to amino acid residues 2-8 (underlined) and mutate S9 to C (bolded) with a single round of Quikchange® mutagenesis.

Construction of Recombinant pTXB1-$\alpha S_{1\text{-}122} F^*_{39}\text{-}H_{Tag}$ Expression Plasmid. A plasmid containing a wild type αS gene was digested by NdeI and XhoI, gel-purified, and ligated into NdeI/XhoI site of the pTXB1 vector which contained mini-intein GyrA from the *Mycobacterium xenopi* followed by a chitin binding domain (CBD). An XhoI site (CTCGAG) was introduced at D_{121}/N_{122} of αS using QuikChange® mutagenesis . Digestion with XhoI and ligation of the two sticky ends yielded a pTXB1-$\alpha S_{1\text{-}122} D_{121} L\text{-}N_{122} E$-SapI site plasmid. Finally, QuikChange® mutagenesis was used to mutate $L_{121} D$

and $N_{122}E$ while simultaneously deleting the SapI site for generating pTXB1-$\alpha S_{1\text{-}122}$ plasmid.

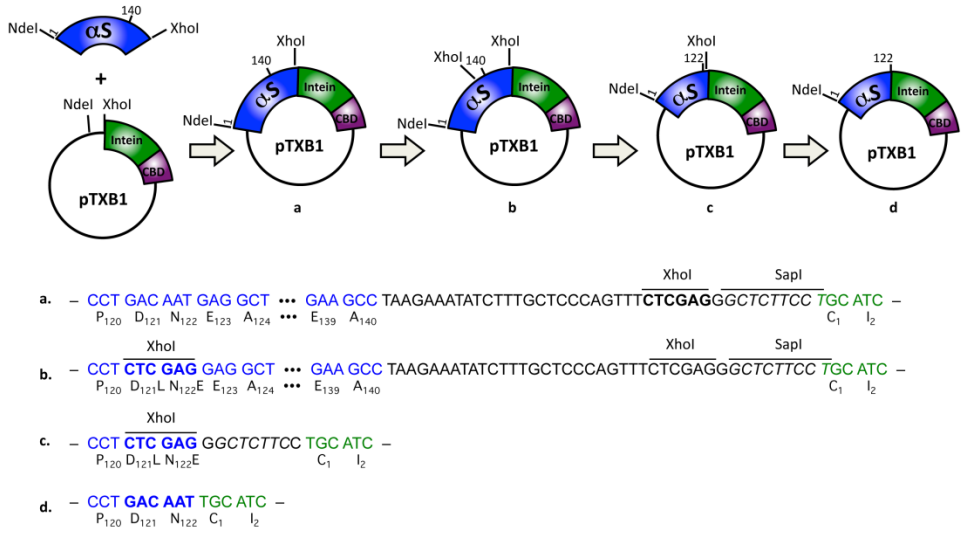

Figure 3.32 Construction of $\alpha S_{1\text{-}122}$ Intein Expression Plasmid with C-terminal Chitin Binding Domain.

To replace CBD with a polyhistidine tag at the C-terminus of the intein, two HindIII sites were introduced at the N- and C- termini of CBD using QuikChange® mutagenesis. The main portion of the CBD was removed by digestion with HindIII and ligation of the two sticky ends. Then, six His were inserted between the HindIII and stop codon. The plasmid sequence was confirmed by sequencing using T7-promotor as well as T7-terminator primers. Lastly, the amber stop codon (TAG) was introduced at position 39 using QuikChange® mutagenesis to generate $\alpha S^{F}_{1\text{-}122}TAG_{39}$-Int construct.

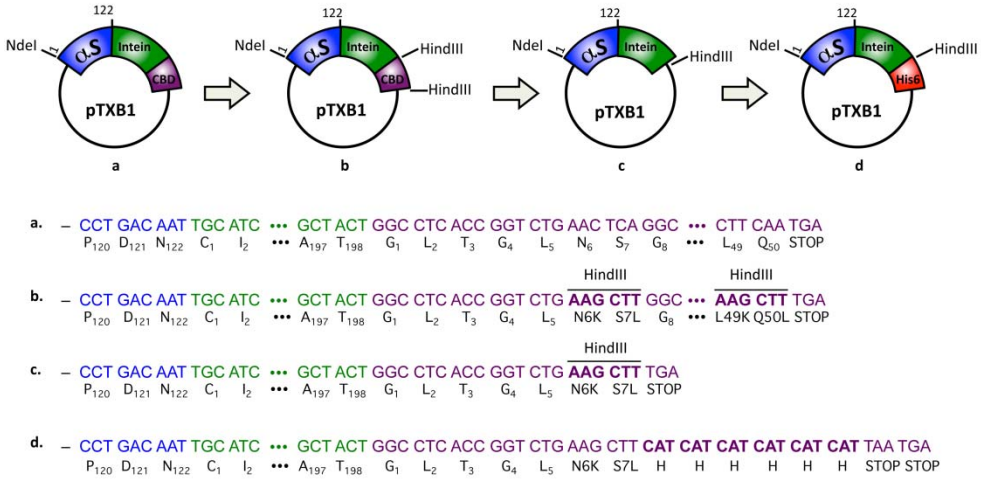

Figure 3.33 Construction of $\alpha S_{1\text{-}122}$ Intein Expression Plasmid with C-terminal His_6 Tag.

DNA Oligomers Used for αS QuikChange ® Mutagenesis

a. Mutation $Y_{39}F$
Forward: 5' – GGAAAGACAAAAGAGGGTGTTCTCTTTGTAGGCTCCAAA – 3'
Reverse: 5' – TTTGGAGCCTACAAAGAGAACACCCTCTTTTGTCTTTCC – 3'

b. Mutation $Y_{125}F$
Forward: 5' – GGATCCTGACAATGAGGCTTTTGAAATGCCTTCTGA – 3'
Reverse: 5' – TCAGAAGGCATTTCAAAAGCCTCATTGTCAGGATCC – 3'

c. Double Mutation $Y_{133}F\_Y_{136}F$
Forward: 5' – CCTTCTGAGGAAGGGTTTCAAGACTTCGAACCTGAAGCC – 3'
Reverse: 5' – GGCTTCAGGTTCGAAGTCTTGAAACCCTTCCTCAGAAGG – 3'

d. Mutation $Y_{39}TAG$
Forward: 5' – AAAAGAGGGTGTTCTCTAGGTAGGCTCCAAAACCAA – 3'
Reverse: 5' – CTTGGTTTTGGAGCCTACCTAGAGAACACCCTCTTTT – 3'

e. Mutation $F_{94}TAG$
Forward: 5' – GCATTGCAGCAGCCACTGGCTAGGTCAAAAAGGACCAGTTGGG – 3'
Reverse: 5' – CCCAACTGGTCCTTTTTGACCTAGCCAGTGGCTGCTGCAATGC – 3'

e. Deletion 2-8 and Mutation S_9C
Forward: 5' – AGAAGGAGATATACATATGTGCAAGGCCAAGGAGGG – 3'
Reverse: 5' – CTCCCTCCTTGGCCTTGCACATATGTATATCTCCTTCT – 3'

f. XhoI site introduction at D_{121}/N_{122}
Forward: 5' – GGAAGATATGCCTGTGGATCCTCTCGAGGAGGCTTATGAAATGCCTTCTG – 3'
Reverse: 5' – CAGAAGGCATTTCATAAGCCTCCTCGAGAGGATCCACAGGCATATCTTCC – 3'

g. Mutation $L_{121}D/E_{122}N$ and deletion of SapI site
Forward: 5' – GGAAGATATGCCTGTGGATCCTGACAATTGCATCACGGGAGATGCA– 3'
Reverse: 5' – TGCATCTCCCGTGATGCAATTGTCAGGATCCACAGGCATATCTTCC – 3'

h. Insertion of HindIII – 1st site
Forward: 5' – TACTGGCCTCACCGGTCTGAAGCTTGGCCTCACGACAAATCC – 3'
Reverse: 5' – GGATTTGTCGTGAGGCCAAGCTTCAGACCGGTGAGGCCAGTA – 3'

i. Insertion of HindIII – 2nd site
Forward: 5' – CGTTCCTGCCTTGTGGCAGAAGCTTTGACTGCAGGAAGGGGATCC – 3'
Reverse: 5' –GGATCCCCTTCCTGCAGTCAAAGCTTCTGCCACAAGGCAGGAACGT – 3'

j. Introduction of His Tag
Forward: 5' – CTCACCGGTCTGAAGCTTCATCATCATCATCATCATTAATGACTGCAGGAAGGG – 3'
Reverse: 5' –CCCTTCCTGCAGTCATTAATGATGATGATGATGATGAAGCTTCAGACCGGTGAG – 3'

Figure 3.34 DNA Oligomers Used for αS Mutagenesis.

Generation of pDule2-Cnf-BL21(DE3) *E. Coli* Cells for Unnatural Amino Acid Mutagenesis. A plasmid encoding for an orthogonal $tRNA_{CUA}$ and pCnf-tRNA synthetase for the incorporation of pCnf through amber stop codon suppression

(pDULE2-Cnf) was provided by Dr. Ryan Mehl (University of Oregon). *E. coli* BL21(DE3) cells were transformed with this plasmid and plated onto LB agar plates containing streptomycin. Following transformation, competent cells for protein expression were brewed in accordance with the Hanahan method. Cells stocks were stored at - 80 ºC in single-use aliquots.

Overexpression and Purification of Full-length αS Cnf Mutants. $\alpha S^{F}C_{9}F^{*}_{39}$, $\alpha S^{F}C_{9}F^{*}_{94}$, and $\alpha S^{F}F^{*}_{39}C_{123}$ were transformed into competent *E. coli* BL21(DE3) cells harboring the orthogonal $tRNA_{CUA}$ and pCnf-tRNA synthetase pair. Transformed cells were selected on the basis of ampicillin (Amp) and streptomycin (Strep) resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100 ug/mL) and Strep (100 ug/mL). The primary 5 mL culture was incubated at 37 ºC with shaking at 250 rpm for 5 hours. The primary culture was used to inoculate 1 L of a variant of M9 minimal media. To an autoclaved 1L solution containing 6 g Na_2HPO_4, 3 g KH_2PO_4, .5 g NaCl and 1 g NH_4Cl, the following autoclaved solutions were added: 1 mL of 2 M $MgSO_4$, 1 mL of 15 mg/mL $FeCl_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL $ZnCl_2$ (in acidified H_2O), 2 mL of 10% Bacto™ Yeast Extract 6.5 mL 40% glucose (w/v), and 1 μL of 1 M $CaCl_2$. When the OD_{600} of the secondary culture reached 0.8, Cnf was added (100 mg, 0.5 mM final concentration), and the culture was incubated overnight at 37 ºC with shaking at 250 rpm. The cells were harvested at 5000 x g for 15 min and the resulting pellet was resuspended in 20 mM Tris, pH 8. Following sonication, the cell lysate was boiled for 20 minutes prior to centrifugation for 20 minutes at 30,000 x g, 4 ºC. The cleared lysate was dialyzed against 20 mM Tris pH 8.0, loaded onto a Superdex 200 column (25 cm) connected to a BioCad Sprint (FPLC) system and eluted with 20 mM

Tris, pH 8. FPLC fractions post size-exclusion chromatography were analyzed by SDS PAGE. The fractions containing αS were then loaded onto a HighTrap Q HP column and eluted over a 100 minute sodium chloride gradient (0 to 0.5M NaCl in 20 mM Tris, pH 8). FPLC fractions post ion-exchange chromatography were dialyzed against 20 mM Tris, 100 mM NaCl, pH 7.4 and analyzed by SDS PAGE, MALDI, and fluorescence spectroscopy.

Mass Spectrometry and Fluorescence Spectrum of $\alpha S^{F}C_9F^{*}_{39}$

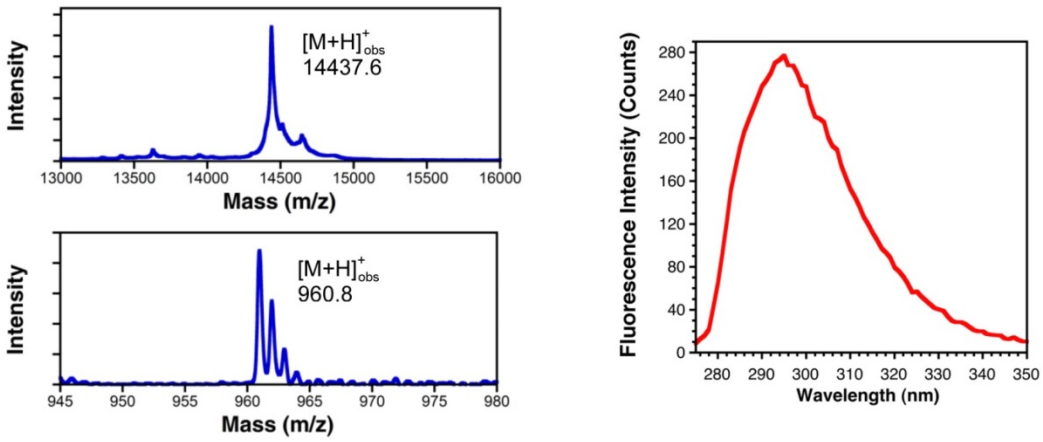

Figure 3.35 MALDI MS of $\alpha S^{F}C_9F^{*}_{39}$.
MALDI MS of $\alpha S^{F}C_9F^{*}_{39}$ (top left), trypsin digestion (top right), and primary fluorescence (bottom) of full-length $\alpha S^{F}C_9F^{*}_{39}$ .Calcd m/z of $\alpha S^{F}C_9F^{*}_{39}$:14438.1 Calcd m/z of trypsin fragment 35-43: 961.1

Mass Spectrometry and Fluorescence Spectrum of $\alpha S^{F}C_{9}F^{*}_{94}$

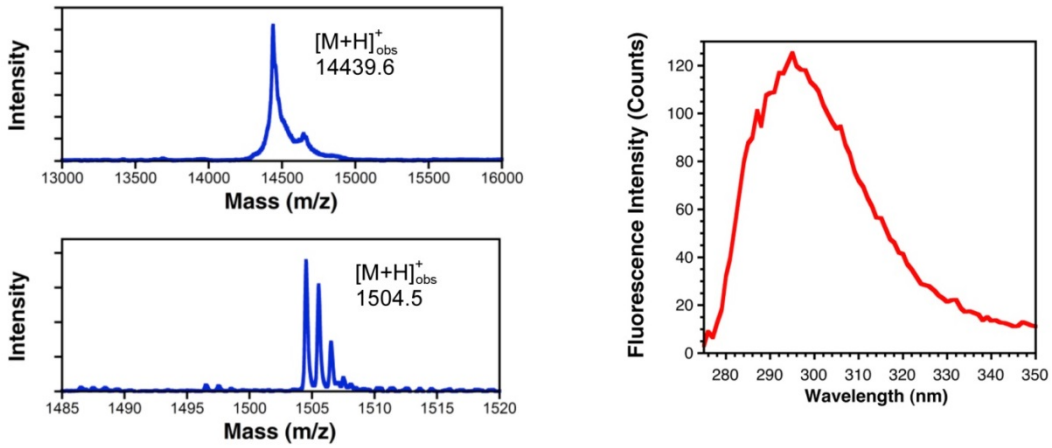

Figure 3.36 MALDI MS of $\alpha S^{F}C_{9}F^{*}_{94}$.
MALDI MS of $\alpha S^{F}C_{9}F^{*}_{94}$.(top left), trypsin digestion (top right), and primary fluorescence (bottom) of full-length $\alpha S^{F}C_{9}F^{*}_{94}$. Calcd m/z of $\alpha S^{F}C_{9}F^{*}_{94}$: 14438.1 Calcd m/z of trypsin fragment 81-96: 1503.8.

Mass Spectrometry and Fluorescence Spectrum of $\alpha S^{F}F^{*}_{39}C_{123}$

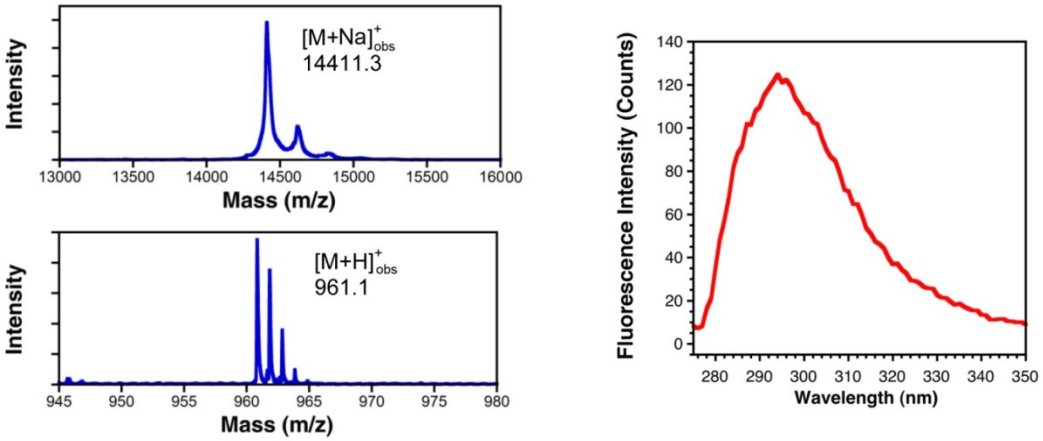

Figure 3.37 MALDI MS of $\alpha S^{F}F^{*}_{39}C_{123}$.
MALDI MS of $\alpha S^{F}F^{*}_{39}C_{123}$. (top left), trypsin digestion (top right), and primary fluorescence (bottom) of full-length $\alpha S^{F}F^{*}_{39}C_{123}$. Calcd m/z of $\alpha S^{F}F^{*}_{39}C_{123}$: 14409.1 Calcd m/z of trypsin fragment 35-43: 961.1

C-Terminal Fragment Expression (Factor Xa). The plasmid pET16b_H_{Tag}-$\alpha S^{F}_{\Delta 1\text{-}8}C_{9}F^{*}_{39}$ (see Supporting Information for plasmid construction) was transformed into *E. coli* BL21(DE3) cells containing pDULE2-Cnf (see above) and colonies were selected on the basis of Amp and Strep resistance. Following growth of a primary culture in LB

media, 1L of M9 minimal media containing glycerol as a carbon source (0.25% v/v) was inoculated and allowed to grow at 37 ºC with shaking at 250 RPM. When the OD_{600} reached 0.8, Cnf was added (150 mg), and the culture was allowed to incubate overnight at 37 ºC. The boiled and cleared lysate was prepared as described and incubated with 2 mL Ni-NTA resin for 1 hour at room temperature.[116] The resin was washed with 20 mL of 50 mM Tris, 150 mM NaCl, pH 8. The protein was eluted in eight 1.5 mL portions of 50 mM Tris, 150 mM NaCl, 250 mM imidazole, pH 8. Elution fractions were analyzed by SDS-PAGE gel, combined, and dialyzed against Factor Xa cleavage buffer (50 mM Tris, 150 mM NaCl, 5 mM $CaCl_2$, pH 8.0) overnight. Following dialysis, His-tag cleavage was achieved by overnight incubation with Factor Xa protease (10 units per 1 mg of H_{Tag}-$\alpha S^{F}_{\Delta 1\text{-}8}C_9F^*_{39}$). Factor Xa was inactivated by boiling the reaction for 20 minutes and the precipitated enzyme was removed by centrifugation for 20 minutes at 30,000 x g, 4 ºC. The supernatant was then subjected to incubation with 1 mL of Ni-NTA resin for 1 h at room temperature to capture any remaining H_{Tag}-$\alpha S_{9-140}{}^{F}C_9F^*_{39}$. The flow-through containing $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^*_{39}$ was collected, analyzed by SDS-PAGE gel and MALDI MS, dialyzed overnight against water, and stored at - 80 ºC.

Mass Spectrometry of $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^*_{39}$ Generated by Factor Xa Cleavage. Factor Xa-mediated cleavage of the N-terminal histidine tag followed by an IEGR recognition site was performed as described in the main text. Removal of the His tag was confirmed by MALDI MS.

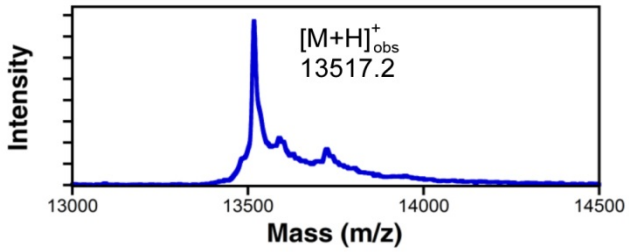

Figure 3.38 MALDI analysis of $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^{*}_{39}$ post Factor Xa cleavage.
Expected mass of $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^{*}_{39}$ post cleavage: 13516.0

C-Terminal Fragment Expression (Met Aminopeptidase). The plasmids pRK_$\alpha S^{F}_{\Delta 2\text{-}8}C_9F^{*}_{39}$ and pRK_$\alpha S^{F}_{\Delta 2\text{-}8}C_9F^{*}_{94}$ (see Supporting Information for plasmid construction) were transformed into competent *E. coli* BL21(DE3) cells with pDULE2-Cnf (see above). Transformed cells were selected on the basis of Amp and Strep resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100 ug/mL) and Strep (100 ug/mL). The primary 5 mL culture was incubated at 37 ºC with shaking at 250 RPM for 5 hours and then used to inoculate 1 L of LB media. When the OD_{600} of the secondary culture reached 0.8, the cells were isolated by centrifugation at 5000 x g for 15 minutes and gently re-suspended in 1 L of M9 minimal media containing Cnf (100 mg). M9 minimal media was prepared as described above. The cells were harvested at 5000 x g for 15 min and the resulting pellet was resuspended in 20 mM Tris pH 8. Following sonication, the cell lysate was boiled for 20 min prior to centrifugation for 20 min at 30,000 x g, 4 ºC. The cleared lysate was loaded onto a Superdex 200 column (25 cm) connected to a BioCad Sprint (FPLC) system and eluted with 20 mM Tris, pH 8. FPLC fractions from size-exclusion chromatography were analyzed by SDS PAGE. The fractions containing αS were then loaded onto a HighTrap™ Q HP column and eluted over a 100 min NaCl gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8). FPLC fractions from ion-exchange chromatography were dialyzed against 20 mM Tris, 100 mM NaCl,

pH 7.4 and analyzed by SDS PAGE gel, MALDI MS, UV/Vis, and fluorescence spectroscopy. Following dialysis, guanidinium hydrochloride and methoxylamine hydrochloride were added to a final concentration of 6 M and 400 mM, respectively. The pH was adjusted to 4.0 with NaOH and the solution was stirred overnight at 4 ºC. Removal of N-terminal cysteine adducts was monitored by MALDI MS. Following deprotection, the solution was dialyzed against water and dried in a vacuum centrifuge.

Mass Spectrometry of $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^{*}_{39}$ Generated by Met Aminopeptidase Cleavage. Deprotection of N-terminal cysteine adducts formed during protein expression was performed as described in the main text. Prior to deprotection, adducts corresponding to a mass increase of +26, +49, and +70 Da are observed. The reaction progress can be monitored by MALDI MS.

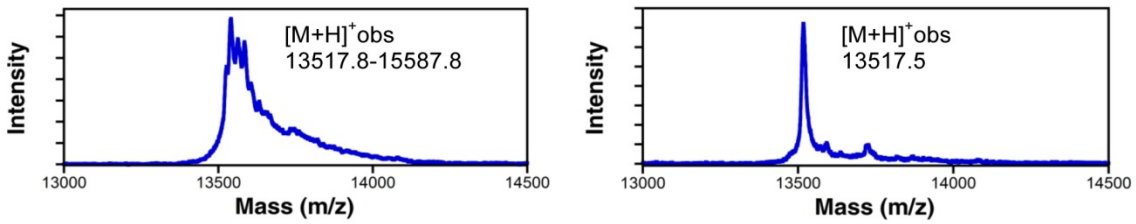

Figure 3.39 MALDI Analysis of Methoxyamine Deprotection.
MALDI analysis pre (left) and post (right) overnight incubation with $MeONH_2$•HCl in 6M Gdn•HCl, pH 4.0, at 4 ºC. Calcd m/z of $\alpha S^{F}_{\Delta 1\text{-}8}C_9F^{*}_{39}$: 13516.0

Thioester Synthesis (PyBOP). Peptide synthesis and selective cleavage from Cl-Trt resin was performed as previously described. Cleaved N-terminal peptide was dissolved in THF, and 2-mercaptopropionate (3 equiv) was added to the solution. After stirring for 5 min, DIPEA (6 equiv) and a solution of PyBOP pre-dissolved in warm THF (3 equiv) were added to the reaction mixture. The solution was allowed to stir for 30 min at room temperature, at which point the solvent was removed by rotary evaporation. Sidechain protecting groups were removed by incubating the peptide for 10 min with a cleavage cocktail containing TFA, TIPS, thioanisole, and 1,2-ethanedithiol (47:1:1:1). Following

concentration of the cleavage solution by rotary evaporation, the peptide was precipitated with ether and isolated by centrifugation. The peptide thioester was purified by HPLC and dried in a vacuum centrifuge.

Thioester Synthesis (C^bPG_o). Ac-MDV′$_3$FMKGLC^{b}PG$_o$ (αS$_{1\text{-}8}$V'$_3$-C^{b}PG$_o$,) was synthesized essentially as previously described, with the following modifications. Bromoacetic acid (0.0695 g, 500 μmol, 5 equiv) was pre-activated with diisopropyl carbodiimide (78 μL, 500 μmol, 5.0 equiv) for 30 min in dry DMF (6 mL). Also, the peptide (αS$_{1\text{-}8}$V'$_3$-C^{b}PG$_o$) was cleaved from the resin with the treatment of CH_2Cl_2/TFA/thioanisole (38:1:1 v/v, 5 mL) on a rotisserie for 1 h. The resulting solution was concentrated by rotatory evaporation, precipitated in cold diethyl ether (20 mL), centrifuged, and decanted ether. The precipitated peptide was dissolved in acetonitrile/water (50:50 v/v) for HPLC purification.

Ligation to Form αS^{F}F′$_4$C$_9$F*$_{39}$, αS^{F}V′$_3$C$_9$F*$_{39}$, and αS^{F}F′$_4$C$_9$F*$_{94}$. The expressed C-terminal protein fragment (1 equiv, 0.1 μmol) was dissolved in 100 μL of freshly prepared ligation buffer (6 M Gdn·HCl, 200 mM Na_2HPO_4, 20 mM TCEP, 1% v/v thiophenol, pH 7.2 for PyBOP-based thioesters or pH 8.0 for C^{b}PG$_o$ thioesters). The dissolved protein fragment was transferred to a microcentrifuge tube containing 1.2 equiv of the dried N-terminal peptide thioester, purged with argon, and allowed to incubate overnight at 37 ºC with shaking at 1000 RPM. The ligation solution was diluted to 3 mL and dialyzed against water prior to HPLC purification (see Supporting Information). Following HPLC purification, the product fractions were combined, concentrated, and exchanged into 20 mM Tris, 100 mM NaCl, pH 7.4 using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column.

HPLC Analysis of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$ and Ac-αS^{F}F′$_4$C$_9$F*$_{94}$ Ligations. Conditions for native chemical ligation reactions using mercaptopropionate thioester peptide formed by PyBOP activation were described in the main text. HPLC analyses of the ligation reactions were performed after 20 h.

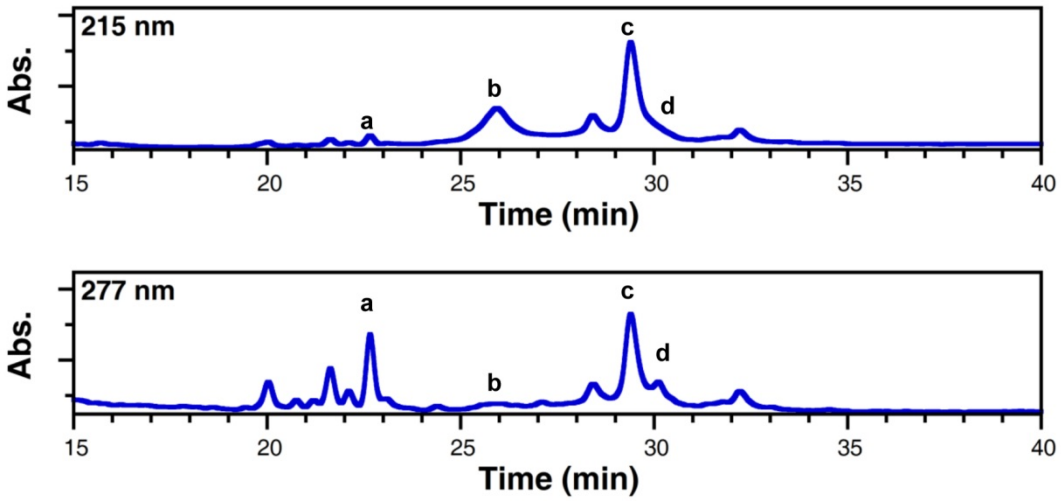

Figure 3.40 HPLC Analysis of Ligation of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$.
Peaks a, b, c, and d correspond to hydrolyzed Ac-αS$_{1-8}$F′$_4$ thioester, αS$^{F}_{\Delta 1-8}$C$_9$F*$_{39}$, ligation product Ac-αS^{F}F′$_4$C$_9$F*$_{39}$, and intact Ac-αS$_{1-8}$F′$_4$ thioester. Peaks earlier than a correspond to hydrolyzed peptide side products.

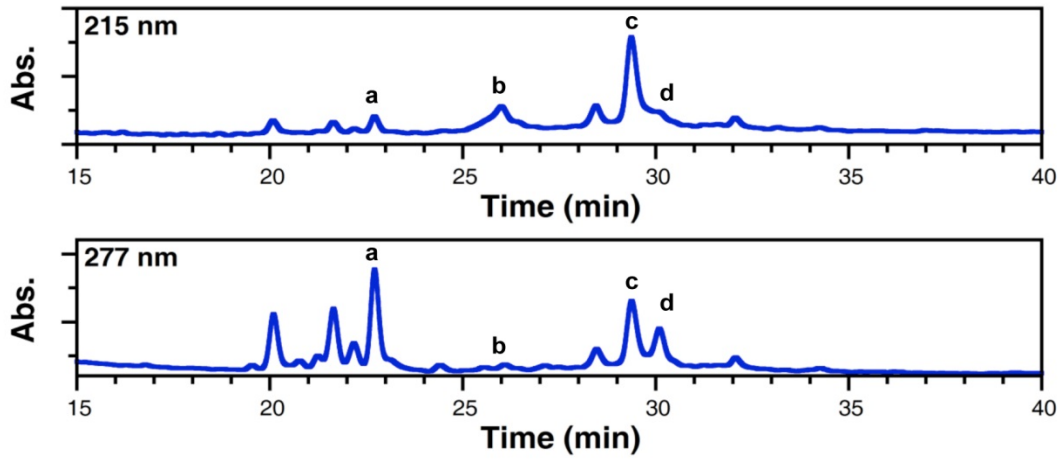

Figure 3.41 HPLC analysis of ligation of Ac-αS^{F}F′$_{4}$C$_{9}$F*$_{94}$.
Peaks a, b, c, and d correspond to hydrolyzed Ac-αS$_{1\text{-}8}$F′4 thioester, αS$^{F}_{\Delta 1\text{-}8}$C$_{9}$F*$_{94}$, ligation product Ac-αS^{F}F′$_{4}$C$_{9}$F*$_{94}$, and intact Ac-αS$_{1\text{-}8}$F′4 thioester. Peaks earlier than a correspond to hydrolyzed peptide side products.

Conditions for native chemical ligation reactions using peptide thioester formed by *in situ* deprotection and rearrangement of the C^{b}PG$_{o}$ linker are described in section 3.2. HPLC analyses of the ligation reactions were performed after 20 h. Due to low yields of peptide synthesis, ligation reactions were not fully optimized.

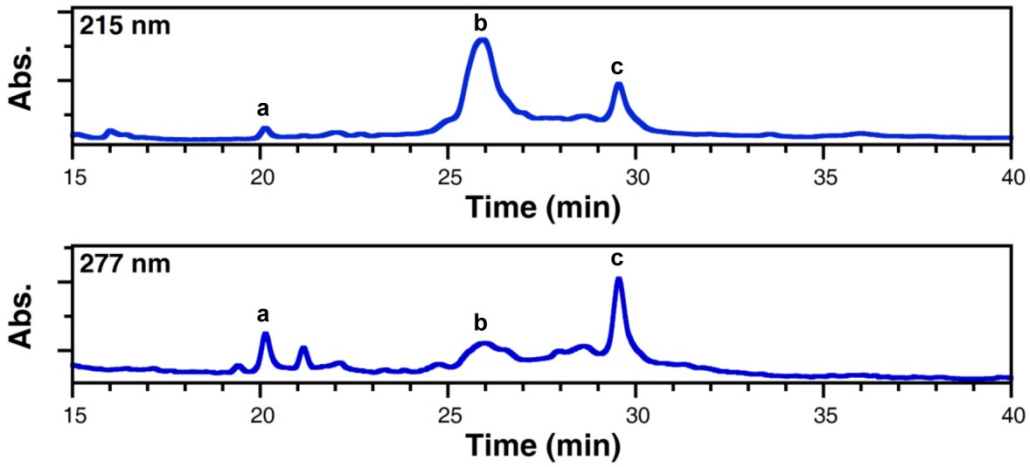

Figure 3.42 HPLC Analysis of Ligation of Ac-αS^{F}V′$_{3}$C$_{9}$F*$_{39}$.
Peaks a, b, and c, correspond to hydrolyzed Ac-αS$_{1\text{-}8}$V′$_{3}$ thioester, αS$^{F}_{\Delta 1\text{-}8}$C$_{9}$F*$_{39}$, and ligation product Ac-αS^{F}V′$_{3}$C$_{9}$F*$_{39}$. Peaks earlier than b correspond to hydrolyzed peptide side products.

Mass Spectrometry of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$.

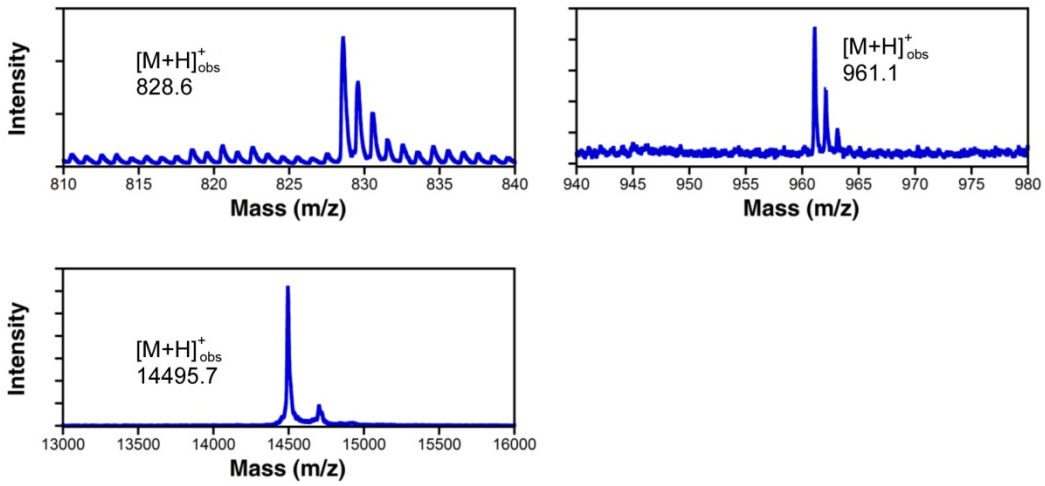

Figure 3.43 MALDI MS analysis of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$.
MALDI MS analysis of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$ (bottom) and trypsin fragments containing thioamide (top left) or Cnf (top right). Calcd m/z of thioamide-containing trypsin fragment 1-8: 828.4. Calcd m/z of Cnf-containing trypsin fragment: 961.1 Calcd m/z of of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$: 14494.7

UV Absorbance and Primary Fluorescence Spectra of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$.

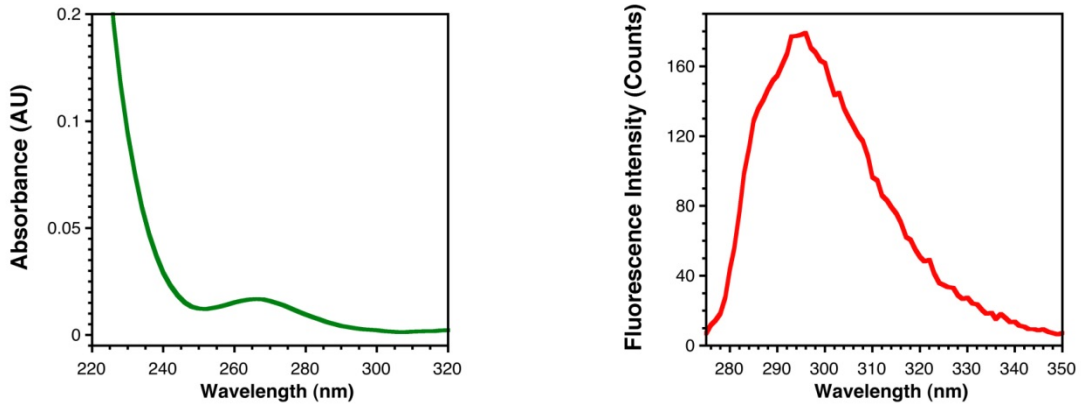

Figure 3.44 UV Asborbance and Fluorescence Spectrumof Ac-αS^{F}F′$_4$C$_9$F*$_{39}$.
UV absorbance (left) and fluorescence spectrum (right) of Ac-αS^{F}F′$_4$C$_9$F*$_{39}$. 1 μM samples were prepared in Tris buffer (20 mM Tris, 100 mM NaCl, 1 mM BME, pH 7.5).

Mass Spectrometry of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$.

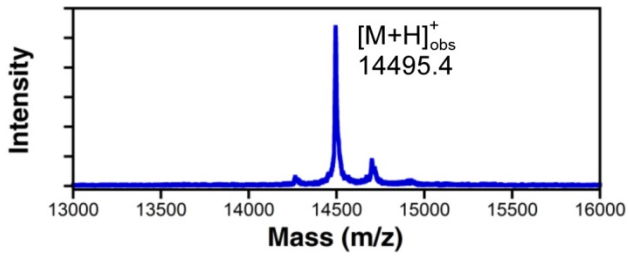

Figure 3.45 MALDI MS analysis of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$.
Calcd m/z of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$:14494.7.

Expression and purification of αS$^{F}_{1\text{-}122}$F*$_{39}$-Int. A plasmid encoding αS$_{1\text{-}122}$F*$_{39}$-Int-H$_{Tag}$ was transformed into *E. coli* BL21(DE3) cells containing the pDULE2-Cnf gene (prepared and made competent using the Hanahan method). A starter culture of 4 mL LB media was inoculated with a single colony and was grown at 37 °C in the presence of Amp (100 mg/L) and Strep (100 mg/L) for approximately 4 h. A secondary culture of 1 L M9 media was inoculated with 4 mL of the starter culture and grown at 37 °C with the same concentrations of Amp and Strep. M9 minimal media was prepared as described above. The culture was induced by the addition of Cnf (150 mg) followed by 0.5 mM of IPTG when it reached an OD$_{600}$ of 0.8 and allowed to grow at 37 °C overnight. Cells were harvested at 5, 000 RPM using a GS3 rotor and Sorvall RC-5 centrifuge for 20 min at 4 °C. After discarding the supernatant, the pellet was resuspended in 30 mL of resuspension buffer: 40 mM Tris, 5 mM EDTA, pH 8.3, protease inhibitor cocktail, 1 mM phenylmethanesulfonyl fluoride (PMSF), and 10 units/mL DNAse1–Grade II. The cells were lysed by sonication, and the lysate was centrifuged at 13, 900 RPM for 30 min at 4 °C. The supernatant was then incubated with Ni-NTA resin for 1 h on ice. This slurry was loaded into a column and the liquid allowed to flow though. The resin was washed extensively, first with 15 mL of Wash Buffer A (50 mM HEPES, pH 7.0), then twice

with 10 mL of Wash Buffer B (50 mM HEPES, 10 mM imidazole, pH 7.0. The protein was eluted with five 2 mL portions of elution buffer (50 mM HEPES, 300 mM imidazole, pH 7.0). The eluted fractions were combined and incubated with 400 mM sodium 2-mercaptoethanesulfonate (MESNA) for thiolysis for 48 h, stirring at 4 °C. The protein thioester was dialyzed against 20 mM Tris, pH 8.0 overnight, and incubated with Ni-NTA resin for 1 h on ice. The flow-through was collected, dialyzed against 20 mM sodium citrate, pH 5.0, purified over HiTrap SP column using a 100 min NaCl gradient (0 to 0.5 M NaCl in sodium citrate, pH 5). The product fractions were pooled and dialyzed against Milli-Q water overnight, dried in a vacuum centrifuge, and stored at - 80 °C until further use.

Mass Spectrometry of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$.

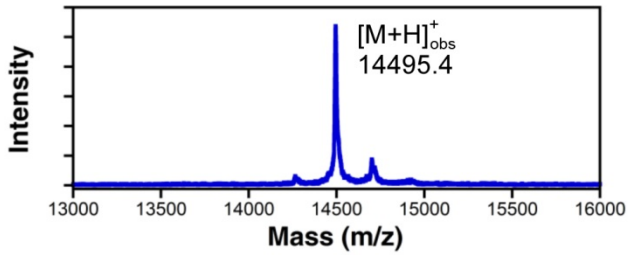

Figure 3.46 MALDI MS analysis of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$.
Calcd m/z of Ac-αS^{F}V′$_3$C$_9$F*$_{39}$:14494.7.

Ligation to Form αS^{F}F*$_{94}$C$_{123}$A′$_{124}$. Protein thioester αS$^{F}{}_{1\text{-}122}$F*$_{39}$ (1.5 equiv, 0.45 μmol, 1.5 mM) and C-terminal protein fragment αS$^{F}{}_{123\text{-}140}$C$_{123}$A′$_{124}$ (1 equiv, 0.3 μmol, 1 mM) were dissolved in 300 μL of a freshly made, degassed ligation buffer (6 M Gdn•HCl, 200 mM Na_2HPO_4, 20 mM TCEP, and 1% v/v thiophenol, pH 7.0). The reaction solution was placed in an incubator at 37 °C, shaking at 1000 RPM. The ligation progress was monitored by MALDI MS, and allowed to react until all the protein thioester was either consumed or hydrolyzed. Upon completion, the ligation solution was brought up to 1.5

mL with water and dialyzed against 20 mM Tris, pH 8.0 for 4 h. Then, the ligated product, $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$, was purified over a HiTrap™ Q HP column using a 100 min NaCl gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8). The product fractions were pooled, concentrated, and the buffer was exchanged for 20 mM Tris, 100 mM NaCl, pH 7.4 using an Amicon (Millipore) Ultra 0.5 mL 10 kDa spin column.

HPLC Analysis of $\alpha S^{F}F^{*}_{39}$ $C_{123}A'_{124}$.

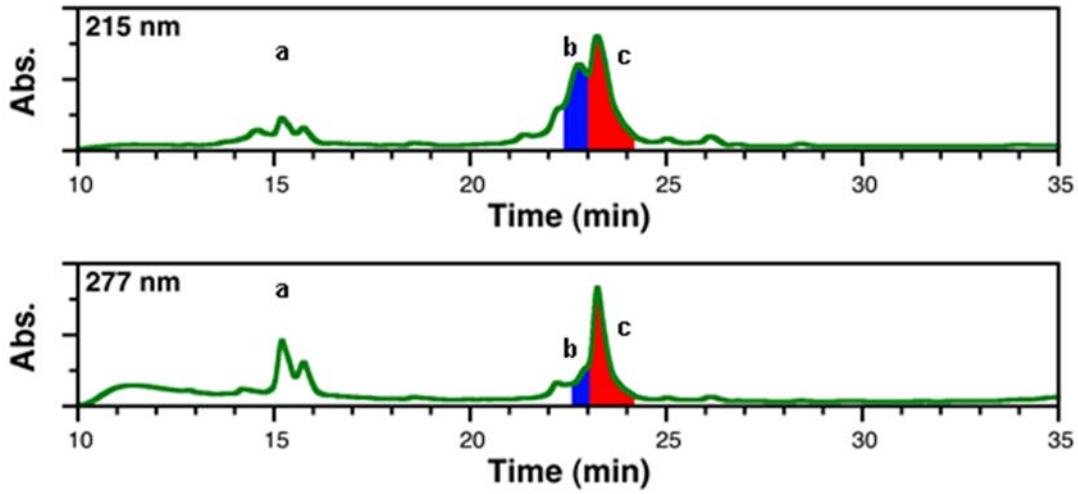

Figure 3.47 HPLC Analysis of $\alpha S^{F}F^{*}_{39}$ $C_{123}A'_{124}$.
Peaks under bracket **a** correspond to thioamide peptide and thioamide peptide side products. Unligated $\alpha S_{1\text{-}122}F^{*}_{39}$ (**b**) and ligated product $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$ (**c**) are represented by the blue and red shaded regions, respectively.

FPLC Analysis of $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$.

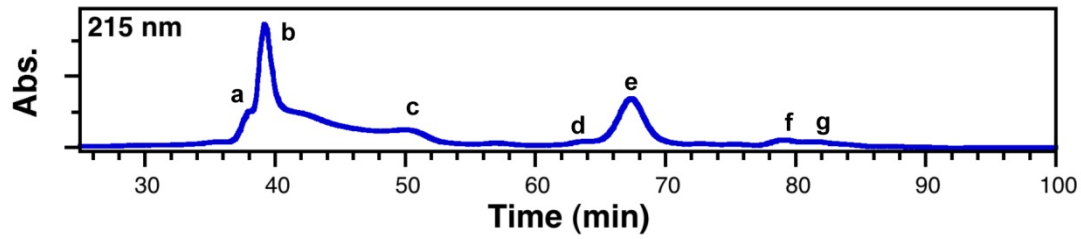

Figure 3.48 FPLC analysis of $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$.
Peaks a, b, and c correspond to hydrolyzed $\alpha S_{1\text{-}122}F^{*}_{39}SR$ ($\alpha S_{1\text{-}122}F^{*}_{39}OH$), TCEP, and protein impurities in the 8-9 kDa range. Peaks d, e, and f, and g correspond to thioamide peptide side products, $\alpha S^{F}F^{*}_{39}C_{123}A'_{124}$, thioamide peptide-ligated protein dimer, and thioamide peptide dimer.

Absorbance Measurements of Thioacetamide in Trimethylamine Oxide. The absorbance of thioacetamide (approximately 35 µM) was measured in Tris buffer (20 mM Tris, 100 mM NaCl, 1 mM BME, pH 7.5) and in Tris buffer containing 4M trimethylamine oxide. This verified that TMAO did not alter the ability of the thioamide to participate in FRET.

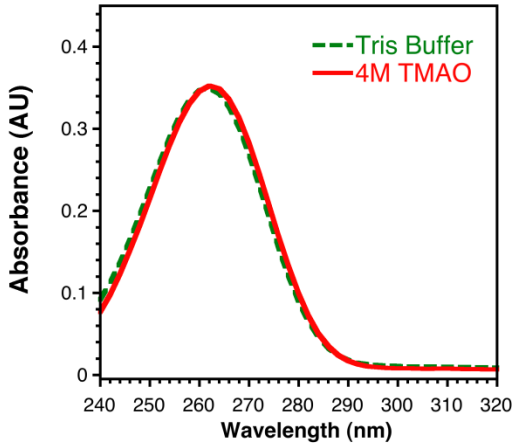

Figure 3.49 Thioacetamide Absorbance in Tris Buffer Compared to Thioacetamide Absorbance in Tris buffer Containing 4M TMAO.

Fluorescence Measurements of Cnf in Trimethylamine Oxide. The fluorescence of 1 µM Cnf amino acid was measured in 20 mM Tris, 100 mM NaCl, pH 7.5 containing 0, 1, 2, 3, or 4M TMAO. A decrease in fluorescence is observed at higher TMAO concentrations.

'

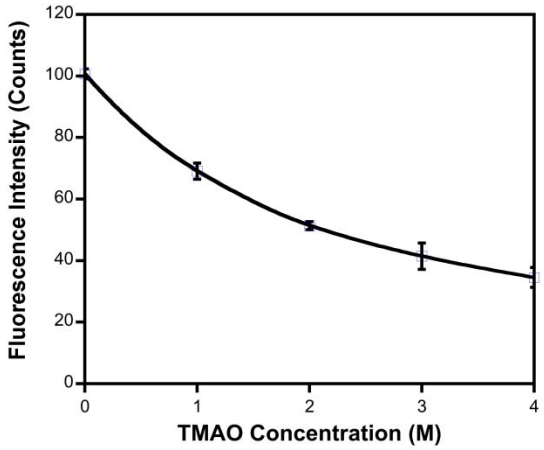

Figure 3.50 Fluorescence Intensity of Cnf In Increasing Concentrations of TMAO.

Fluorescence Measurements of αS Mutants in Trimethylamine Oxide and Urea. Prior to performing the experiment, the concentrations of all double-labeled proteins were adjusted using the BCA assay to match their respective oxoamide (donor only) control proteins. Tris buffers (20 mM Tris, 100 mM NaCl, 1 mM βME, pH 7.5) containing trimethylamine oxide (TMAO) were prepared such that upon addition of protein (to a final concentration of 1 μM), the final TMAO concentrations were 0 M, 1 M, 2 M, 3 M, and 4 M. Buffers containing urea were similarly prepared such that upon addition of protein, the final urea concentration was 6 M. Fluorescence spectra were measured immediately following sample dilution in TMAO or urea buffer (to a final volume of 130 μL). All samples were prepared and measured in triplicate. The background fluorescence of each buffer was independently measured and subtracted from the protein fluorescence spectrum.

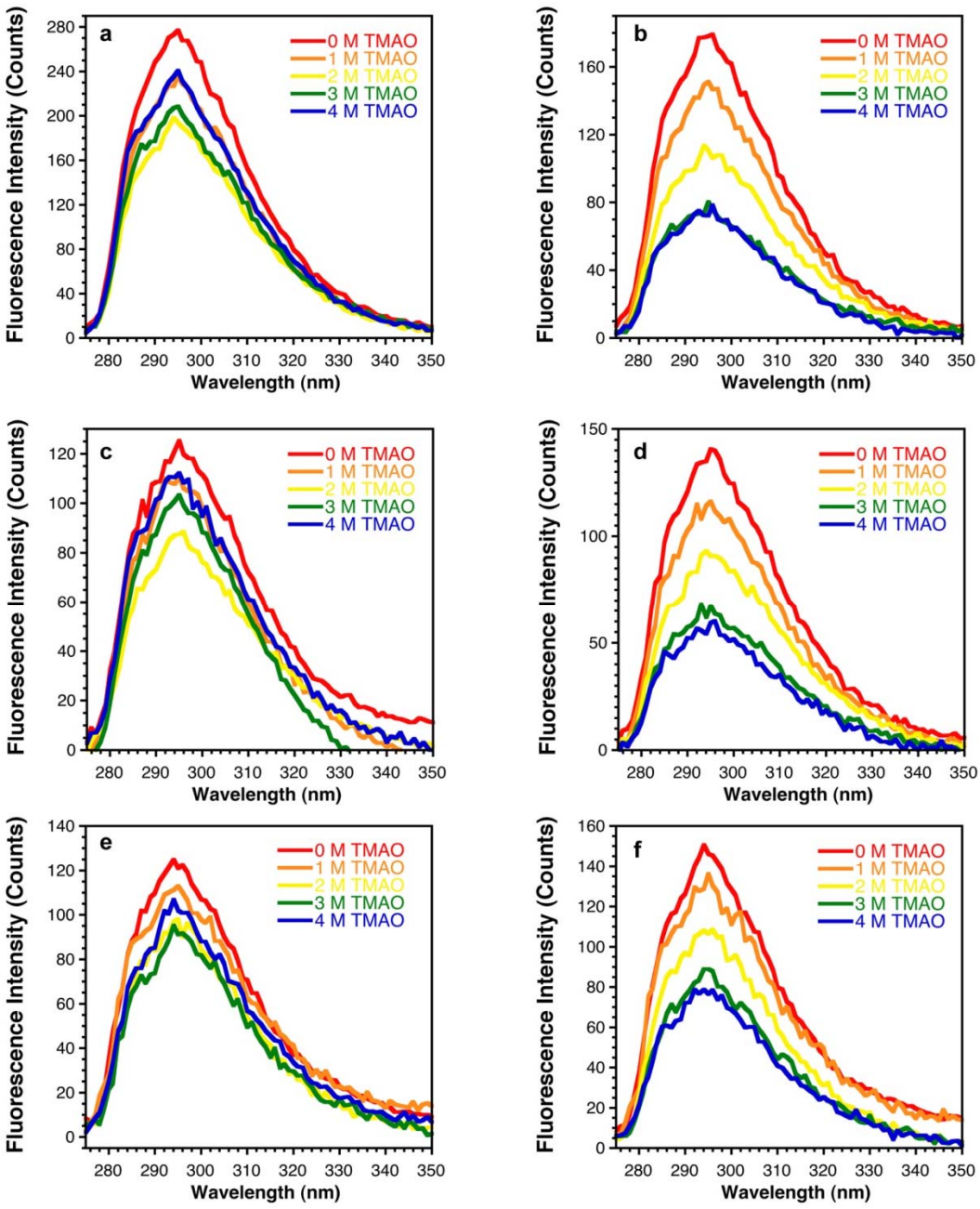

Figure 3.51 Fluorescence Scans of Oxoamide and Thioamide-containing αS mutants in Increasing Concentrations of TMAO.

$\alpha S^{F}C_{9}F^{*}_{39}$ (a), Ac-$\alpha S^{F}F'_{4}C_{9}F^{*}_{39}$ (b), $\alpha S^{F}C_{9}F^{*}_{94}$ (c), Ac-$\alpha S^{F}F'_{4}C_{9}F^{*}_{94}$ (d), $\alpha S^{F}F^{*}_{39}$ C_{123} (e), $\alpha S^{F}F^{*}_{39}$ $C_{123}A'_{124}$ (f).

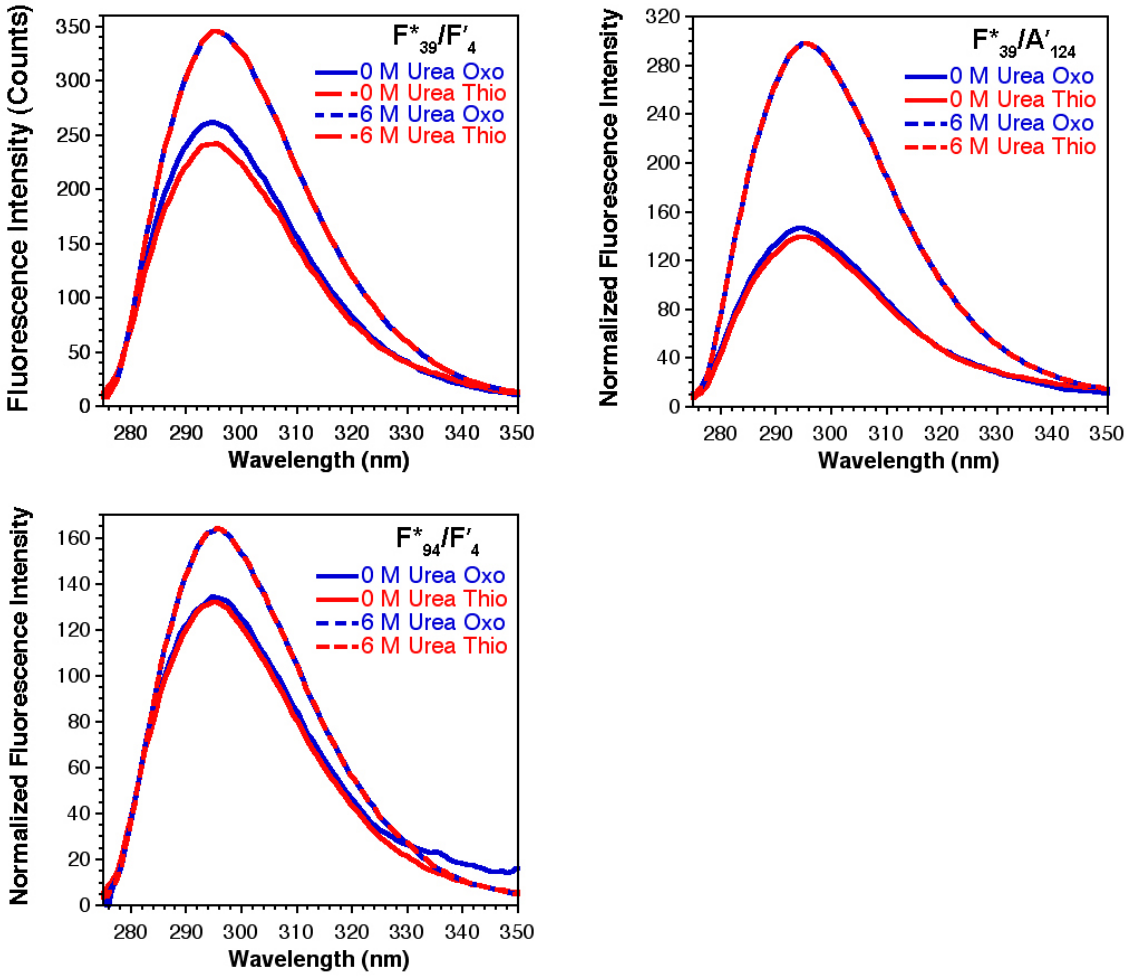

Figure 3.52 Normalized Fluorescence Scans of Oxoamide and Thioamide-containing αS Mutants in 0 and 6 M Urea.
Top: $\alpha S^{F}C_9F^*_{39}$ and Ac-$\alpha S^{F}F'_4C_9F^*_{39}$, Middle: $\alpha S^{F}C_9F^*_{94}$ and Ac-$\alpha S^{F}F'_4C_9F^*_{94}$, Bottom: $\alpha S^{F}F^*_{39}C_{123}$ and $\alpha S^{F}F^*_{39}C_{123}A'_{124}$.

Experimental Förster Distance Calculations (αS). The Förster distance, R_0, was calculated as described above for CaM. Substituting $J = 7.0\times10^{12} \pm 2\times10^{11}$ $M^{-1}\bullet cm^{-1}\bullet nm^4$ into equation (3.1), as well as 0.11 for the quantum yield of Cnf, 1.33 for the index of refraction of water, and 2/3 for κ^2 gives R_0 = 15.6 Å for $\alpha S^{F}C_9F^*_{39}$ constructs in 0 M TMAO. For $\alpha S^{F}C_9F^*_{94}$ constructs, we use 0.049 as a quantum yield for Cnf, and calculate R_0 = 13.6 Å. For $\alpha S^{F}F^*_{39}C_{123}$ constructs, we use 0.045 as a quantum yield for Cnf, and calculate R_0 = 13.5 Å. In all cases data in 6 M urea were used to normalize oxo-

and thio- fluorescence data to ensure that small deviations in concentration did not alter data interpretation. Quantum yields for each mutant in different concentrations of TMAO were determined by comparing the fluorescence output of the mutant relative to the fluorescence of that mutant in 0 M TMAO using identical concentrations and fluorometer settings. Comparison of $\alpha S^{F}C_{9}F^{*}_{39}$ fluorescence to free Cnf showed that no substantial quenching occurred in this mutant in the absence of the thioamide (data not shown). As for CaM, these values of R_0 were then used in determining experimental chromophore separations (R_{FRET}) for each set of Cnf/thioamide probe locations according to equation 3.5 where E_Q is defined by equation 3.6 as described in the main text. These calculations are summarized in Table 3.9.

Table 3.9 R_{FRET} Determination in α-Synuclein Refolding Assay.

[TMAO]	E_Q	SE	E_Q	SE	E_Q	SE
0	0.13	0.02	-0.02	0.02	0.03	0.06
1	0.14	0.01	0.07	0.03	0.02	0.02
2	0.24	0.02	0.08	0.03	0.07	0.03
3	0.48	0.03	0.31	0.03	0.24	0.02
4	0.59	0.03	0.59	0.07	0.40	0.05
[TMAO]	R_{FRET}	SE	R_{FRET}	SE	R_{FRET}	SE
0	21.3	0.7				
1	20.5	0.4	20.8	1.7	24.5	4.2
2	17.9	0.4	19.2	1.5	20.1	1.4
3	15.2	0.3	14.8	0.4	15.8	0.3
4	14.4	0.3	12.6	0.6	14.1	0.5
[TMAO]	F_{Oxo}	R_0	F_{Oxo}	R_0	F_{Oxo}	R_0
0	282.2	15.6	125.3	13.6	115.9	13.5
1	243.9	15.2	110.8	13.4	106.2	13.3
2	204.7	14.8	87.4	12.8	90.2	12.9
3	220.8	15.0	93.4	13.0	93.2	13.0
4	257.5	15.4	112.3	13.4	102.0	13.2

Parameters calculated as described in text. All distances in Å.

Synthesis and Purification of Thioaspartate Precursor

***α*-*N*-Fmoc-L-aspartate-2-amino-5-nitroanilide.** Fmoc-Asp(O*t*-Bu)-OH (2.67 g, 6.50 mmol) was dissolved in 35 mL of tetrahydrofuran (THF) under argon flow and the solution was cooled to - 10 °C in a 1:3 NaCl/ice bath. *N*-methylmorpholine (NMM, 0.72 mL, 6.5 mmol) and isobutyl chloroformate (IBCF, 0.85 mL, 6.5 mmol) were added dropwise with stirring. After 15 min, 4-nitro-*o*-phenylenediamine (1.0 g, 6.5 mmol) was added and the reaction was allowed to proceed with stirring under argon flow at - 10 °C for 2 h. The reaction was then allowed to proceed for an additional six hours with stirring at room temperature. The reaction mixture was dried by rotary evaporation, resuspended in 20 mL of DMF, and then poured into 200 mL of a saturated potassium chloride solution. The precipitated product was filtered and washed with cold water. The precipitate was then dissolved in minimal ethyl acetate (EtOAc) and purified over a silica gel column in 3:2 hexanes/EtOAc to afford the compound as a yellow solid in 88.9% yield. R_f = 0.5 in 1:1 hexanes/EtOAc; ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.07 (br. s, 2H), 7.95 (dd, *J* = 2.5, 9.0 Hz, 1H), 7.76 (dd, *J* = 3.6, 7.5 Hz, 2H), 7.57 (dd, *J* = 3.4, 7.4 Hz, 2H), 7.41-7.36 (m, 2H), 7.31-7.28 (m, 2H), 6.67 (d, *J* = 9.0 Hz, 1H), 5.94 (d, *J* = 8.2 Hz, 2H), 4.65 (br. s, 1H), 4.55-4.45 (m, 2H), 4.21 (t, *J* = 6.3 Hz, 2H), 2.99 (dd, *J* = 4.6, 17.2 Hz, 1H), 2.77 (d, *J* = 12.8 Hz, 1H) 1.46 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 172.5 (no peak), 170.8 (no peak), 157.3 (no peak), 149.0 (no peak), 144.4 (no peak), 142.3 (no peak), 139.6 (no peak), 128.8 (+), 128.2 (+), 125.9 (+), 125.4 (+), 124.5(+), 121.7 (no peak), 121.1 (+), 115.9 (+), 83.6, 68.3 (-), 52.7 (+), 48.2 (+), 38.3 (-), 29.1 (+); HRMS (ESI) m/z calcd for $C_{29}H_{30}N_4O_7$ $[M + H]^+$ 547.219, found 547.218.

α-*N*-Fmoc-L-thioaspartate-2-amino-5-nitroanilide. P_4S_{10} (2.47 g, 5.56 mmol) and anhydrous Na_2CO_3 (0.589 g, 5.56 mmol) were stirred in 30 mL of THF at room temperature under argon flow until a clear yellow solution was obtained. After cooling the solution to 0 ºC on ice, α-*N*-Fmoc-L-aspartate-2-amino-5-nitroanilide (3.04 g, 5.56 mmol) was added, and the reaction was carefully monitored by TLC. After approximately 1 h, the reaction was filtered through Celite® (Sigma-Aldrich) and dried by rotary evaporation. The crude reaction material was dissolved in EtOAc and purified over a silica gel column in 1:1 hexanes/EtOAc to afford α-*N*-Fmoc-L-thioaspartate-2-amino-5-nitroanilide as a yellow foam (2.46 g, 78.6% yield). R_f = 0.7 in 1:1 hexanes/EtOAc; ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.84 (br. s, 1H), 8.07 (br. s, 1H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.76 (d, *J* = 6.9, 2H), 7.52 (dd, *J* = 7.3, 19.6 Hz, 2H) 7.43-7.39 (m, 2H), 7.32-7.28 (m, 2H), 6.56 (d, *J* = 9.0 Hz, 1H), 6.06 (d, *J* = 8.2 Hz, 1H), 5.07 (br. s, 1H), 4.84 (br. s, 2H), 4.37 (br. s, 2H), 4.21-4.13 (m, 1H), 3.18-3.05 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 204.9 (no peak), 172.1 (no peak), 157.2 (no peak), 149.6 (no peak), 144.3 (no peak), 142.2 (no peak), 139.0 (no peak), 128.8 (+), 128.1 (+), 126.5 (+), 125.9 (+), 125.8 (+), 122.8 (no peak), 121.0 (+), 115.6 (+), 83.5, 68.4 (-), 58.4 (+), 47.8 (+), 41.6 (-), 28.9 (+); HRMS (ESI) m/z calcd for $C_{29}H_{30}N_4O_6S_1$ $[M + H]^+$ 563.196, found 563.197.

α-*N*-Fmoc-L-thioaspartate-nitrobenzotriazolide. α-*N*-Fmoc-L-thioaspartate-2-amino-5-nitroanilide (1.00 g, 1.78 mmol) was added to glacial acetic acid diluted with 5% H_2O (25 mL). $NaNO_2$ (0.16 g, 2.23 mmol) was added in small portions over 5 min with constant stirring at room temperature. After 30 min, the reaction was quenched by the addition of 500 mL ice water. The resulting pale orange precipitate was filtered, washed extensively with ice water, and allowed to dry under vacuum. The final product was

characterized and used in peptide synthesis without any further purification. R_f = 0.9 in 1:1 hexanes/EtOAc; ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.64 (s, 1H), 8.46 (d, *J* = 8.6 Hz 1H), 8.33 (d, *J* = 8.7 Hz, 1H), 7.78-7.77 (m, 2H), 7.64-7.59 (m, 2H) 7.44-7.38 (m, 2H), 7.35-7.29 (m, 2H), 6.52-6.45 (m, 1H), 6.23 (d, *J* = 8.7 Hz, 1H), 4.54-4.48 (m, 1H), 4.41-4.34 (m, 1H), 4.26-4.21 (m, 1H), 3.15-3.07 (m, 1H), 2.97-2.86 (m, 1H), 1.42 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 206.7 (no peak), 169.3 (no peak), 156.4 (no peak), 150.6 (no peak), 149.9 (no peak), 144.6 (no peak), 142.3 (no peak), 132.8 (no peak), 128.7 (+), 128.0 (+), 126.0 (+), 123.2 (+), 122.6 (+), 121.0 (+), 113.6 (+), 83.4, 68.2 (-), 59.0 (+), 48.1 (+), 41.9 (-), 28.9 (+); HRMS (ESI) m/z calcd for $C_{29}H_{30}N_4O_6S_1$ $[M + Na]^+$ 596.158, found 596.158.

Peptide Synthesis and Purification. Peptide synthesis and selective cleavage from Cl-Trt resin was performed as previously described. 50 μmol of cleaved N-terminal peptide (Ac-MetAsp'(O*t*-BuValPhe-OH) was dissolved in THF, and methyl 3-mercaptopropionate (16.6 μL, 150 μmol) was added to the solution. After stirring for 5 min, *N*,*N*-diisoproplethylamine (DIPEA, 26.2 μL, 300 μmol) and a solution of PyBOP, 78 mg, 150 μmol) dissolved in minimal DMF were added to the reaction mixture. The solution was allowed to stir for 25 min at room temperature, at which point the solvent was reduced to less than 1 mL by rotary evaporation. Next, the peptide was precipitated with water and isolated by centrifugation. The aspartate protecting group was removed by incubating the peptide for 10 min with a cleavage cocktail containing trifluoroacetic acid (TFA), triisopropylsilane, thioanisole, and 1,2-ethanedithiol (47:1:1:1). Following concentration of the cleavage solution by rotary evaporation, the peptide was dissolved in a solution of water and acetonitrile (70:30) and purified by HPLC using a binary system

of Buffer A (0.1% TFA/H_2O) and Buffer B (0.1% TFA/CH_3CN). Solvent gradients, column descriptions, and retention times are listed in Tables 3.10 and 3.12. MALDI MS calcd m/z $(M + H)^+$ 671.2, found 671.2.

Table 3.10 Summary of HPLC Peptide Purification.

Peptide	Gradient Description†	Retention Time (min)	Column (Size: Brand)
Ac-MetAsp'ValPhe-SR	S7	27.5	Semi-prep: Vydac 218 TP C18

† Described in Table S2.

Table 3.11 Solvent Gradient Used for Peptide Purification.

Gradient	Time (min)	Buffer A (%)
S7	0:00	98
	5:00	98
	11:00	66
	40:00	62
	43:00	0
	45:00	0
	50:00	98

Construction of Recombinant pET16b αS Expression Plasmid. The generation of pET16b-His_{Tag}-$αS^{F}_{9-140}C_9TAG_{39}$, constructed for cysteine-mediated NCL at position 9, has been described previously. In order to generate His_{Tag}-$αS^{F}_{6-140}TAG_{39}$, a primer was designed to insert base pairs corresponding to residues K, G, and L with simultaneous mutation of C_9 back to S (Figure 3.54).

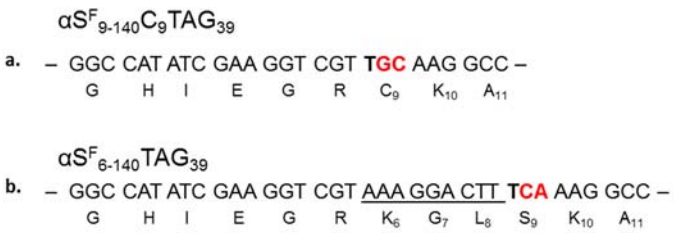

Figure 3.53 Construction of His_{Tag}-$αS^{F}_{6-140}TAG_{39}$.
A primer was designed to insert nucleotides corresponding to amino acid residues 6-8 (underlined) and mutate C_9 to S (bolded) with a single round of Quikchange® mutagenesis.

DNA Oligomers Used for αS QuikChange® Mutagenesis

Forward: 5' – GGCCATATCGAAGGTCGTAAAGGACTTTCAAAGGCCAAGG – 3'
Reverse: 5' – CCTTGGCCTTTGAAAGTCCTTTACGACCTTCGATATGGCC – 3'

Figure 3.54 Sequence of Forward and Reverse Primers Used in Mutagenesis.

Overexpression and Purification of His_{Tag}-$αS^{F}_{6-140}Cnf_{39}$. pET16b-His_{Tag}-$αS^{F}_{6-140}TAG_{39}$ was transformed into competent *E. coli* BL21(DE3) cells harboring the orthogonal $tRNA_{CUA}$ and CnfRS pair. Transformed cells were selected on the basis of ampicillin (Amp) and streptomycin (Strep) resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100 μg/mL) and Strep (100 μg/mL). The primary 5 mL culture was incubated at 37 ºC with shaking at 250 rpm for 5 h. The primary culture was used to inoculate 1 L of a variant of M9 minimal media described previously. When the OD_{600} of the secondary culture reached 0.8, Cnf was added (150 mg, 0.8 mM final concentration), and the culture was incubated overnight at 37 ºC with shaking at 250 rpm. The cells were harvested at 5000 x g for 15 min and the resulting pellet was re-suspended in Factor Xa cleavage buffer (20 mM Tris, 100 mM NaCl, and 1 mM $CaCl_2$, pH 8.0). Following sonication, the cell lysate was boiled for 20 min prior to centrifugation for 20 min at 30,000 x g, 4 ºC. The cleared lysate was incubated with 2 mL of Ni-NTA resin for 1 h at room temperature and then applied to an empty PD-10 column. The resin was washed with 20 mL of Factor Xa buffer followed by a 20 mL wash with Factor Xa buffer containing 30 mM imidazole. The protein was eluted from the column in eight 1.5 mL portions of Factor Xa buffer containing 250 mM imidazole. Elution fractions were analyzed by SDS-PAGE, combined, and dialyzed

against Factor Xa cleavage buffer overnight at 4 ºC. Protein expression yields ranged from 4 to 6 mg per L of culture.

Factor Xa Cleavage and Purification of His_{Tag}-$\alpha S^{F}_{6\text{-}140}Cnf_{39}$. Dialyzed His_{Tag}-$\alpha S^{F}_{6\text{-}140}Cnf_{39}$ was diluted to 0.5 mg/mL in Factor Xa cleavage buffer. Removal of the His_{11} tag was achieved by incubating His_{Tag}-$\alpha S^{F}_{6\text{-}140}Cnf_{39}$ with 8 µL of Factor Xa (1.05 mg/L). The reaction progress was monitored by MALDI MS. When full-length protein was no longer detected by MALDI MS, the protein was dialyzed into 20 mM Tris, pH 8.0 for 2 h. The protein was then loaded onto a HighTrap Q HP (GE Healthcare) column and eluted over a 120 min sodium chloride gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8.0). FPLC fractions from ion-exchange chromatography were analyzed by SDS PAGE and MALDI MS and then immediately exchanged into AaT ligation buffer (50 mM HEPES pH 8.0, 150 mM KCl, 10 mM $MgCl_2$) using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column.

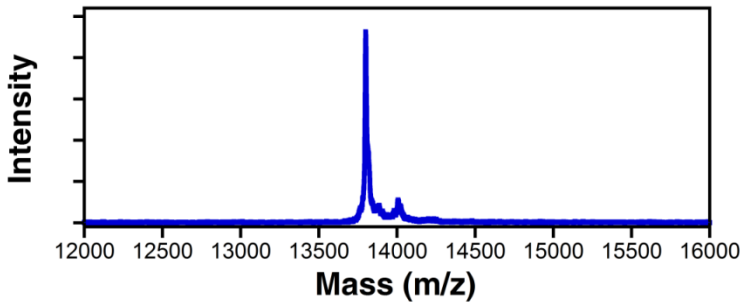

Figure 3.55 MALDI MS of $\alpha S^{F}_{6\text{-}140}Cnf_{39}$.
Calcd m/z $(M + H)^+$ 13799, found 13800.

Enzymatic Transfer of Hcm to $\alpha S^{F}_{6\text{-}140}Cnf_{39}$. Aminoacyl transferase (AaT), $L_{13}G$ methionyl tRNA synthetase (Met*RS), and Hcm were prepared as previously described. $\alpha S^{F}_{6\text{-}140}Cnf_{39}$ (0.5 mg/mL) was incubated with Hcm (1 mM), 20 mg *E. coli* total tRNA, ATP (2.5 mM) and 0.1 mg each *E. coli* Met*RS and AaT for 2 h at 37 ºC. Following the

incubation period, the reaction buffer was exchanged (20 mM Tris, pH 8.0), and the protein was loaded onto a HighTrap Q HP column, and eluted over a 120 min sodium chloride gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8.0). Prior to native chemical ligation, the buffer was exchanged against MilliQ H_2O and the protein was dried in a vacuum centrifuge.

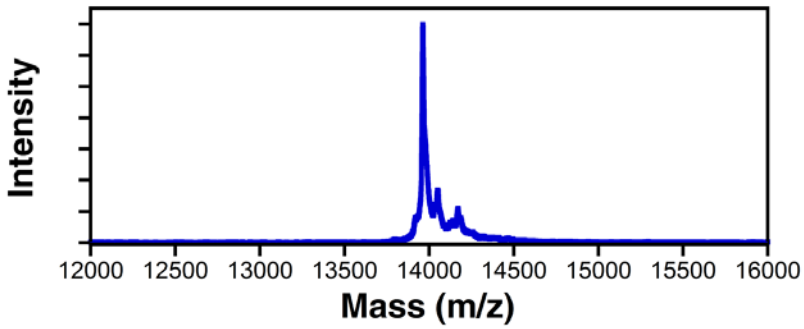

Figure 3.56 MALDI MS of $\alpha S^{F}_{5\text{-}140}Hcm_5Cnf_{39}$.
Calcd m/z $(M + H)^+$ 13963, found 13962.

Ligation of Ac-MetAspValPhe-SR to $\alpha S^{F}_{5\text{-}140}Hcm_5Cnf_{39}$ to form Ac-$\alpha S^{F}Asp'_2Hcs_5Cnf_{39}$. $\alpha S^{F}_{5\text{-}140}Hcm_5Cnf_{39}$ (1 equiv, 0.072 µmol) was dissolved in 200 µL of degassed ligation buffer (6 M guanidinium hydrochloride, 200 mM sodium phosphate, 20 mM tris(2-carboxyethyl)phosphine (TCEP), 1% v/v thiophenol, pH 7.5) and incubated for 5 min. The reduced protein fragment was transferred to a microcentrifuge tube containing 1.5 equiv of the dried N-terminal peptide thioester, purged with argon, and allowed to incubate for 24 h at 37 ºC with shaking at 600 rpm. The ligation solution was then thoroughly exchanged into 500 µL methylation reaction buffer (20 mM Tris, 20 mM TCEP, pH 8.6).

Methylation of Ac-$\alpha S^{F}Asp'_2Hcs_5Cnf_{39}$ to form Ac-$\alpha S^{F}Asp'_2Cnf_{39}$. Following exchange into methylation reaction buffer, the ligation products were quantified using the bicinchioninic acid assay. Methylation was performed by incubating the ligation reaction

with 3,000 equiv methyl iodide for 10 min. The reaction was quenched by adding 10,000 equiv β-mercaptoethanol, diluted to 1 mL with MilliQ H_2O, and immediately purified by HPLC.

Table 3.12 Summary of Protein HPLC Purification.

Peptide	Gradient Description†	Retention Time (min)	Column (Size: Brand)
$\alpha S^{F}_{5\text{-}140}Hcs_5Cnf_{39}$	S8	25.2	Semi-prep: Vydac 218 TP
$\alpha S^{F}_{5\text{-}140}Cnf_{39}$ (Met_5)	S8	26.7	Semi-prep: Vydac 218 TP
Ac-$\alpha S^{F}Asp'_2Hcs_5Cnf_{39}$	S8	34.8	Semi-prep: Vydac 218 TP
Ac-$\alpha S^{F}Asp'_2Cnf_{39}$ (Met_5)	S8	35.9	Semi-prep: Vydac 218 TP

† Described in Table 3.12.

Table 3.13 Solvent Gradient Used for Peptide Purification.

Gradient	Time (min)	Buffer A (%)
S8	0:00	98
	5:00	98
	10:00	67
	50:00	56
	55:00	0
	60:00	0
	65:00	98

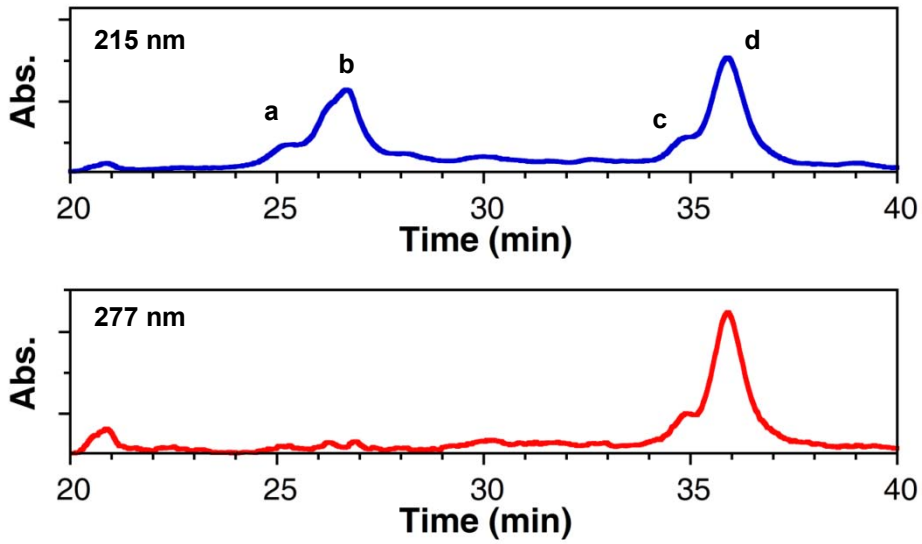

Figure 3.57 HPLC Purification of One-Pot Ligation/Methylation Reaction to Form Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$..
Peaks: a) Unmethylated starting material (αS$^{F}{}_{5\text{-}140}$Hcs$_{5}$Cnf$_{39}$) b) Methylated starting material (αS$^{F}{}_{5\text{-}140}$Cnf$_{39}$) c) Unmethylated ligation product (Ac-αS^{F}Asp'$_{2}$Hcs$_{5}$Cnf$_{39}$) d) Methylated ligation product (Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$).

Absorbance and Fluorescence Spectroscopy. UV-Vis spectra were acquired in quartz cells with a 1.00 cm path length using an HP 8452A spectrometer. Corrected fluorescence spectra were collected in triplicate using quartz fluorometer cells with path lengths of 1.00 cm in a PTI Quantamaster 40 spectrometer with photon counting detection. For all Cnf experiments, the excitation wavelength was 240 nm and emission data were collected from 275 - 400 nm. The excitation and emission slit widths were 5 nm, the step size was 1 nm, and the data interval was 1 nm.

Fluorescence and UV Spectra of Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$. Primary UV and fluorescence spectra were obtained in αS buffer (20 mM Tris, 100 mM NaCl, pH 7.5). Characteristic thioamide absorbance (λ_{max} = 272 nm) and Cnf fluorescence upon excitation with 240 nm light are displayed in the spectra shown below.

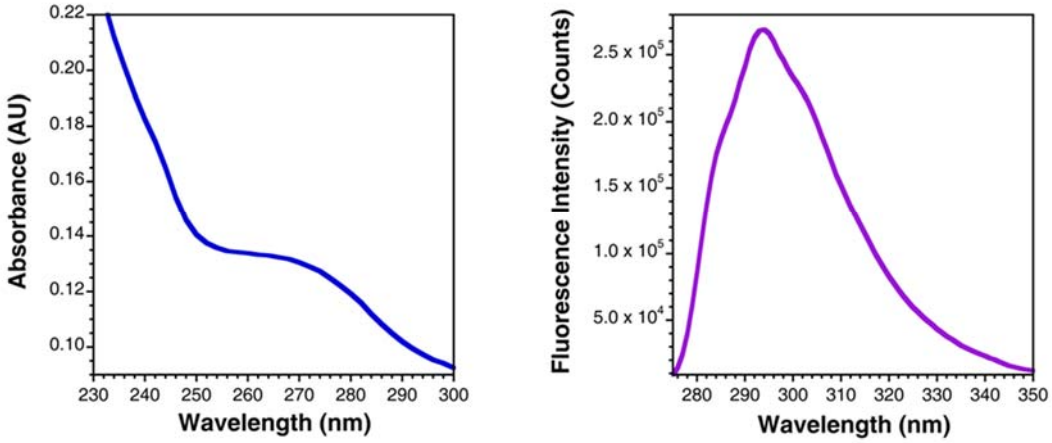

Figure 3.58 Absorbance and Fluorescence Spectra of Purified Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$.

Trypsin Digestion. Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ (30 µg) was incubated with a 5 µL aliquot of sequencing-grade modified trypsin (0.1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 7.5, at 37 ºC for 1.5 h. An aliquot (1.0 µL) of the digestion reaction was removed from the solution and analyzed by MALDI MS. An analysis of the thioamide-containing fragment as well as the Cnf-containing fragment is described in the main text. Notably, alkylation in the histidine-containing fragment of Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ (44-58) was not observed.

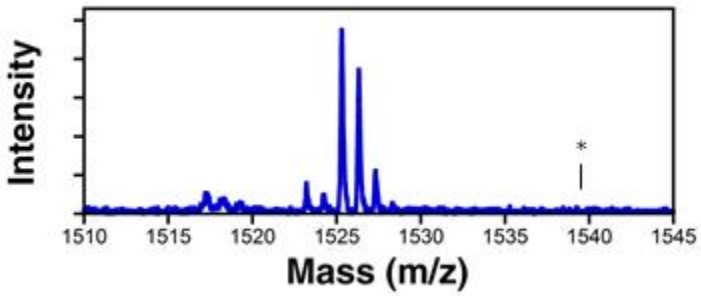

Figure 3.59 MALDI MS of Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ Trypsin Fragment 44-58.
Calcd m/z (M + H)$^{+}$ 1525.8, found 1525.3. Asterisk indicates the absence of a peak corresponding to methylated 44-58. Calcd m/z (M + H)$^{+}$ 1539.8.

Overexpression and Purification of αS^{F}Cnf$_{39}$ (Oxoamide Control). Overexpression and purification of αS^{F}Cnf$_{39}$ was performed as previously described. Briefly, αS^{F}Cnf$_{39}$ was transformed into competent *E. coli* BL21(DE3) cells harboring the orthogonal

$tRNA_{CUA}$ and CnfRS pair. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100 ug/mL) and Strep (100 ug/mL). The primary 5 mL culture was incubated at 37 ºC with shaking at 250 rpm for 5 h. The primary culture was used to inoculate 1 L of a variant of M9 minimal media described previously. When the OD_{600} of the secondary culture reached 0.8, Cnf was added (150 mg, 0.8 mM final concentration), and the culture was incubated overnight at 37 ºC with shaking at 250 rpm. The cells were harvested at 5000 x g for 15 min and the resulting pellet was re-suspended in 20 mM Tris, pH 8.0. Following sonication, the cell lysate was boiled for 20 min prior to centrifugation for 20 min at 30,000 x g, 4 ºC. The cleared lysate was dialyzed against 20 mM Tris pH 8.0, loaded onto a Superdex 200 column (25 cm) and eluted with 20 mM Tris, pH 8.0. The fractions containing $\alpha S^{F}Cnf_{39}$ were then loaded onto a HighTrap Q HP column and eluted over a 100 min sodium chloride gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8.0). FPLC fractions from ion-exchange chromatography were dialyzed against 20 mM Tris, 100 mM NaCl, pH 7.5 and analyzed by MALDI MS and fluorescence spectroscopy.

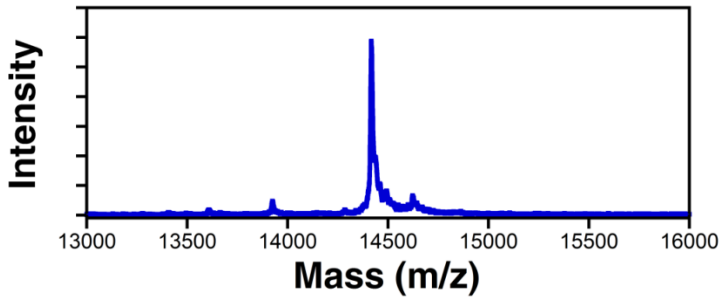

Figure 3.60 MALDI MS of $\alpha S^{F}Cnf_{39}$.
Calcd m/z $(M + H)^{+}$ 14423, found 14420.

Fluorescence Spectrometry of $\alpha S^{F}Cnf_{39}$. Primary fluorescence spectra were obtained in αS buffer (20 mM Tris, 100 mM NaCl, pH 7.5). Characteristic Cnf fluorescence upon excitation with 240 nm light is displayed in the spectrum shown below.

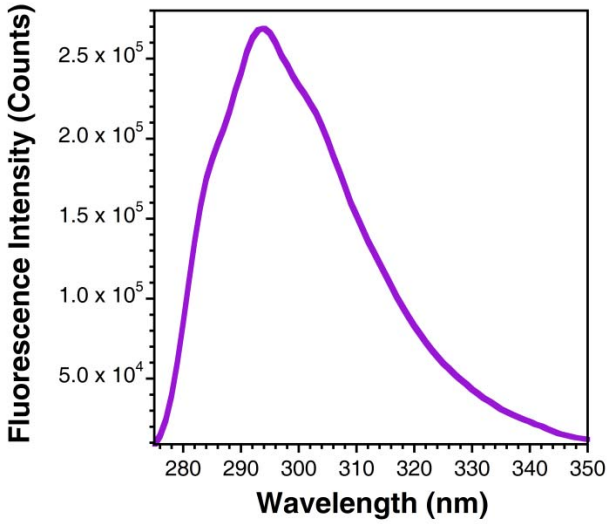

Figure 3.61 Fluorescence Spectrum of Purified $\alpha S^{F}Cnf_{39}$.

αS Refolding Assays. Prior to performing the experiment, the concentrations of Ac-$\alpha S^{F}Asp'_{2}Cnf_{39}$ and $\alpha S^{F}Cnf_{39}$ were adjusted to 10 μM. Tris buffers (20 mM Tris, 100 mM NaCl, pH 7.5) containing trimethylamine oxide (TMAO) were prepared such that upon addition of αS, the final TMAO concentrations were 0, 2, and 4 M. Fluorescence spectra were obtained immediately following sample dilution (to a final concentration of 1 μM) in TMAO buffer. All samples were prepared and measured in triplicate. The background fluorescence of each buffer was independently measured and subtracted from the protein fluorescence spectrum.

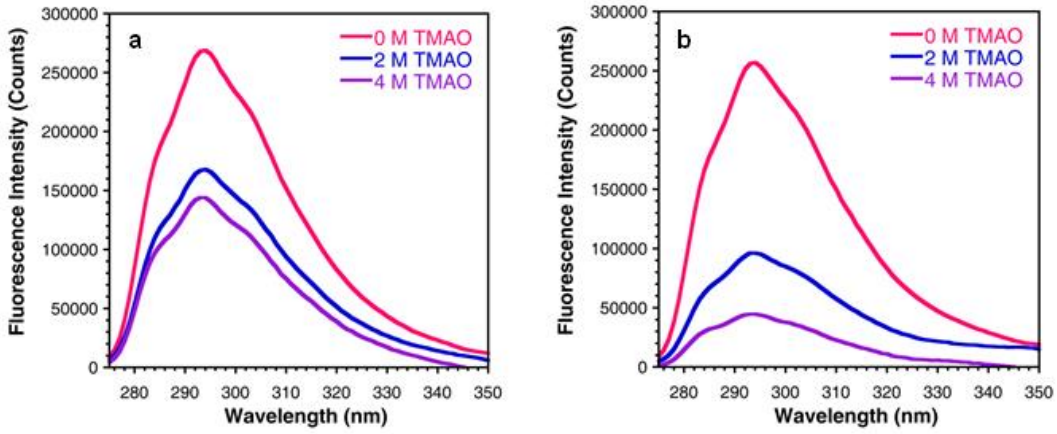

Figure 3.62 Fluorescence Spectra of αS^{F}Cnf$_{39}$ and Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ in Increasing Concentrations of TMAO.
Fluorescence Spectra of αS^{F}Cnf$_{39}$ (a) and Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ (b) in increasing concentrations of TMAO.

Förster Distance Calculations. Förster distance calculations for double-labeled αS were performed as described above. Quantum yields for each mutant in different concentrations of TMAO were determined by comparing the fluorescence output of αS^{F}Cnf$_{39}$ relative to its fluorescence in 0 M TMAO. The value of R_0 for αS^{F}Cnf$_{39}$ mutants in 0, 2, and 4 M TMAO was determined to be 15.6, 14.7, and 14.4 Å, respectively. The experimental chromophore separation (R_{FRET}) in Ac-αS^{F}Asp'$_{2}$Cnf$_{39}$ was determined as previously described.

Table 3.14 R_{FRET} Determination in αS Refolding Assay.

[TMAO]	E_Q	SE
0	0.04	0.04
2	0.42	0.04
4	0.69	0.04

[TMAO]	R_{FRET}[a]	SE[a]
0	26.2	7.8
2	15.4	0.4
4	12.6	0.4

[TMAO]	F_{Oxo}[b]	R_0[a]
0	278856	15.6
2	190131	14.7
4	168416	14.4

[a] Distances in Å. [b] Fluorescence intensity in counts

Chapter 4 : Incorporation of Visible Wavelength Fluorescent Labels into α-Synuclein for Aggregation and Disaggregation Studies

§ 4.1 Strategies for Efficiently Incorporating Non-Perturbing Labels Into α-Synuclein

Our understanding of the dynamics of protein folding, protein-protein interactions, cellular trafficking, and the effects of protein post-translational modifications can be facilitated by labeling proteins with appropriate reporters. Two classes of proteins of particular current interest are intrinsically disordered proteins and amyloidogenic proteins. These proteins, which are refractory to traditional structural characterization techniques, can be labeled with fluorophores so that methods such as Förster resonance energy transfer (FRET) or fluorescence quenching can be used to monitor their conformations *in vitro* and *in vivo*.[214,215] To do so, one must find a way to selectively introduce multiple site-specific labels that can effectively report on protein dynamics without being disruptive to the native fold of the protein. Here, we use the unstructured and amyloidogenic protein α-synuclein (αS) to demonstrate effective combinations of several methods for protein labeling. While we focus on αS, these strategies should be broadly applicable to the protein folding community.

To be useful, protein labeling must be specific, either through direct genetic incorporation of a label, or through the introduction of a chemoselective reactive handle for further functionalization. Reactive groups that can be selectively modified in the presence of other biomolecule functional groups are described as bioorthogonal.[25] The most well known reactions in this category are the Cu-catalyzed or strain-promoted azide/alkyne cycloadditions commonly referred to as "click" reactions.[29,216] The simultaneous site-specific introduction of multiple functionalities into a protein further requires that the labeling techniques be mutually orthogonal, or capable of being carried

out sequentially. In order to frame our discussion of how the following labeling techniques may be used together, we must first introduce them individually.

Several genetically controlled strategies currently exist for protein labeling. The most time-tested method is Cys modification, which is typically done through reaction with an alkyl halide or maleimide label.[217] However, Cys modification will only be site-specific if the protein contains no other Cys residues or if reaction conditions can be used to suppress the reactivity of the undesired sites. The amber codon suppression method developed by Schultz and co-workers can be used to site-specifically incorporate an unnatural amino acid (Uaa) into a protein at an amber (UAG) stop codon.[23] Efficient incorporation into recombinant proteins is typically accomplished using evolved orthogonal tRNA/aminoacyl tRNA synthetase (aaRS) pairs to generate the requisite Uaa-tRNA for ribosomal incorporation of the Uaa. The Uaa can either feature an intrinsic label such as a fluorophore, or a bioorthogonal reactive handle for further modification such as an azide. The Uaa mutagenesis method has been expanded by Chin, Liu, and others to enable the incorporation of multiple Uaas.[35,46,218-222] Alternatively, Tirrell and co-workers have demonstrated the utility of so-called auxotrophic organisms for incorporating Uaas into proteins.[223,224] Auxotrophic organisms are engineered to be deficient in a specific natural amino acid. By taking advantage of the promiscuity of endogenous RSs by supplying an isosteric Uaa during protein expression in an auxotroph, one can achieve codon-specific Uaa incorporation. While auxotrophs could conceivably be used in combination with other methods, we have not investigated them in depth, as they would only permit site-specific labeling in unusual cases or through mutagenesis to remove certain residues as the Davis group has done.[225]

In addition to these co-translational methods, several post-translational chemoenzymatic strategies exist for transferring labels onto proteins at unique recognition motifs. Some enzymes, like sortase, can be used for conjugation to the protein termini.[16,17] Other enzymes, like biotin ligase and lipoic acid ligase, can be used to attach labels to specific amino acid sidechains.[16,19] However, sortase, biotin ligase, and lipoic acid ligase all require the insertion of a 5-15 amino acid recognition motif that may prove to be disruptive to protein folding. In contrast, *E. coli* aminoacyl transferase (AaT) requires only a single amino acid for recognition. AaT is an enzyme in the N-end rule pathway that transfers Leu, Phe, or Met from the corresponding aminoacyl tRNA to proteins containing an N-terminal Lys or Arg.[226] AaT is relatively promiscuous, so it can be used to transfer a wide range of Uaas to the N-termini of peptides and proteins if the appropriate aminoacyl tRNA donor or small molecule analog can be generated.[227] Importantly, one can control the timing of AaT modification to introduce a second label independently of a primary modification made co-translationally (e.g., by amber suppression).

Finally, one of the most versatile and general methods for protein modification is native chemical ligation (NCL).[143] NCL reactions involve the reaction of a peptide with a C-terminal thioester and a peptide with an N-terminal Cys residue. Thus, like AaT modification, NCL requires only a single amino acid motif for synthesis, and through the use of β-thiol Cys analogs and desulfurization, even this single amino acid motif can be "erased." The two sections of the protein can be generated either through synthesis or recombinant expression, allowing one to produce semi-synthetic proteins where large segments of unmodified protein are derived from cellular expression.[162] NCL methods

allow for essentially complete freedom in the synthetic portion of the protein since no enzyme recognition requirements must be satisfied. The two primary drawbacks to the use of NCL for protein modification are that the semi-synthesis process can be very labor intensive and that refolding after synthesis is required, which may not be feasible for some proteins. Thus, although one can conceivably introduce all desired modifications through NCL, it is not always the optimal technique.

The ideal synthesis of a multiply labeled version of a protein is likely to involve a judicious combination of techniques. Cys modification, amber codon suppression, AaT N-terminal modification, and NCL are techniques that can be combined for protein multifunctionalization using Uaas. While we stress that these are general techniques, we will demonstrate all of them in syntheses of αS. αS is an intrinsically disordered 140 amino acid protein, aggregates of which are among the hallmarks of Parkinson's disease.[44,51,78,145] Residues 61-95 of αS comprise the so-called non-amyloid-β component (NAC) domain and drive the formation of β-sheet rich amyloidogenic aggregates. *In vitro* aggregation studies suggest a nucleation-propagation mechanism for αS fibril assembly that is highly sensitive to experimental conditions.[154] Labeling of αS has been used previously to investigate the mechanisms that drive αS fibril formation and cellular toxicity. Macromolecular labeling of αS using either a yellow fluorescent protein (YFP) fusion construct or immuno-tags has been used for monitoring protein aggregation *in vitro* and *in vivo*.[144] Imaging of fibrillization in living cells was accomplished by inserting the CysCysProGlyCysCys motif at the C-terminus of αS and labeling using biarsenical fluorophore reagents (i.e., FlAsH).[228] These labeling techniques, which require the insertion of a large protein label or the binding of a label molecule to a section of the

protein, can be disruptive to the folding of αS. Thus, our laboratory and others have focused on αS labeling through the modification of single amino acids.

Cys-maleimide fluorophore labeling was combined with Trp fluorescence to generate a FRET pair to probe αS conformational dynamics just prior to aggregation.[229] Alternatively, others groups have prepared doubly labeled αS using stochastically controlled fluorophore-maleimide labeling of double Cys mutants.[69,81] This leads to a mixture of forms of singly and doubly labeled proteins, which can be difficult to separate. Using Cys and pH-dependent N-terminal α-amine labeling, it was shown that a conformational change precedes fibril formation and that fibril nucleation is concentration dependent.[103] However, it is difficult to achieve complete specificity for the N-terminus in this way. Lashuel and Brik have used both semi-synthesis and total synthesis through NCL to incorporate authentic post-translational modifications at many sites in αS, including ubiquitination and phosphorylation.[163,230] Pielak and coworkers have used amber suppression to incorporate ^{19}F NMR probes into αS for folding studies.[231]

Recently, our laboratory has demonstrated the ability to incorporate a variety of minimally perturbing probes into αS. Spectroscopic probes such as cyanophenylalanine or acridonylalanine (Acd) have been incorporated using amber suppression.[47,117] Double labels, such as the fluorophore/quencher pair of fluorescein and a thioamide, have been incorporated using sequential NCL and Cys modification.[151] We have also combined NCL and amber codon suppression to attach Cnf/thioamide pairs.[117] We have further combined these methods with AaT modification to enable the traceless incorporation of N-terminal αS modifications.[117] While these combinations of methods have allowed us to

multiply label αS, due to the labor-intensive nature of the peptide synthesis in NCL, we have continued to examine other combinations of techniques that can be used in a straightforward manner to generate large quantities of protein for *in vitro* and *in vivo* aggregation studies. Here, we compare combinations of a number of labeling methods to generate homogeneous, multiply labeled versions of αS.

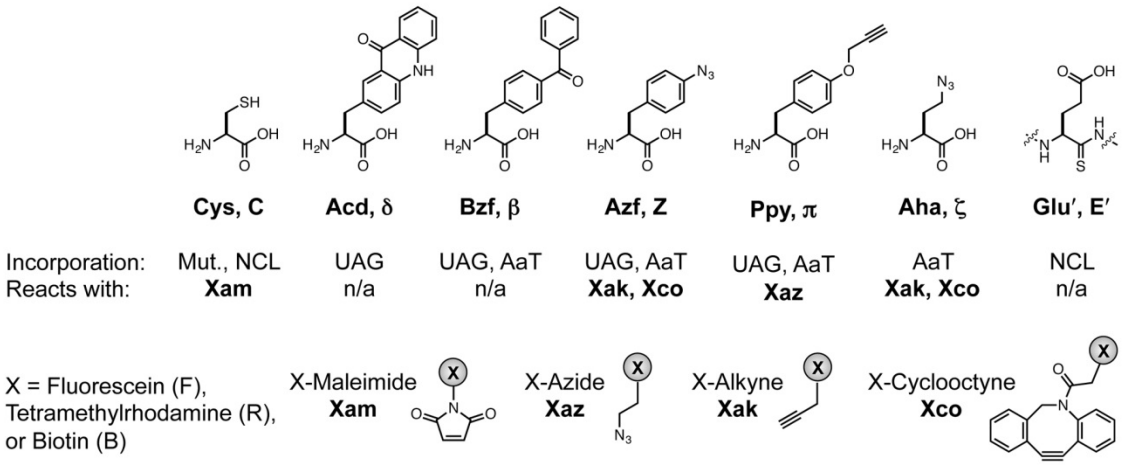

Figure 4.1 Methods for Amino Acid Incorporation and Modification.
Cysteine (Cys, C), acridonylalanine (Acd, δ), benzoylphenylalanine (Bzf, β), azidophenylalanine (Azf, Z), propargyl tyrosine (Ppy, π), azidohomoalanine (Aha, ζ), and thioglutamate (Glu′, E′) are incorporated by simple mutation (Mut.), native chemical ligation (NCL), amber suppression (UAG), or aminoacyl transferase (AaT) and may be modified with the reagents shown.

§ 4.2 Labeling Results and Discussion

In using combinations of labeling methods, it is important to consider whether any given pair of methods satisfies the criteria for orthogonality and for sequential use. We deem methods to be orthogonal if one type of reaction can be carried out in the presence of a functional group for the other type of reaction; we do not require that they be used in one pot. Some reactions are inherently sequential, such as the co-translational incorporation of an intrinsic spectroscopic probe and the post-translational attachment of a second probe. Other pairs of reactions could be carried out in either order, but we have identified a preferred sequence. Techniques requiring click chemistry modification and

Cys modification fall into this category. Finally, some combinations are only practical in a certain order. We have generated at least one example of every labeling combination, and for several examples, we have explored multiple strategies. Of course, before generating any doubly labeled constructs, we assessed the corresponding singly labeled constructs. All purified, labeled proteins were characterized by whole protein matrix-assisted laser desorption ionization mass spectrometry (MALDI MS) as well as trypsin digest followed by MALDI MS to confirm the extent and specificity of labeling.

Cys Modification

While several types of reactions are available for Cys modification, including radical-mediated thiol-ene coupling, photochemically-controlled reactions, and elimination/addition chemistry; simple S_N2 or Michael additions are by far the most common Cys reactions, with a great number of reagents commercially available.[217] In addition, we expected either reaction with the Cys nucleophile to be compatible with the other labeling reactions considered here. Since αS has no endogenous Cys residues, we easily generated two singly modified αS constructs by expressing αS-C_9 and αS-C_{114}, labeled them in a semi-pure state with fluorescein-maleimide (Fam) to generate αS-C^{Fam}_9 and αS-C^{Fam}_{114}, and purified the products by high performance liquid chromatography (HPLC). Analyses of the crude labeling reactions indicated that the yields of αS-C^{Fam}_9 and αS-C^{Fam}_{114} were quantitative (see Section 4.7, Materials and Methods). Whole protein MALDI MS and trypsin digest confirmed complete and site-specific labeling in the purified products. Trypsin digestion also revealed that ring-opening hydrolysis of the maleimide unit (corresponding to a mass increase of + 18 Da in the observed fragment) occurred during the labeling and purification procedures.

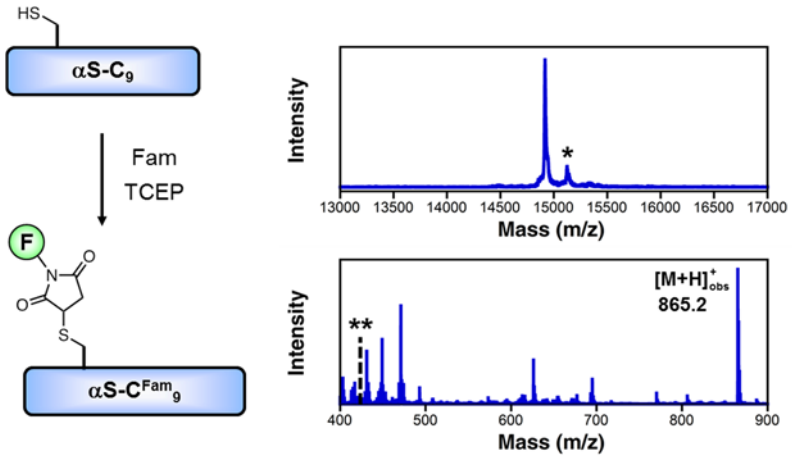

Figure 4.2 Single Labeling of αS-C$_9$ with Fam Maleimide.
Top MALDI MS: Whole protein and trypsin digest MALDI MS data show that the purified protein is homogeneously labeled. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da). Bottom MALDI MS: Calcd m/z for trypsin fragment 7-10: 847.3, Obs 865.2 (+18 Da). The double asterisk (**) corresponds to the unmodified trypsin fragment 7-10; no peak is observed. MALDI MS data for αS-C$_{114}$ can be found in Materials and Methods.

Uaa Mutagenesis

Note: *Performed by John Warner and Lee Speight and described in detail elsewhere.*[47,232]

For Uaa modification, we examined two types of Uaas: those with bioorthogonal reactive handles (i.e., azides and alkynes) and intrinsic probes such as the fluorescent amino acid acridonylalanine (Acd, δ). Acd is a blue wavelength fluorophore that can be selectively excited in the presence of Trp and Tyr.[47,113] Acd can function on its own as an environmentally sensitive fluorescent probe, or it can be paired with a FRET acceptor such as the quencher Dabcyl or the fluorophore BODIPY R6G. We expressed αS- δ$_{39}$ and αS- δ$_{94}$ using standard amber suppression protocols as described previously.[47] First, we inserted TAG codons at the appropriate positions in the αS gene. We then transformed each αS plasmid into *E. coli* cells along with a plasmid encoding a mutant *Methanococcus jannaschii* (*Mj*) TyrRS/tRNA pair (pDULE2-Acd) and induced expression along with addition of Acd to give αS-δ$_{39}$ and αS-δ$_{94}$ in ~3 mg/L yields. In

this case, whole protein and trypsin digest MALDI MS were used to assess the fidelity of Acd incorporation.

While intrinisic probes like Acd are valuable, they are often limited in their functionality by their need to remain compatible with the translational machinery. Thus, several RS/tRNA pairs derived from the *Mj* TyrRS/tRNA have been selected for Uaas amenable to post-translational modification through click chemistry reactions, including pXFRS for both azidophenylalanine (Azf, Z) and propargyltyrosine (Ppy, π). Selective post-translational reactions with Azf and Ppy can be used to derivatize them with virtually any type of label. Using the αS plasmid from above with the pDULE2-pXF plasmid, αS-Z_{94} was expressed and labeled with tetramethylrhodamine-dibenzocyclooctyne (Tco) to give αS-Z^{Fco}_{94}. MALDI MS analysis of the crude labeling mixture showed that the labeling yield was incomplete (see Section 4.7, Materials and Methods). Aryl azides such as Azf are subject to both chemical reduction and photodegradation that could prevent complete labeling.[233] Although light exposure is easily limited by working in the semi-dark, some Azf reduction can occur in cells or during handling of the proteins. Unfortunately, Azf-containing peptide fragments are ionized to *p*-aminophenylalanine species during MALDI MS analysis, so it is difficult to directly determine the extent of Azf reduction, and we infer it from the yields of reactions with alkynes. Suspecting that the problem lay in the kinetics of the azide/cyclooctyne reaction, we subjected αS-Z_{94} to Cu-catalyzed cycloaddition reactions with tetramethylrhodamine-alkyne (Tak) to give αS-Z^{Fak}_{94}. We observed slightly higher modification yields (by MALDI MS) than for stain-promoted cycloaddition with Fco, but the labeling reaction did not go to completion (see Section 4.7, Materials and Methods).

Unlike azides, which can be reacted with alkynes through either strain-promoted cyclizations or Cu-catalyzed cyclizations, terminal alkynes like Ppy must be reacted with azides using Cu catalysis. We expressed αS-π_{39} and αS-π_{94} using the αS plasmids and the pDULE2-pXF plasmid, and derivatized the Ppy residues with tetramethylrhodamine-azide (Raz) to give αSπ^{Raz}_{39} and αS-π^{Raz}_{94}. When the proteins were analyzed by MALDI MS to confirm the extent and specificity of labeling, we observed near quantitative labeling at both sites (see Section 4.7, Materials and Methods). We believe that the lower labeling efficiency of Azf relative to Ppy under otherwise identical Cu-catalyzed conditions is the result of reduction of some of the Azf residues. This is consistent with our observations for other reactions with co-translationally incorporated Azf (see below).

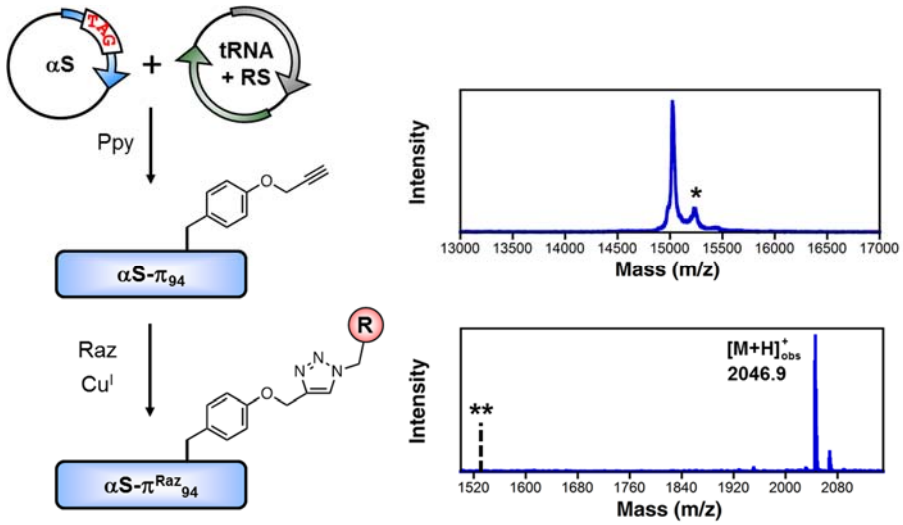

Figure 4.3 Single Labeling of αS-π^{Raz}_{94} with Raz.
Top MALDI MS: Whole protein and trypsin digest MALDI MS data show that the purified protein is homogeneously labeled. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da). Bottom MALDI MS: Calcd m/z for trypsin fragment 81-96: 2047.0, Obs 2046.9. The double asterisk (**) corresponds to the unmodified trypsin fragment 81-96; no peak is observed. MALDI MS data for αS-π^{Raz}_{39} can be found in Materials and Methods.

AaT Modification

Note: *Performed by John Warner and described in detail elsewhere.*[232]

We have previously shown that AaT is capable of modifying protein N-termini with Uaas using semi-synthetic aminoacyl adenosine donors.[227] Our laboratory and others have also used RSs to generate aminoacyl tRNA substrates for AaT *in situ*.[203,210,227] Although aminoacyl adenosine donors are useful for screening the substrate specificity of AaT, if an RS appropriate for the Uaa can be identified, the RS method is more straightforward for most laboratories to implement. Therefore, we sought to develop RS/AaT protocols for labeling the αS N-terminus that could be combined with other labeling methods. An αS construct with an N-terminal Lys was generated by inserting a 10 His tag with a Factor Xa proteolysis site at the protein N-terminus, deleting Met1, and making the mutation Asp2Lys (His_{10}-$\alpha S_{2\text{-}140}$-K_2). Following Ni-NTA affinity purification, the His_{10} tag was cleaved with Factor Xa and the N-terminal Lys was revealed. This construct, $\alpha S_{2\text{-}140}$-K_2, could then be modified with a variety of Uaas using an RS/AaT protocol, including Azf, Ppy, azidohomoalanine (Aha, ζ) or homopropargylglycine (not shown). For example, Aha was transferred to $\alpha S_{2\text{-}140}$-K_2 using a mutant *E. coli* MetRS (Met*RS), ATP, *E. coli* total tRNA, and AaT. Transfer was complete within 2 h as monitored by MALDI MS. $\alpha S_{2\text{-}140}$-$K_2\zeta_N$ was subsequently modified with Fco or a dibenzylcyclooctyne–tetramethylrhodamine (Rco) fluorophore using strain-promoted azide-alkyne cycloaddition. Interestingly, nearly complete labeling of the N-terminus was observed, in contrast to our results with Azf incorporated through amber suppression. The purified proteins, $\alpha S_{2\text{-}140}$-$K_2\zeta^{Fco}{}_N$ and $\alpha S_{2\text{-}140}$-$K_2\zeta^{Rco}{}_N$, were characterized by MALDI MS and labeling was found to be selective, as expected. In additional experiments in which we labeled the N-terminus with Azf using AaT, we also observed complete labeling at the N-terminus with Fco. This indicates that the

limited Fco reactivity observed for internal Azf modifications may be due to steric hinderance not present at the terminus. Moreover, it shows that Azf is stable in the short term after transfer by AaT, so that partial reduction observed elsewhere probably occurs during growth and lysis of *E. coli* cells.

NCL then Cys Modification

We have previously shown that we can label αS with spectroscopic probes at either the C-terminus or the N-terminus through NCL. While there are a great many modifications possible, here we restrict ourselves to thioamide fluorescence quenching probes since there is no equivalent way of installing these backbone modifications through any of the other labeling methods. We examined two modifications, Asp'_2 and Glu'_{137}, where we denote the presence of the thioamide by adding a prime symbol to the one or three letter code for the corresponding natural amino acid. To insert the Asp'_2 modification, we expressed the truncated construct His_{10}-$\alpha S_{9\text{-}140}$-C_9, cleaved the N-terminal His_{10} tag with Factor Xa, and ligated the resulting $\alpha S_{9\text{-}140}$-C_9 fragment to the synthetic thioester peptide $\alpha S_{1\text{-}8}$-D'_2-SR in quantitative yield (MALDI analysis, see Section 4.7, Materials and Methods). We have previously described insertion of the Glu'_{137} modification, where we expressed the intein fusion construct $\alpha S_{1\text{-}113}$-$IntH_6$, cleaved the C-terminal intein with mercaptoethanesulfonate (MESNA), and ligated the resulting thioester fragment $\alpha S_{1\text{-}113}$-SR to the synthetic peptide $\alpha S_{114\text{-}140}$-$C_{114}E'_{137}$.[151] The crude proteins, αS-D'_2C_9 and αS-$C_{114}E'_{137}$, were taken forward to the second labeling step where they were incubated with Fam. The doubly labeled proteins αS-$D'_2C^{Fam}_9$ and αS-$C^{Fam}_{114}E'_{137}$ were then purified and characterized by MALDI MS.

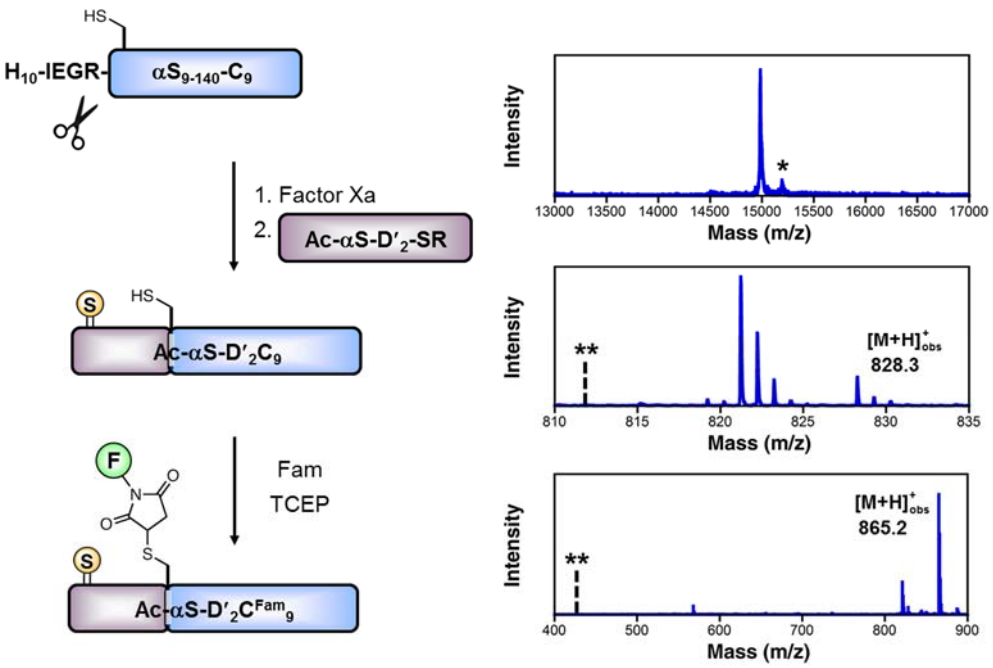

Figure 4.4 NCL and Cys Modification to Generate Ac-αS-D′$_2$C$^{Fam}_9$.
His$_{10}$-IEGR-αS$_{9\text{-}140}$-C$_9$ was cleaved using Factor Xa and subject to NCL with Ac-αS-D′$_2$-SR to generate Ac-αS-D′$_2$C$_9$. Top MALDI MS: Following the ligation reaction, Ac-αS-D′$_2$C$_9$ was labeled with Fam to yield the doubly-labeled protein. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da). Middle MALDI MS: Calcd m/z for trypsin fragment 1-8: 828.3, Obs 828.3. The double asterisk (**) corresponds to the unmodified trypsin fragment 1-8; no peak is observed. Calcd m/z for trypsin fragment 7-10: 847.3, Obs 865.2 (+18 Da). The double asterisk (**) corresponds to the unmodified trypsin fragment 7-10; no peak is observed.

Cys Modification and Uaa Mutagenesis

Cys modification and Uaa mutagenesis are perhaps the most versatile and obvious combination of methods for doubly labeling a protein. In spite of this, there are relatively few examples of their use together.[21,28,31] For both methods, many substitutions are possible. To produce doubly-labeled proteins for fluorescence experiments, we began with Cys modification of a Ppy-labeled protein. TAG mutations were introduced into the Cys mutant plasmids at positions 39 and 94, then αS-C$_9$π$_{94}$, and αS-π$_{39}$C$_{114}$ were expressed using the standard amber suppression protocols with pDULE2-pXF. We performed the Cys modification with Fam first, and then the Cu-catalyzed addition of

Raz to Ppy. We found that both reactions occurred in high yields to give αS-$C^{Fam}{}_{9}\pi^{Raz}{}_{94}$ and αS-$\pi^{Raz}{}_{39}C^{Fam}{}_{114}$.

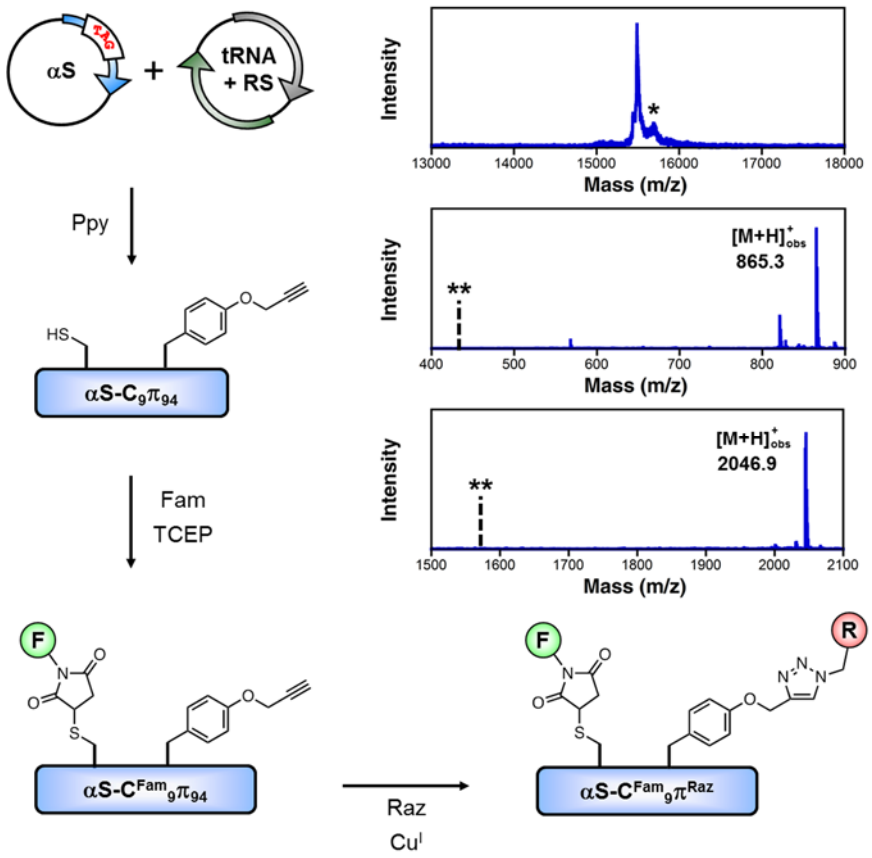

Figure 4.5 Double Labeling of αS-$C_9\pi_{94}$ with Fam and Raz.
Top MALDI MS: Whole protein and trypsin digest MALDI MS data show that the purified protein is homogeneously labeled. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da). Middle MALDI MS: Calcd m/z for trypsin fragment 7-10: 847.3, Obs: 865.3 (+18 Da). The double asterisk (**) corresponds to the unmodified trypsin fragment 7-10; no peak is observed. Bottom MALDI MS: Calcd m/z for trypsin fragment 81-96: 2047.0, Obs: 2046.9. The double asterisk (**) corresponds to the unmodified trypsin fragment 81-96; no peak is observed. MALDI MS data for αS-$C_{39}\pi_{114}$ can be found in Materials and Methods.

Uaa Mutagenesis then AaT

Note: *Performed by John Warner and described in detail elsewhere.*[232]

Many of the examples of amber suppression and AaT modification considered thus far have involved bioorthogonal chemistry. Because one can control the timing of the AaT modification, azide/alkyne cycloadditions can easily be used for both

modification steps when combining AaT and amber suppression. To develop procedures for these modifications, we generated two constructs by introducing amber codons into the His_{10}-$\alpha S_{2\text{-}140}K_2$ construct at either Tyr39 or Phe94. As before, Azf was incorporated during expression in *E. coli* cells using the pDULE2-pXF plasmid to produce His_{10}-$\alpha S_{2\text{-}140}$-K_2Z_{39} and His_{10}-$\alpha S_{2\text{-}140}$-K_2Z_{94}, and the proteins were purified using the His tags. Labeling of Azf at either residue 39 or 94 was performed with Rco to give His_{10}-$\alpha S_{2\text{-}140}$-$K_2Z^{Rco}{}_{39}$ and His_{10}-$\alpha S_{2\text{-}140}$-$K_2Z^{Rco}{}_{94}$. Factor Xa proteolysis was then used to reveal the N-terminal Lys for subsequent transfer of Aha to the N-terminus. Addition of Aha to $\alpha S_{2\text{-}140}$-$K_2\zeta_N Z^{Rco}{}_{39}$ and $\alpha S_{2\text{-}140}$-$K_2\zeta_N Z^{Rco}{}_{39}$ proceeded in quantitative yields to give $\alpha S_{2\text{-}140}$-$K_2\zeta_N Z^{Rco}{}_{39}$ and $\alpha S_{2\text{-}140}$-$K_2\zeta_N Z^{Rco}{}_{94}$. The N-terminal Aha was then modified with Fco for both constructs producing $\alpha S_{2\text{-}140}$-$K_2\zeta^{Fco}{}_N Z^{Rco}{}_{39}$ and $\alpha S_{2\text{-}140}$-$K_2\zeta^{Fco}{}_N Z^{Rco}{}_{94}$ in quantitative yields. The proteins were purified by reverse phase HPLC and characterized by MALDI MS and trypsin digest followed by MALDI MS.

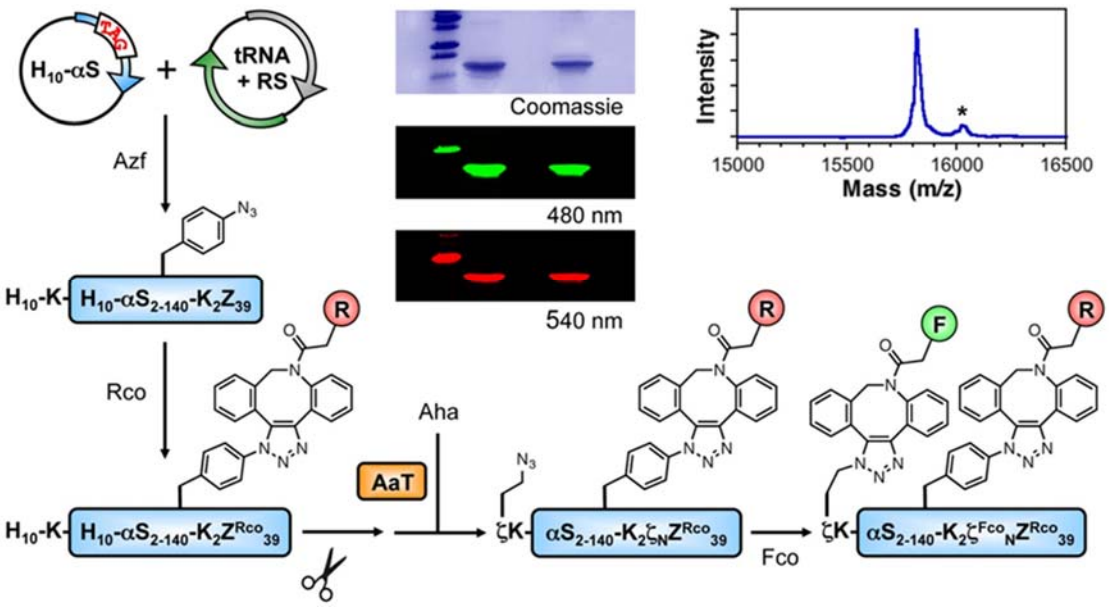

Figure 4.6 Double Labeling Using Amber Suppression and N-Terminal AaT Transfer. Azf was incorporated into αS and modified with tetramethylrhodamine-dibenzocyclooctyne (Rco). Then Aha was attached by AaT-mediated transfer and modified with fluorescein-dibenzocyclooctyne (Fco). The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

In our efforts towards $\alpha S_{2\text{-}140}\text{-}K_2\zeta^{Fco}{}_NZ^{Rco}{}_{39}$ and $\alpha S_{2\text{-}140}\text{-}K_2\zeta^{Fco}{}_NZ^{Rco}{}_{94}$, we observed incomplete labeling of Azf at residues 39 or 94, similar to what was seen in using Cys modification together with Uaa mutagenesis. We again attributed this to azide reduction, although in this case relatively low levels of thiol were present (AaT can be handled with β-mercaptoethanol concentrations as low as 5 μM). To address the limitation of incomplete Azf labeling, we incorporated Ppy at position 39 to give $His_{10}\text{-}\alpha S_{2\text{-}140}\text{-}K_2\pi_{39}$. The extent of protein labeling of Azf versus Ppy was compared using Cu-catalyzed fluorophore labeling. The Azf protein ($His_{10}\text{-}\alpha S_{2\text{-}140}\text{-}K_2Z_{39}$) was incomplete whereas labeling of the Ppy protein ($His_{10}\text{-}\alpha S_{2\text{-}140}\text{-}K_2\pi_{39}$) occurred in near quantitative yields. Thus, as before, the incomplete Azf labeling observed with Fco and Rco seems to be due to instability of the Azf residue, not a limitation of the cyclooctyne chemistry. In general, Cu-catalyzed reactions with Ppy should be used for *in vitro* labeling unless the

reaction conditions are detrimental to protein stability. For combinations of Uaa mutagenesis and AaT N-terminal labeling, one should label the internal Ppy first, and then attach either Aha or Hpg using AaT. Since no azide stability issues were observed at the N-terminus in either singly or doubly labeled constructs, cyclooctyne chemistry can be used with N-terminal Aha, if desired, or Cu-catalyzed reactions could be used with Hpg.

§ 4.3 Monitoring αS Dynamics using Fluorescence Polarization

Misfolding of the intrinsically disordered neuronal protein αS has been implicated in the pathogenesis of several debilitating neurodegenerative disorders including Parkinson's disease (PD).[234] Although it is well established that misfolding of αS leads to the formation of β-sheet rich fibrils, the molecular mechanisms underlying the conversion of αS into neurotoxic species remain poorly understood. Commonly used amyloid binding dyes, such as Thioflavin T (ThT) and Congo Red (CR), can be used to detect the presence of αS fibrils.[174] Although extrinsic dyes are routinely employed to monitor the kinetics of amyloid formation, the structural information obtained from these experiments is limited.

It has recently been shown that pre-formed fibrils can recruit native αS and convert it into a pathogenic form *in vivo*.[56] Accordingly, small molecules that can remodel and/or disaggregate αS fibrils into less toxic species are being investigated as potential therapeutics to combat the spread of PD.[235-241] Amyloid dyes cannot be used to reliably detect morphological changes of αS induced by small molecules, as it is possible, even likely, that these compounds compete for similar binding sites. Moreover, the optical properties of fluorescence-based amyloid dyes are extremely sensitive to the presence of additional compounds. Improved methods for monitoring the aggregation and

disaggregation of αS that yield meaningful structural information would greatly facilitate our understanding of these crucial processes.

Fluorescence polarization (FP) spectroscopy is a powerful technique for monitoring changes in the conformational mobility of fluorescently-labeled molecules.[160] Upon irradiation with plane polarized light, fluorophores that have their absorption dipole transition moments oriented along the incoming electric field vector are preferentially excited. If the fluorophore remains relatively immobile for the duration of its excited state lifetime (typically 1-10 ns), its emission will be polarized along the same direction. If the fluorophore is able to rotate freely in solution, its emission will become depolarized with respect to the incident excitation vector. In a typical FP experiment, the extent of polarization is determined by measuring the fluorescence emission intensities parallel and perpendicular to the plane of the linearly polarized excitation light. Quantitatively, polarization is expressed as the difference in the parallel and perpendicular emission intensities normalized by the total fluorescence intensity (Equation 4.1).[160]

$$P = \frac{\parallel - \perp}{\parallel + \perp} \qquad \text{(Eq. 4.1)}$$

The ability of a fluorescently-labeled molecule to rotate in solution is dependent on its conformation, its average molecular mass, and on the viscosity of its local environment. By incorporating a fluorescent label into a protein, one can use FP spectroscopy to obtain a wide range of dynamic information. Since the degree of polarization is independent of the total observed intensity, FP measurements are generally insensitive to fluctuations in sample concentration and/or to the presence of fluorescence quenchers.[242]

In 2007, the Lee laboratory demonstrated that FP can be used to monitor the aggregation of αS.[243] In this study, fluorescently-labeled αS was generated by treating purified recombinant αS with succinimidyl esters of either Oregon Green 488 or Alexa Fluor 594. Succinimidyl esters react with free amino groups (such as the ε-amino group of lysine residues or the N-terminus) to form an amide linkage.[244] To avoid attaching multiple fluorophores to αS, the ratio of dye to protein used in the labeling reaction was kept low (1:10). However, since αS contains up to 16 sites that can be labeled by succinimidyl esters, the material that was ultimately produced by this method likely consisted of a complex mixture of labeled proteins. Regardless, aggregation reactions containing WT αS doped with a small percentage of labeled αS generated fibrils that were morphologically indistinguishable from WT fibrils by electron microscopy (EM).

Over the course of aggregation, FP values of the labeled protein increased in a manner that was consistent with a nucleation-dependent process characteristic of αS assembly.[154] In order to compare the observed changes in FP to amyloid formation, the aggregation reaction was independently monitored using ThT as well as another well-known amyloid dye, K114. The observed increase in polarization preceeded fibril detection by ThT or K114, suggesting that FP can be used to detect the formation of early oligomeric intermediates. Next, the authors performed FP aggregation assays in the presence of small molecules that had previously been shown to prevent fibril formation. Norepinephrine, epinephrine, dopamine, and L-DOPA were added to aggregation reactions containing labeled αS. In the presence of each compound, only a modest increase in FP was detected. Sedimentation and EM analyses confirmed that the majority of αS remained soluble under these conditions. Together, these studies confirm that FP

can be used to analyze the conformational state of aS and that FP assays can be performed in solutions containing extraneous compounds.

Although this initial report demonstrates the utility of FP for monitoring αS aggregation, the proteins used in these studies consisted of heterogeneous mixtures of labeled constructs. Therefore, we sought to determine whether site-specifically labeled αS could be used in conjunction with FP to distinguish local conformational rearrangements that occur during folding and misfolding of αS. Although several small molecules have been reported to remodel and/or disaggregate pre-formed αS fibrils, robust methods for examining this process in real-time are currently lacking. Here, we show that we can use FP to monitor the conformational remodeling of αS fibrils by small polyphenolic compounds and demonstrate that site-specific labels yield insight into their mechanism of action.

§ 4.4 Fluorescence Polarization Results and Discussion

In order to produce site-specifically labeled αS for FP studies, we expressed constructs containing single Cys mutations at positions 9, 114, and 136. Whereas positions 9 and 136 can be used to monitor local dynamics of the protein termini, position 114 was selected for its proximity to the fibril core. Following disulfide bond reduction by TCEP, Cys labeling was performed by incubating the protein with an excess of Fam for 8 hours at 37 °C. HPLC purified proteins were characterized by MALDI-MS and PAGE analysis. MALDI-MS of trypsin digestion fragments confirmed that the labeling of αS proceeded in a quantitative and highly specific manner (See Section 4.7, Materials and Methods).

We began our studies by measuring the polarization values associated with each labeled protein in its monomeric state. In order to avoid intermolecular fluorophore

interactions (such as homoFRET, which may lead to a decrease in polarization[245]) labeled αS was diluted into a solution of WT αS in a 1:99 molar ratio for all FP measurements. Since αS is intrinsically disordered, we expected that FP values corresponding to monomeric αS would be similar for each labeled mutant. Intriguingly, we found that the polarization of the Fam label was dependent on its location within the αS sequence. Whereas the FP values corresponding to αS-C^{Fam}_{9} and αS-C^{Fam}_{114} were similar, a lower FP value was observed for αS-C^{Fam}_{136} (Figure 4.7). These results are consistent with NMR studies suggesting that the residues in the C-terminal tail exhibit greater conformational freedom in comparison to those residing in the N-terminal and NAC domains.[246]

To further demonstrate that site-specific labels can yield information on local dynamics, we obtained FP measurements of each αS mutant bound to sodium dodecyl sulfate (SDS) micelles. Previous NMR and fluorescence studies have shown that, upon binding to SDS, the first 98 residues of αS fold into a pair of curved antiparallel helices that are connected by a short linker region spanning residues 38-44.[82,247] The helical pair is followed by a predominantly disordered C-terminal tail consisting of residues 98-140.[247] In the presence of SDS, the FP value of αS-C^{Fam}_{9} increased, consistent with the observation that the N-terminal domain of αS folds on the surface of SDS micelles. FP values of αS-C^{Fam}_{114} and αS-C^{Fam}_{136} show a modest decrease upon binding to SDS micelles, indicating that the C-terminal domain is somewhat more dynamic in the SDS-bound conformation. In the disordered ensemble, the C-terminal domain participates in long-range intramolecular contacts that result in compaction of the protein.[65,66,68] Therefore, the observed decreases in FP for αS-C^{Fam}_{114} and αS-C^{Fam}_{136} likely result from

release of these interactions on binding of the N-terminal and NAC domains to SDS micelles, so that the C-terminus moves freely in bulk solvent.

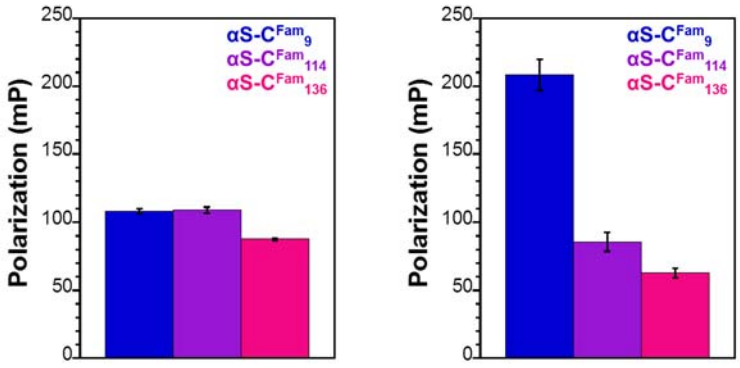

Figure 4.7 Fluorescence Polarization Values of Monomeric Labeled αS.
Left: FP values of Fam-labeled αS in buffer. Right: FP values of monomeric αS in the presence of 1 mM SDS.

Next, we used FP to monitor the assembly of αS into amyloid fibrils. Aggregation reactions were carried out by diluting each labeled construct into a mixture of WT αS in a 1:99 molar ratio and agitating the solution at 1500 rpm. Periodically, separate aliquots of the aggregation reaction were removed for FP and CR analysis. CR binding assays demonstrated that the kinetic profiles of aggregation assays containing fluorescently-labeled αS are indistinguishable from those obtained using WT alone (Figure 4.9). Both FP and CR measurements generated sigmoidal curves typical of a nucleation-dependent mechanism of amyloid aggregation. However, the kinetic profiles obtained using FP varied based on the position of the incorporated label (Figure 4.8, see Section 4.7, Materials and Methods for raw data). Although further studies will be required to fully understand this phenomenon, these preliminary results suggest that rigidification of the C-terminus occurs concurrenty with overall αS aggrgegation and that rearrangement of the N-terminal domain is slightly delayed.

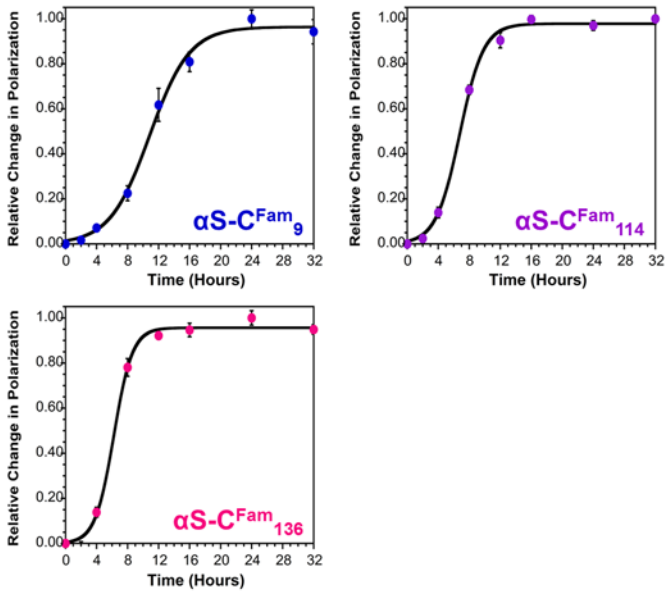

Figure 4.8 Aggregation of αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$ and αS-C$^{Fam}_{136}$ Monitored by Fluorescence Polarization.
Aggregation reactions containing 1% labeled αS in a solution of WT αS were performed in triplicate. At the indicated time points, aliquots were removed from the aggregation reaction, diluted 10-fold in buffer, and assessed using FP. The kinetic profiles of the aggregation reactions varied depending on the position of the incorporated label.

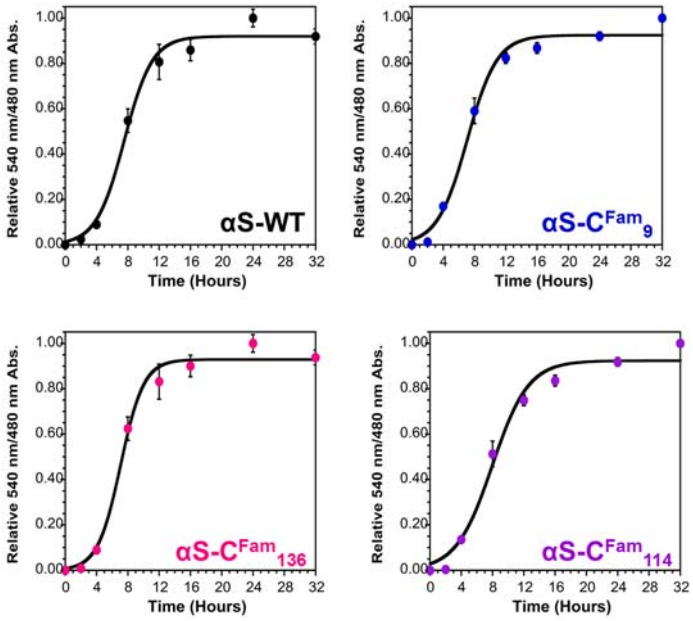

Figure 4.9 Aggregation of αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$ and αS-C$^{Fam}_{136}$.
Monitored by CR Binding. Aggregation reactions containing 1% labeled αS in a solution of WT αS were performed in triplicate. At the indicated time points, aliquots were removed from the aggregation reaction, diluted 20-fold in 20 μM CR, and assessed by UV-Vis absorbance. The kinetic profiles of the aggregation reactions were similar for WT and each labeled mutant.

αS Variant	FP $t_{1/2}$ (h)	CR $t_{1/2}$ (h)
αS-WT	--	7.53 ± .4
αS-C^{Fam}_{9}	10.8 ± .3	7.06 ± .4
αS-C^{Fam}_{114}	6.77 ± .2	7.96 ± .5
αS-C^{Fam}_{136}	6.22 ± .2	7.07 ± .3

Figure 4.10 Summary of $t_{1/2}$ Values Determined by FP and CR. Normalized data were fit to a sigmoidal expression to estimate the $t_{1/2}$ of aggregation associated with each mutant (See Section 4.7, Materials and Methods for details).

Having shown that FP can be used to monitor amyloid assembly, we next chose to examine whether we could use this technique to monitor the disaggregation of αS fibrils. αS fibrils containing each labeled construct were agitated for 36 hours and separated from the aggregation reaction by centrifugation. Freshly re-suspended fibrils were then treated with excess SDS, boiled for 5 min, and allowed to cool to room temperature. Following the disaggregation procedure, the FP values uniformly decreased to those corresponding to the SDS-bound conformation (Figure 4.11). Sedimentation PAGE analyses confirmed that αS fibrils were efficiently dissolved following treatment with SDS (data not shown). These studies demonstrate that we can use FP to monitor conformational changes in αS fibrils without the addition of any exogenous small molecule probe.

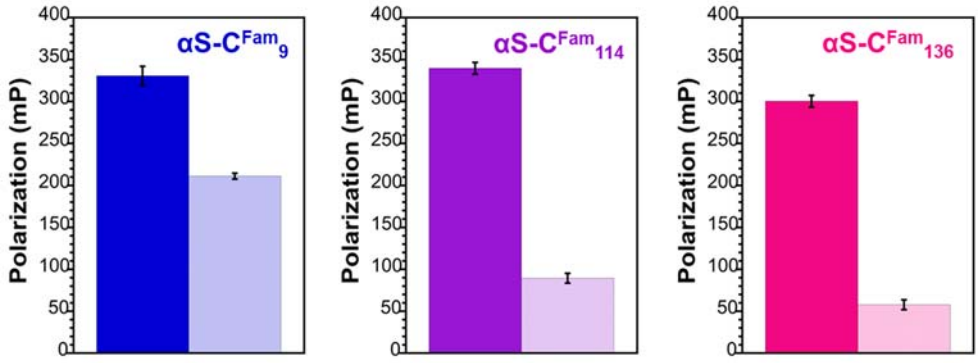

Figure 4.11 Disaggregation of αS-C^{Fam}_{9}, αS-C^{Fam}_{114} and αS-C^{Fam}_{136} Fibrils Monitored by FP. The darker bars (on the left of each graph) represent the FP values corresponding to αS fibrils containing 1% of the labeled protein.The lighter bars (on the right of each graph) represent the FP values after boiling in excess SDS for 5 min.

Several small molecules have been shown to remodel and/or disaggregate pre-formed αS fibrils. However, the molecular details underlying these structural changes are poorly understood. Dopaminergic nerve cell death is one of the pathological hallmarks of PD.[234] Thus, several groups have examined the effect of dopamine (Dop) and its oxidized derivatives on the assembly of αS.[248-252] In 2004, Li *et al*. reported that dopamine treatment led to the disaggregation of pre-formed αS fibrils.[235] In this study, the disaggregation of αS was monitored using ThT fluorescence, dynamic light scattering (DLS), atomic force microscopy (AFM), and EM. Surprisingly, the Dop-mediated disaggregation of αS fibrils has not been described in the literature since this initial report.

In the following years, several aromatic compounds containing vicinal hydroxyl groups have been identified as efficient amyloid remodeling agents. Among the most well-studied examples is epigallocatechin gallate (EGCG), a flavonoid that is found in high concentrations in green tea. In 2008, Bieschke *et al*. demonstrated that EGCG transforms αS fibrils into smaller species that are nontoxic to mammalian cells.[237] Time-resolved EM and AFM imaging revealed that EGCG converts αS fibrils into poorly defined amorphous aggregates. CD spectroscopy experiments showed a progressive loss of β-sheet content in the presence of EGCG, demonstrating that binding of the compound lead to a significant alteration of αS secondary structure. ThT binding measurements were consistent with decreasing amyloid content. However, EGCG and ThT may compete for similar binding sites, and these studies do not shed any light on local structural changes induced by EGCG.

Mature αS fibrils have a tendency to enhance the rate of αS assembly in a process that is commonly referred to as seeding. EGCG-treated fibrils were unable to seed the aggregation of αS *in vitro*, suggesting that the compound may reduce fibril toxicity *in vivo*. HEK-293 cells overexpressing WT αS were transfected with pre-formed αS fibrils and evaluated using a standard lactate dehydrogenase (LDH) assay after three days. Whereas transfection with αS fibrils led to a 2.5-fold increase in LDH activity, no such effect was observed when the cells were treated with EGCG. This result was reproducible in an alternative pheochromocytoma (PC12) cell model. Overall, these studies reveal that EGCG remodels αS fibrils in a manner that leads to a significant reduction of their inherent pathogenicity.

Having shown that we can use FP to monitor αS folding, aggregation, and disaggregation, we reasoned that we could use this technique to gain insight into the mechanism by which small molecules remodel αS fibrils. Here, we chose to examine the remodeling capacity of dopamine, EGCG, and an additional flavonoid, nordihydroguiaretic acid (NDGA). Previously, NDGA has been shown to inhibit αS aggregation and to disassemble αS oligomers that form in the presence of $FeCl_3$.[241] Prior to this work, the effect of NDGA treatment on the structure of αS fibrils has not been described.

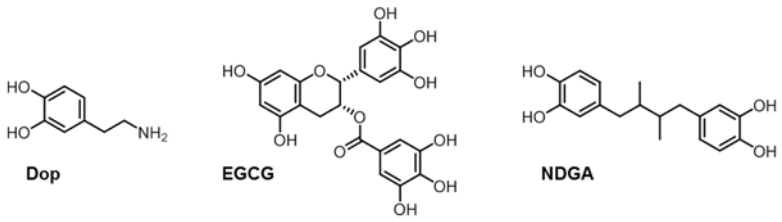

Figure 4.11 Chemical Structures of Dop, EGCG, and NDGA.

Pre-formed fibrils containing each labeled mutant were treated with Dop, EGCG, and NDGA and continuously monitored using FP (Figure 4.12). *Note: At this time, these experiments have not been performed with αS-C^{Fam}_{136}.* Prolonged incubation with stoichiometric or excess amounts of Dop had no effect on the observed FP values for any of the labeled mutants tested. Treatment of pre-formed fibrils with stoichiometric quantities of EGCG also failed to produce any significant changes in FP. When a 10-fold molar excess EGCG was added to pre-formed fibrils containing αS-C^{Fam}_{114}, we observed a significant decrease in FP over time. However, only a small change in FP was observed when αS-C^{Fam}_{9} fibrils were treated with excess EGCG. Together, these studies suggest that EGCG binds to and preferentially remodels the C-terminus of αS. These data are consistent with earlier reports suggesting that EGCG preferentially interacts with the C-terminus of αS monomers.[253,254] Remarkably, treatment with excess NDGA produced a significant change in FP for both αS-C^{Fam}_{9} and αS-C^{Fam}_{114}-containing fibrils. These preliminary results suggest that NDGA is a potent amyloid remodeling agent.

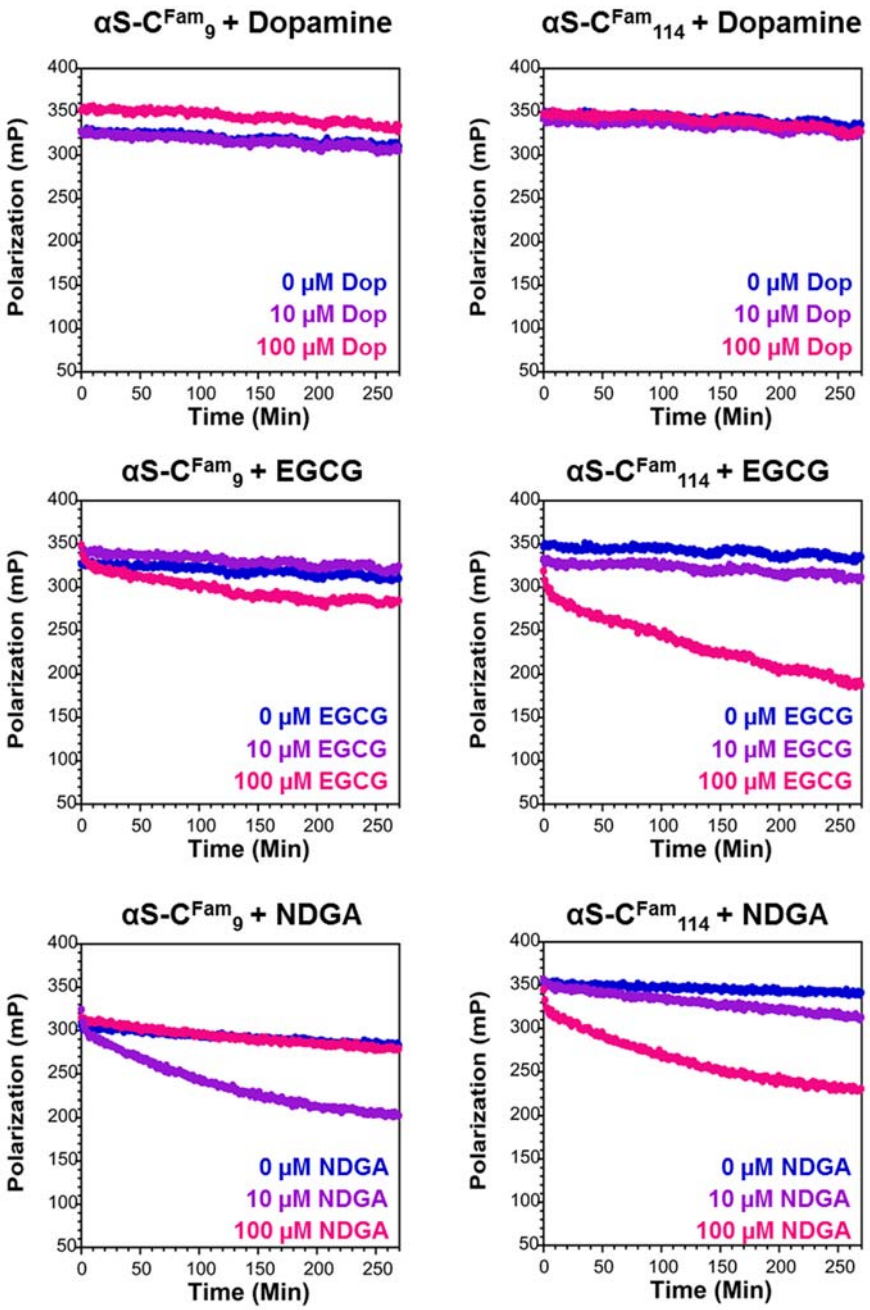

Figure 4.12 Small Molecule Remodeling of αS-C$^{Fam}_{9}$ and αS-C$^{Fam}_{114}$ Monitored by FP. Dopmaine, EGCG, and NDGA were added to pre-formed αS fibrils (10 μM, relative to monomer concentration) and continuously monitored by FP.

§ 4.5 Conclusions

Since αS and other amyloidogenic proteins are generally not amenable to traditional structural studies, many laboratories are interested in functionalizing these

proteins in order to study the aggregation process *in vitro* and *in vivo*. Here, we have employed several strategies to multifunctionalize αS with fluorophores using various combinations of conventional Cys modifications, amber codon suppression, AaT mediated N-terminal modification, and NCL. The combinations of techniques considered here fulfill the requirements of being either mutually orthogonal or able to be carried out sequentially. We have shown that several combinations are practical for site-specific double labeling of αS. Cys modification can be used in combination with amber codon suppression, and this pair of techniques is applicable to αS, but may not be feasible in proteins with multiple Cys residues. Since AaT modification takes place post-translationally, when one uses it in combination with amber suppression, one can use the same types of bioorthogonal reactions and control the position of the labels by reaction sequence.

A variety of *in vitro* and *in vivo* experiments can be performed using fluorescently-labeled αS. Here, we have shown that site-specifically labeled αS can be used in conjunction with FP to monitor various processes including SDS-induced folding, fibril formation, disaggregation, and fibril remodeling induced by exogenous small molecules. Additionally, we have shown that site-specific Fam labels yield local dynamic information. Although other groups have shown that certain small molecules can remodel αS fibrils into amorphous, non-toxic species, the molecular basis for this effect is poorly understood. Our preliminary results suggest that EGCG and NDGA preferentially bind to and remodel the C-terminus of αS. In the future, we will combine FP with FRET studies using doubly-labeled proteins to characterize these conformational changes in rigorous

detail. Additionally, we will continue to use FP to screen synthetic derivatives of NDGA in efforts to discover new compounds with enhanced therapeutic potential.

§ 4.6 Materials and Methods

General Information. Azf was purchased from Bachem (Torrance, CA). 5(6)'-carboxytetramethylrhodamine was purchased from Novabiochem, EMD Millipore (Darmstadt, Germany). Fluorescein-5-maleimide was purchased from TCI America (Portland, Oregon). TAMRA-azide, 5-isomer was purchased from Lumiprobe (Hallandale Beach, Florida). Dibenzylcyclooctyne-amine and carboxyrhodamine110-azide were purchased from ClickChemistryTools (Scottsdale, AZ). Ni-NTA resin was purchased from Qiagen (Valencia, CA). *E. coli* BL21(DE3) cells were purchased from Stratagene (La Jolla, CA). (-)-Epigallocatechin gallate (EGCG) was purchased from Caymen Chemical (Ann Arbor, Michigan). Nordihydroguaiaretic acid (NDGA) was purchased from Sigma-Aldrich (St. Louis, MO). Sequencing-grade trypsin was purchased from Promega (Madison, WI). Bovine Factor Xa protease was purchased from Haematologic Technologies (Esses Junction, VT). QuikChange® site-directed mutagenesis kits were purchased from Stratagene. DNA oligomers were purchased from Integrated DNA Technologies, Inc (Coralville, IA). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA). MilliQ filtered (18 MΩ) water was used for all solutions (Millipore, Billerica, MA). The pEG6 plasmid, containing His_{10}-tagged *E. coli* AaT, was a gift from Alexander Varshavsky (California Institute of Technology). The pDULE2 plasmid was a gift from Ryan Mehl (Oregon State University). Matrix-assisted laser desorption/ionization (MALDI) mass spectra were collected with a Bruker Ultraflex III MALDI-TOF/TOF mass spectrometer (Billerica, MA). UV/vis absorbance spectra were obtained with a Hewlett-Packard 8452A diode array spectrophotometer (Agilent

Technologies, Santa Clara, CA). NMR spectra, ^{1}H and ^{13}C, were collected with a Bruker DRX 500 MHz instrument (Billerica, MA). Gel images were obtained with a Typhoon FLA 7000 (GE Lifesciences, Princeton, NJ). Fluorescence spectra were collected with a Tecan M1000 plate reader. Fluorescence polarization data were collected with a Tecan F200 plate reader (Mannedorf, Switzerland).

Construction of pRK αS Cys Mutant Expression Plasmids. A plasmid containing the human wild-type αS gene cloned between NdeI and HindIII in the expression vector pRK172 was provided by Dr. Virginia Lee (Perelman School of Medicine, University of Pennsylvania). QuikChange® mutagenesis was used to mutate Ser9 to Cys9, Glu114 to Cys114 and Tyr136 to Cys136 in the pRK172 construct to yield pRK172-αS-C_9, pRK172-αS-C_{114} and pRK172-αS-C_{136}, respectively. The sequence of the mutant plasmid was confirmed by DNA sequencing analysis. DNA primers for each mutagenesis step are shown in Figure 4.13.

Construction of pT-T7 αS TAG Expression Plasmids. A plasmid containing the human wild-type αS gene cloned between NdeI and HindIII in pT7-7 vector was provided by Dr. Elizabeth Rhoades (Department of Molecular Biophysics and Biochemistry, Yale University). Higher yields were observed for Uaa-containing proteins when expressed from the pT-T7 vector in comparison to Uaa expression from αS pRK vectors. Thus, all TAG-containing mutant constructs were generated in pT-T7. QuikChange® mutagenesis was used to mutate Tyr39 to TAG, and Phe94 to TAG in the pT-T7 construct to yield pT-T7-αS-TAG_{39}, and pT-T7-αS-TAG_{94} respectively. A second round of QuikChange® mutagenesis reactions was performed to mutate pT-T7TAG_{94} to

pT-T7-αSC$_9$TAG$_{94}$ and pT-T7-αS-TAG$_{39}$ to pT-T7-αS-TAG$_{39}$-C$_{114}$. DNA primers for each mutagenesis step are shown in Figure 4.13.

DNA Oligomers Used for αS QuikChange® Mutagenesis.

i. Mutation S9C in αS

Forward: 5' – GTATTCATGAAAGGACTTTGCAAGGCCAAGGAGGGAGTTG – 3'

Reverse: 5' – CAACTCCCTCCTTGGCCTTGCAAAGTCCTTTCATGAATAC – 3'

ii. Mutation E114C in αS

Forward: 5' – CCCCACAGGAAGGAATTCTGTGCGATATGCCTGTGGATCCTGA – 3'

Reverse: 5' – TCAGGATCCACAGGCATATCGCACAGAATTCCTTCCTGTGGGG – 3'

iii. Mutation Y136C in αS

Forward: 5' – TCTGAGGAAGGGTATCAAGACTGCGAACCTGAAGC – 3'

Reverse: 5' – GCTTCAGGTTCGCAGTCTTGATACCCTTCCTCAGA – 3'

iv. Mutation Y39TAG in αS

Forward: 5' – AAAAGAGGGTGTTCTCTAGGTAGGCTCCAAAACCAA – 3'

Reverse: 5' – TTGGTTTTGGAGCCTACCTAGAGAACACCCTCTTTT – 3'

v. Mutation F94TAG in αS

Forward: 5' – GCATTGCAGCAGCCACTGGCTAGGTCAAAAAGGACCAGTTGGG – 3'

Reverse: 5' – CCCAACTGGTCCTTTTTGACCTAGCCAGTGGCTGCTGCAATGC – 3'

Figure 4.13 DNA Oligomers Used for Quikchange® Mutagenesis.

Overexpression and Purification of Full-length WT αS and αS Cys Mutants. pRK-αS**,** pRK-αS-C$_9$, prK-αS-C$_{114}$, and pRK-αS-C$_{136}$ were transformed into competent *E. coli* BL21(DE3) cells. Single colonies were used to inoculate 5 mL of LB media supplemented with ampicillin (Amp, 100 μg/mL). The primary culture was incubated at 37 °C with shaking at 250 rpm for 4 h. The primary culture was used to inoculate 1 L of LB media containing Amp (100 μg/mL) which was then grown overnight at 37 °C with shaking at 250 rpm. The cells were harvested by centrifugation at 5000 x g for 15 min,

and the resulting pellet was re-suspended in lysis buffer (40 mM tris(hydroxymethyl)aminomethane (Tris), 5 mM ethylenediaminetetraacetic acid (EDTA), pH 8.2) supplemented with 1 mM phenylmethanesulfonyl fluoride (PMSF) and 10 units/mL DNase I – Grade II. The cells were lysed by sonication, boiled for 20 min at 100 °C, and centrifuged for 20 min at 14,200 x g at 4 °C. The clear supernatant was dialyzed against purification buffer (20 mM Tris, pH 8.0) overnight at 4 °C. The resulting solution was purified by gel filtration over a Superdex 75 16/600 column followed by ion-exchange chromatography using a HiTrap Q HP column (5 mL) on an ÄKTA FPLC using a 100 min NaCl gradient (0 to 500 mM NaCl in 20 mM Tris, pH 8.0). The fractions containing the product were identified by MALDI MS, pooled, and dialyzed against αS buffer (20 mM Tris, 100 mM, pH 7.5) overnight.

Labeling of αS-C_9, αS-C_{114}, and αS-C_{136} with Fam. Prior to performing the labeling reactions, the concentration of each protein (αS-C_9, αS-C_{114}, and αS-C_{136}) was adjusted to approximately 1.0 mg/mL and incubated with αS buffer containing TCEP (1 mM) for 10 min at room temperature. Fluorescein-5-maleimide (5 equiv) was then added to the protein from a fresh 10 mM stock solution prepared in DMSO. The reaction was carried out at 37 °C for 4 h without shaking. After the first 4 h, an additional 5 equiv were added and the reaction was allowed to incubate for an addition 2-4 hours at 37 °C until quantitative conversion was observed by MALDI MS. The resulting fluorescein-labeled protein was dialyzed against purification buffer (20 mM Tris, pH 8.0) and purified over a HiTrap Q HP column on an ÄKTA FPLC using a 100 min NaCl gradient (0 to 500 mM NaCl in 20 mM Tris, pH 8.0). The fractions containing the product were identified by MALDI MS and dialyzed against water overnight at 4 °C. The dialyzed, labeled proteins

were further purified by reverse-phase HPLC on a Vydac 218TP C4 semi-prep column using the following gradient: isocratic at 95% aqueous phase for 5 min, and then ranging from 95% to 75% aqueous phase over 5 min, then to 40% over 20 min, then to 0% aqueous phase over 5 min, then returning to 95% aqueous phase during a 5 min wash out period. The purified protein was then concentrated using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column, exchanged into αS buffer, and stored at – 80 °C until further use.

Mass Spectrometry Analysis of Fam Labeling Reactions. Fam labeling was periodically monitored using MALDI MS. Complete labeling was consistently observed within 8 hours.

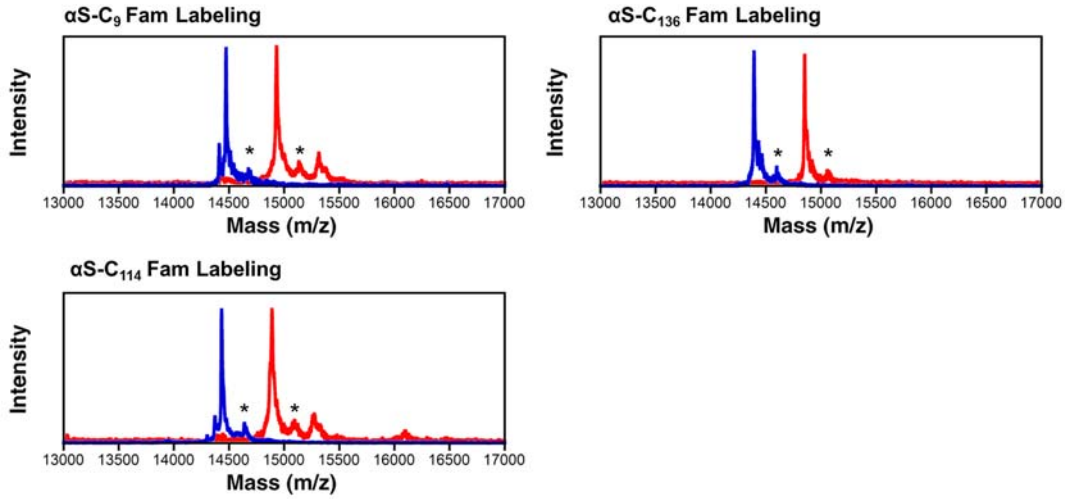

Figure 4.14 MALDI MS Traces of Crude Fam Labeling Reactions.
Blue trace: Cys mutant prior to the labeling reaction. Red trace: Crude reaction mixture post labeling with 10 equiv of Fam maleimide. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

MALDI Analysis of HPLC Purified αS-C^{Fam}_9, αS-C^{Fam}_{114} and αS-C^{Fam}_{136}. MALDI MS of the purified proteins consistently yielded masses that were 15-20 Da higher than the expected mass. MALDI MS analysis of trypsin fragments suggest that this increase in mass corresponds to hydrolytic ring-opening of the maleimide group (see below).

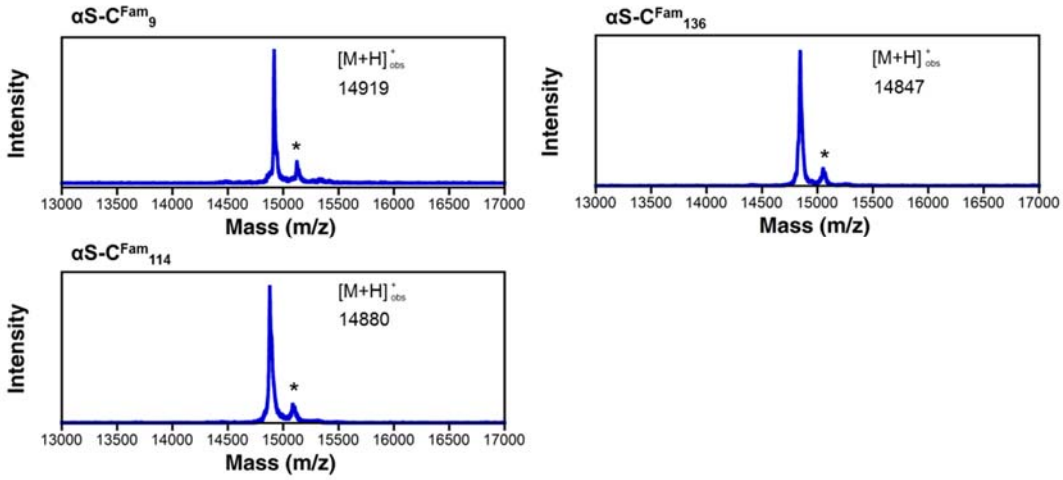

Figure 4.15 MALDI Traces of HPLC Purified αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$ and αS-C$^{Fam}_{136}$. Expected $[M+H]^+$ of αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$, and αS-C$^{Fam}_{136}$ are 14906, 14864, and 14830, respectively. Expected $[M+H]^+$ of maleimide ring-opened αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$, and αS-C$^{Fam}_{136}$ are 14924, 14882, and 14848, respectively. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Trypsin Digestion of αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$ and αS-C$^{Fam}_{136}$. αS mutants (10−50 μg) were incubated with 5−10 μL aliquots of sequencing-grade modified trypsin (0.1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 7.5, at 37 °C for 4 h. An aliquot (1.0 μL) of the digestion reaction was taken and analyzed by MALDI MS.

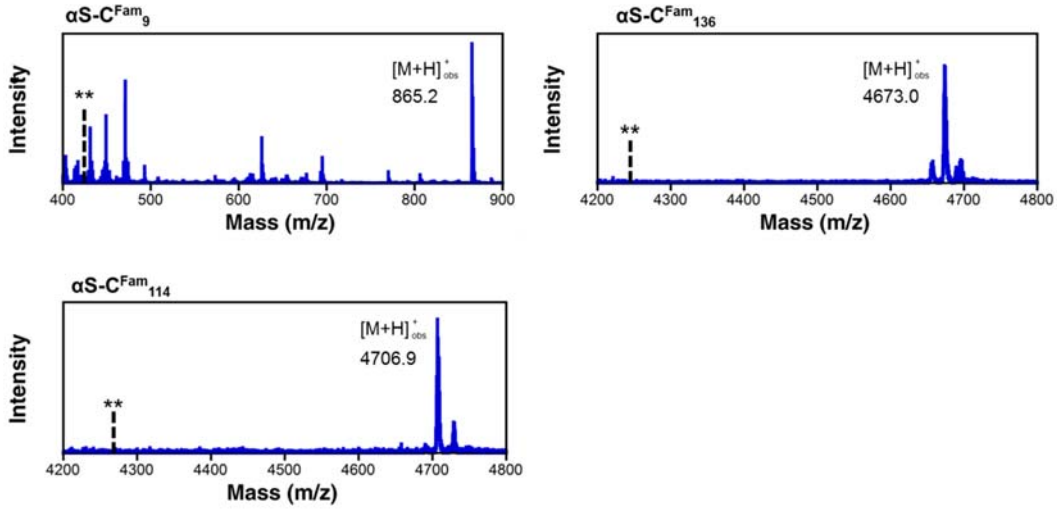

Figure 4.16 MALDI MS of Trypsin Fragments of αS-C$^{Fam}_{9}$, αS-C$^{Fam}_{114}$ and αS-C$^{Fam}_{136}$. Expected $[M+H]^+$ of αS-C$^{Fam}_{9}$ Frag. 7-10, αS-C$^{Fam}_{114}$ Frag. 103-140 and αS-C$^{Fam}_{136}$ Frag. 103-140 are 847.3, 4687.6, and 4652.8, respectively. Expected $[M+H]^+$ of maleimide ring-opened αS-C$^{Fam}_{9}$ Frag. 7-10, αS-C$^{Fam}_{114}$ Frag. 103-140 and αS-C$^{Fam}_{136}$ Frag. 103-140 are 865.3, 4705.6, and 4670.1, respectively. The double asterisk (**) corresponds to the unmodified trypsin fragment.

MALDI MS analysis of the trypsin fragments show that Fam labeling is site-specific and quantitative. In each labeled fragment, a mass adduct of approximately +18 Da was observed. This increase in mass is consistent with hydrolytic ring-opening of the maleimide group.

Overexpression and Purification of αS-Z$_{94}$, αS-π$_{39}$, αS-π$_{94}$, αS-C$_9$π$_{94}$ and αS--π$_{39}$C$_{114}$. pT-T7-αS-TAG$_{39}$ and pT-T7-αS-TAG$_{94}$ were transformed into competent *E. coli* BL21(DE3) cells harboring the orthogonal tRNA$_{CUA}$ and pXF-tRNA synthetase pair. Transformed cells were selected on the basis of Amp and streptomycin (Strep) resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100 ug/mL) and Strep (100 ug/mL). The primary 5 mL culture was incubated at 37 ºC with shaking at 250 rpm for 5 hours. The primary culture was used to inoculate 1 L of a variant of M9 minimal media. To an autoclaved 1L solution containing 6 g Na_2HPO_4, 3 g KH_2PO_4, .5 g NaCl and 1 g NH_4Cl, the following autoclaved solutions were added: 1 mL of 2 M $MgSO_4$, 1 mL of 15 mg/mL $FeCl_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL $ZnCl_2$ (in acidified H_2O), 2 mL of 10% Bacto™ Yeast Extract 12.5 mL 40% glucose (w/v), and 1 µL of 1 M $CaCl_2$. When the OD_{600} of the secondary culture reached 0.8, the corresponding Uaa was added (220 mg for Ppy or 206 mg for AzF, final concentration of 1 mM), and the culture was incubated overnight at 37 ºC with shaking at 250 rpm. The cells were harvested at 5000 x g for 15 min and the resulting pellet was resuspended in 20 mM Tris, pH 8. Following sonication, the cell lysate was boiled for 20 minutes prior to centrifugation for 20 minutes at 30,000 x g, 4 ºC. The cleared lysate was dialyzed against 20 mM Tris pH 8.0 overnight. Next, the lysate was loaded onto a HiLoad 16/600 Superdex 75 column and eluted with 20 mM Tris, pH 8. FPLC fractions from size-

exclusion chromatography were analyzed by SDS PAGE. The fractions containing αS were then loaded onto a HighTrap Q HP column and eluted over a 100 minute sodium chloride gradient (0 to 0.5M NaCl in 20 mM Tris, pH 8). Following FPLC purification of αS-Z_{94}, αS-π_{39}, and αS-π_{94}, the fractions were analyzed by MALDI-MS and pooled prior to click labeling. Following FPLC purification of αS-$C_9\pi_{94}$ and αS-$\pi_{39}C_{114}$, the proteins were dialyzed into αS buffer overnight for Fam labeling.

Click Labeling of αS-π_{39} and αS-π_{94} with Raz. Prior to performing the labeling reactions, the concentration of FPLC-purified αS-π_{39} and αS-π_{94} was adjusted to 1.0 mg/mL. The catalyst mixture was separately prepared by combining (for each mL of reaction) 1.25 μL of 80 mM Cu_2SO_4, 30 μL of 50 mM Tris-(hydroxypropyltriazolylmethyl)amine (THPTA) and 30 μL of 100 mM sodium ascorbate. This mixture was allowed to incubate at room temperature for 5 min. After adding 5 equiv of 5-TAMRA (Raz) azide to the protein solution, the Cu(I)-containing catalyst mixture was added, and the reaction was allowed to incubate for 4 hours at 37 °C. Following the labeling reaction, excess dye was removed by performing buffer exchange using using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column. Next, the protein was purified by reverse-phase HPLC as previously described. The purified protein was then concentrated, exchanged into αS buffer, and stored at – 80 °C until further use.

Mass Spectrometry Analysis of Raz Labeling Reactions. Click labeling was periodically monitored using MALDI MS. Following treatment with a total of 5 equivalents of Raz azide in the presence of Cu(I), near quantitative labeling was observed within 4 h.

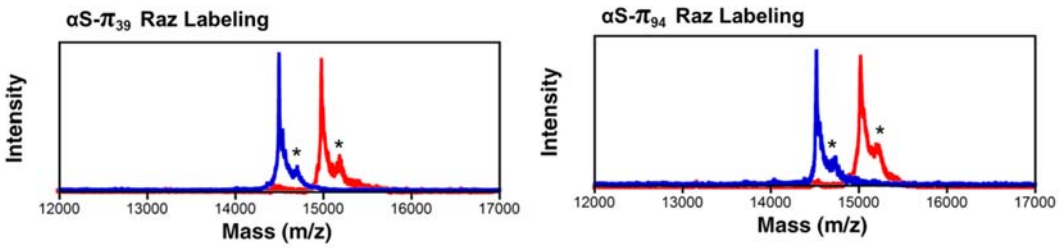

Figure 4.17 MALDI Traces of Crude Raz Labeling Reactions.
Blue trace: Cys mutant prior to the labeling reaction. Red trace: Crude reaction mixture post labeling with 5 equiv of Raz in the presence of Cu(I). The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

MALDI Analysis of Purified αSπ$^{Raz}_{39}$ and αSπ$^{Raz}_{94}$.

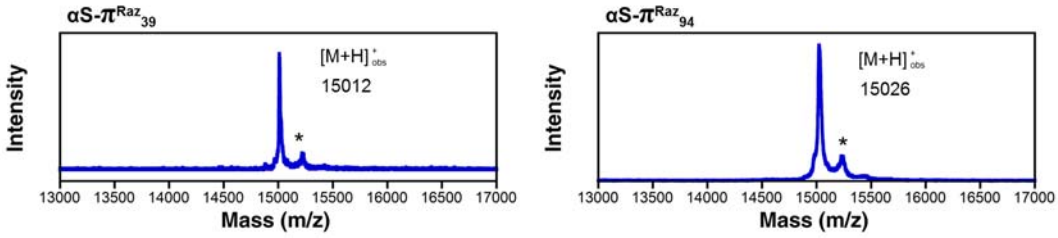

Figure 4.18 MALDI Traces of HPLC Purified αS-π$^{Raz}_{39}$ and αS-π$^{Raz}_{94}$.
Expected [M+H]$^+$ of αS-π$^{Raz}_{39}$ and αS-π$^{Raz}_{94}$ are 15013 and 15029, respectively. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Trypsin Digestion of αSπ$^{Raz}_{39}$ and αSπ$^{Raz}_{94}$. Trypsin digestion was performed as previously described. MALDI MS analysis of the trypsin digestion fragments show that Raz labeling is quantitative and site-specific.

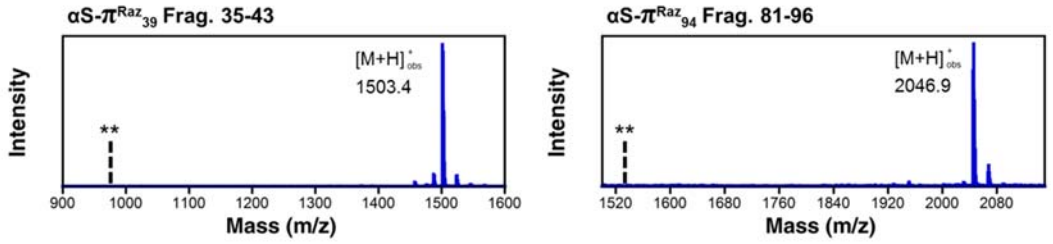

Figure 4.19 MALDI MS of Trypsin Fragments of αS-π$^{Raz}_{39}$ and αS-π$^{Raz}_{94}$.
Expected [M+H]$^+$ of αS-π$^{Raz}_{39}$ Frag. 35-43 and αS-π$^{Raz}_{94}$ Frag 81-96 are 1503.8 and 2047.0, respectively. The double asterisk (**) corresponds to the unmodified trypsin fragment.

Click Labeling of αS-Z_{94} with Tco and Raz. Prior to performing the labeling reactions, the concentration of FPLC-purified αS-Z_{94} was adjusted to 1.0 mg/mL. αS-Z_{94} was incubated with 5 equivalents of Tco for 4 h at 37 °C. Following the 4 h incubation period, incomplete labeling was observed by MALDI MS. After adding additional equivalents of Tco, the labeling yield did not improve. Incomplete labeling is likely due to reduction of Z to the unreactive amine during protein expression and purification. Raz labeling of αS-Z_{94} was performed as described above. Although we observed a slightly higher modification yield by MALDI MS, the labeling reaction did not go to completion. Shown below are representative crude MALDI traces demonstrating that labeling of αS-Z_{94} using Tco or Raz fails to go to completion.

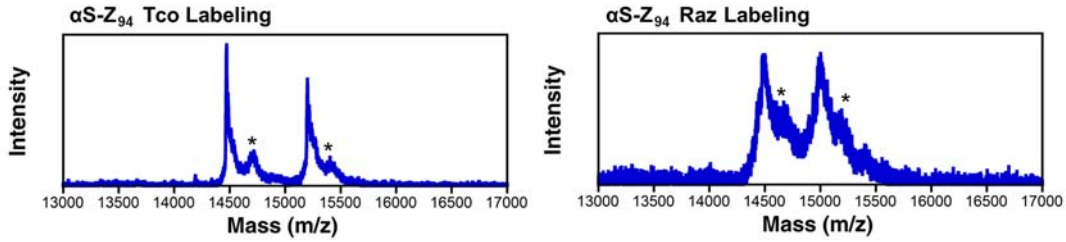

Figure 4.20 MALDI MS Analysis of Tco and Raz Labeling of αS-Z_{94}. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Double-Labeling of αS-$C_9\pi_{94}$ and αS--$\pi_{39}C_{114}$. First, αS-$C_9\pi_{94}$ and αS--$\pi_{39}C_{114}$ were labeled with Fam as described above. The resulting fluorescein-labeled proteins were then dialyzed against purification buffer (20 mM Tris, pH 8.0) and purified over a HiTrap Q HP column on an ÄKTA FPLC over a 120 min NaCl gradient (0 to 1 M NaCl in 20 mM Tris, pH 8.0). The fractions containing the product were identified by MALDI MS and pooled. Next, Cu-catalyzed click labeling with Raz azide was performed as described above. Following the click reaction, excess dye was removed by performing buffer exchange using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column. The doubly-

labeled proteins were purified by reverse-phase HPLC as previously described. The purified proteins were then concentrated, exchanged into αS buffer, and stored at – 80 °C until further use.

Mass Spectrometry Analysis of Double-Labeling. Each labeling step was monitored using MALDI MS. Near quantitative labeling was observed at each step for both reactions. Representative MALDI MS data corresponding to each labeling step of αS-$\pi_{39}C_{114}$ are shown below.

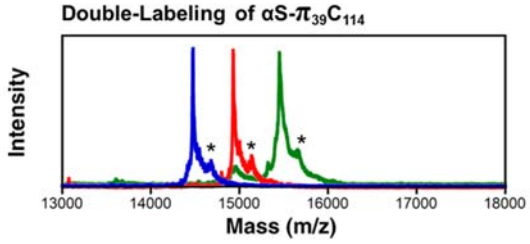

Figure 4.21 MALDI Traces of Crude Double Labeling Reactions.
Blue trace: αS-$\pi_{39}C_{114}$ prior to the labeling reaction. Red trace: Crude reaction mixture post labeling with 10 equiv of Fam maleimide. Green trace: Crude reaction mixture post labeling with 5 equiv of Raz azide in the presence of Cu(I). The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Mass Spectrometry Analysis of HPLC Purified αS-$C^{Fam}{}_9\pi_{94}$ and αS--$\pi^{Raz}{}_{39}C_{114}$.

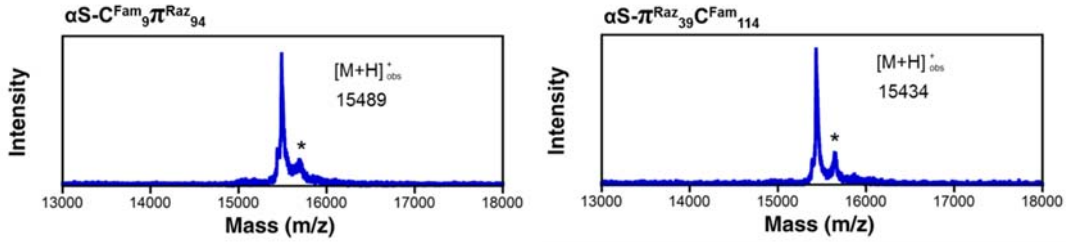

Figure 4.22 MALDI MS Traces of HPLC Purified αS-$C^{Fam}{}_9\pi^{Raz}{}_{94}$ and αS-$\pi^{Raz}{}_{39}C^{Fam}{}_{114}$.
Expected $[M+H]^+$ of αS-$C^{Fam}{}_9\pi^{Raz}{}_{94}$ and αS-$\pi^{Raz}{}_{39}C^{Fam}{}_{114}$. are 15474 and 15416, respectively. Expected $[M+H]^+$ of maleimide ring-opened αS-$C^{Fam}{}_9\pi^{Raz}{}_{94}$ and αS-$\pi^{Raz}{}_{39}C^{Fam}{}_{114}$. are 15492 and 15434, respectively. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Trypsin Digestion of αS-C$^{Fam}_9\pi_{94}$ and αS--$\pi^{Raz}_{39}C_{114}$. Trypsin digestion was performed as previously described. MALDI MS analysis of the trypsin digestion fragments show that Fam and Raz incorporation are site-specific and quantitative. In each Fam labeled fragment, a mass adduct of +18 Da was observed. This increase in mass is consistent with hydrolytic ring-opening of the maleimide group.

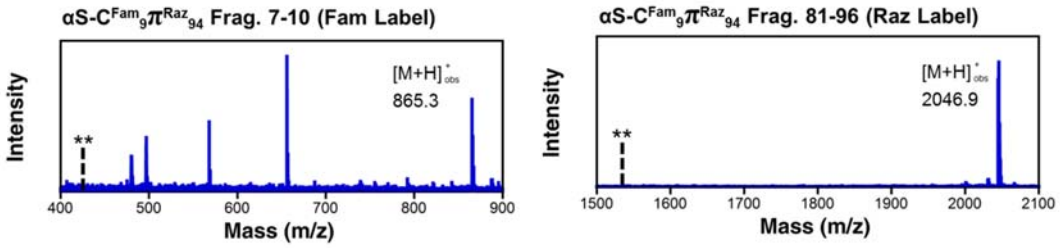

Figure 4.23 MALDI MS of Trypsin Fragments of αS-C$^{Fam}_9\pi^{Raz}_{94}$.
Expected [M+H]$^+$ of Frag. 7-10 and Frag 81-96 are 847.3 and 2047.0, respectively. The double asterisk (**) corresponds to the unmodified trypsin fragment.

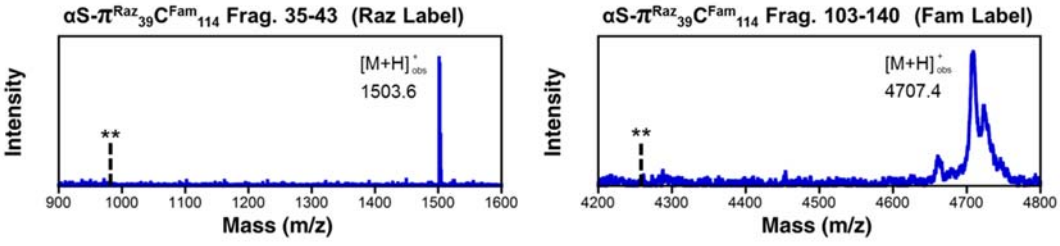

Figure 4.24 MALDI MS of Trypsin Fragments of αS-π^{Raz}_{39} C$^{Fam}_{114}$.
Expected [M+H]$^+$ of Frag. 35-43 and Frag 103-140 are 1503.8 and 4687.6, respectively. The double asterisk (**) corresponds to the unmodified trypsin fragment.

Synthesis αS-D′$_2$C$^{Fam}_9$. Protein fragment αS$_{9\text{-}140}$-C$_9$ and the synthetic thioester peptide Ac-αS$_{1\text{-}8}$-D′$_2$-SR were prepared in accordance with previously described protocols.[116,117] αS$_{9\text{-}140}$-C$_9$ (1 equiv, 0.10 μmol) was dissolved in 200 μL of degassed ligation buffer (6 M guanidinium hydrochloride, 200 mM sodium phosphate, 20 mM TCEP, 1% v/v thiophenol, pH 7.5) and incubated for 5 min. The reduced protein fragment was transferred to a microcentrifuge tube containing 2 equiv of the dried N-terminal peptide thioester Ac-αS$_{1\text{-}8}$-D′$_2$-SR, purged with argon, and allowed to incubate for 24 h at 37 ºC

with shaking at 600 rpm. MALDI MS analysis of the crude ligation reaction showed quantitative conversion to the full length product Ac-αSD′$_2$C$_9$ (see below). Following ligation, the reaction solution was dialyzed into αS buffer. Next, the ligated protein was labeled with Fam as previously described. MALDI MS analysis of the crude reaction showed nearly quantitative conversion to the Fam-containing product (see below). Following the labeling reaction, excess dye was removed by performing buffer exchange using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column. The doubly-labeled protein was purified by reverse-phase HPLC as previously described. The purified protein was then concentrated, exchanged into αS buffer, and stored at – 80 °C.

Mass Spectrometry Analysis of αS-D′$_2$C$^{Fam}_9$ Synthesis.

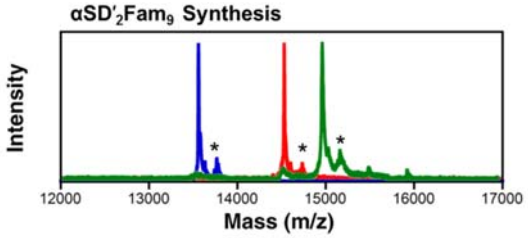

Figure 4.25 MALDI Traces of Crude NCL and Fam Labeling Reactions Towards αS-D′$_2$C$^{Fam}_9$. Blue trace: αS$_{9-140}$-C$_9$. Red trace: Formation of αS-D′$_2$C$_9$ following ligation of αS$_{9-140}$-C$_9$ with Ac-αS$_{1-8}$-D′$_2$-SR. Green trace: Post Fam labeling of αS-D′$_2$C$_9$ to generate αS-D′$_2$C$^{Fam}_9$. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Mass Spectrometry Analysis of HPLC Purified αS-D′$_2$C$^{Fam}_9$.

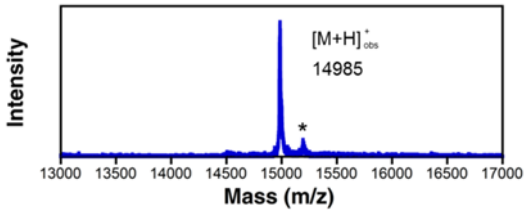

Figure 4.26 MALDI MS Trace of HPLC Purified αS-D′$_2$C$^{Fam}_9$.
Expected [M+H]$^+$ of αS-D′$_2$C$^{Fam}_9$ is 14962. Expected [M+H]$^+$ of maleimide ring-opened αS-D′$_2$C$^{Fam}_9$ is 14980. The asterisk (*) indicates a matrix adduct observed in all MALDI spectra of αS (+207 Da).

Fluorescence Polarization Measurements of αS-C^{Fam}_{9}, αS-C^{Fam}_{114} and αS-C^{Fam}_{136}. The concentrations of HPLC purified Fam-labeled proteins were determined by UV-Vis absorbance using a molar extinction coefficient of 68,000 $M^{-1}cm^{-1}$ at 494 nm. The concentration of WT αS was also determined using UV-Vis absorbance using a molar extinction coefficient of 5,600 $M^{-1}cm^{-1}$ at 277 nm. Prior to obtaining FP measurements, the Fam-labeled mutants were diluted into a solution of WT αS in a 1:99 molar ratio to achieve a final total protein concentration of 100 μM (1 μM Fam mutant to 99 μM WT). 100 μL of 10 μM samples were prepared in triplicate by diluting the concentrated stock into αS buffer (20 mM Tris, 100 mM NaCl, pH 7.5). Each sample was then gently vortexed and pipetted into a clear-bottomed, black-walled 96-well plate. FP measurements were obtained using a Tecan Infinite® F200 Pro microplate reader equipped with excitation (485 ± 20 nm) and emission (535 ± 25 nm) filters for Fam fluorescence.

Fluorescence Polarization Measurements of αS-C^{Fam}_{9}, αS-C^{Fam}_{114} and αS-C^{Fam}_{136} with SDS. In order to measure the FP values of each mutant in the presence of SDS, 10 μM samples containing 1 mM SDS were prepared in triplicate by combining 10 μL of the protein stock solution (prepared as described above) with 80 μL of aggregation buffer and 10 μL of 10 mM SDS. The samples were then gently vortexed, pipetted into a 96-well plate and immediately assessed on the microplate reader.

Aggregation Assays. Aggregation reactions were carried out by diluting each labeled construct into a mixture of WT αS in a 1:99 molar ratio to a final concentration of 100 μM as described above. Aggregation reactions were assembled in triplicate. Aggregation was initiated by shaking the solution at 1500 rpm on an IKA MS3 digital orbital shaker in

parafilm-sealed Eppendorf tubes. Periodically, aliquots were removed from the aggregation reaction and assessed by FP and CR absorbance in separate assays. For FP measurements, 10 µL of the aggregation solution was diluted to 100 µL using αS buffer. The samples were gently vortexed and transferred to a 96-well plate. For CR measurements, 10 µL of the aggregation solution was diluted to 200 µL using 20 µM CR dissolved in αS buffer. The samples were allowed to sit at room temperature for 15 minutes prior to being transferred to a 96-well plate and measured by FP and UV-Vis. Normalized data is shown in Section 4.5. Raw data corresponding to FP and CR measurements are shown below. Note that free fluorescein has an FP value of ~20 mP under the same conditions (data not shown).

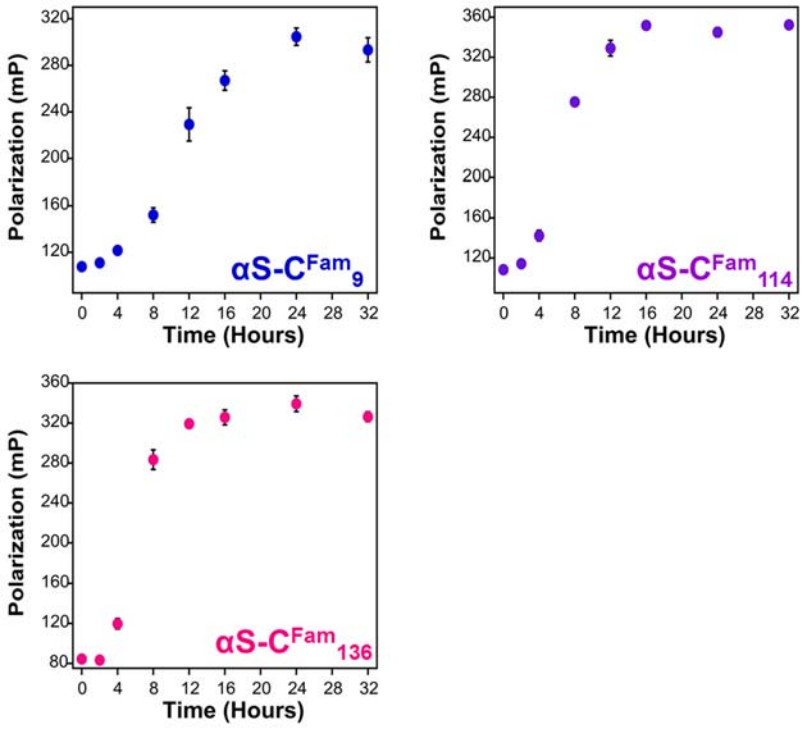

Figure 4.27 FP Measurements Corresponding to αS-C^{Fam}_{9}, αS-C^{Fam}_{114} and αS-C^{Fam}_{136} Aggregation.

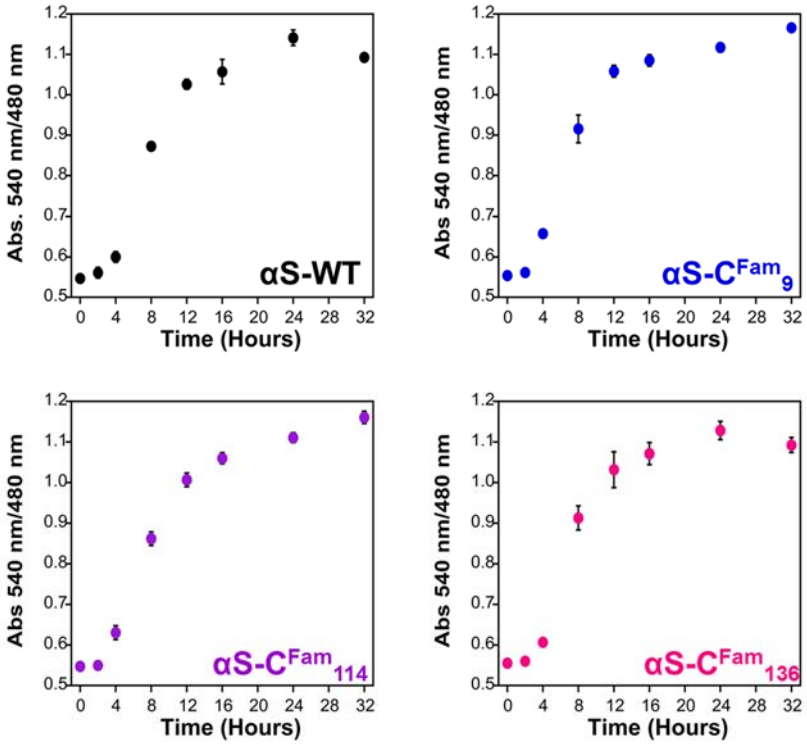

Figure 4.28 CR Absorbance Measurements Corresponding to αS-C^{Fam}_{9}, αS-C^{Fam}_{114} and αS-C^{Fam}_{136} and Wild-Type Aggregation Assays.

Estimation of $t_{1/2}$. The times corresponding to the half-maximum of the kinetics curves ($t_{1/2}$, based on FP and CR measurements) were determined by fitting the data to a sigmoidal equation (Eq. 4.2) of the form:

$$F(t) = m_1\left(\frac{1}{1+e^{\frac{t_{50}-t}{\sigma}}}\right)$$

in Kaleidagraph, where m_1 defines the upper baseline of the sigmoidal curve (normalized to 1) and σ represents a fitting parameter. Measurements of $t_{1/2}$ that correspond to each mutant are listed in Section 4.4.

SDS Disaggregation Assays. Fibrils for disaggregation assays were prepared by combining labeled αS in a 1:99 molar ratio with WT αS (300 µL total) and agitating the solution at 1500 rpm at 37 ºC for a total of 36 hours. Following the aggregation reaction, fibrils were isolated from the solution by centrifuging the samples for 1 h at 13,200 rpm

on a desktop centrifuge. After removing the supernatant, the fibrils were re-suspended in the same starting volume, such that the approximate protein concentration was 100 μM (relative to monomeric αS). The fibrils were re-suspended by gently pipetting αS buffer over the fibrils, and then vortexing at high speeds for 10 s intervals for a total of 4 times. Typically, FP values following re-suspension were between 310-350 mP units for each mutant. In order to obtain FP measurements of αS fibrils, 10 μL aliquots were removed from the re-suspended samples and diluted into 90 μL of αS buffer. Disaggregation reactions were performed by combining 10 μL of fibrils, 80 μL of αS buffer, and 10 μL of 10 mM SDS and boiling the solution for 5 min in parafilm-sealed Eppendorf tubes. The samples were allowed to cool to room temperature prior to being gently vortexed and transferred to a 96-well plate for FP measurements.

Small Molecule Remodeling Assays. Fibrils containing each labeled mutant were prepared for small molecule remodeling assays as described above. Stock solutions of 500 μM dopamine, 500 μM EGCG, and 500 μM NDGA were freshly prepared prior to performing each remodeling assay. NDGA stock solutions were prepared with 20% EtOH. Remodeling assays were performed in triplicate by treating three independent samples of 10 μM αS fibrils with stoichiometric (10 μM) or excess (100 μM) amounts of dopamine, EGCG, and NDGA. Prior to adding the small molecule solution to the αS fibrils, the solutions containing the fibrils were transferred to a 96-well plate. Each small molecule was then quickly added to the plate prior to continuously monitoring the reactions using FP. FP measurements were obtained at 1 min intervals with 2 s of shaking performed prior to each reading at room temperature.

References

1) Anfinsen, C. B. *Science* **1973**, *181*, 223.

2) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333.

3) Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. *Nature* **2011**, *475*, 324.

4) Kim, Y. E.; Hipp, M. S.; Bracher, A.; Hayer-Hartl, M.; Hartl, F. U. *Annu. Rev. Biochem.* **2013**, *82*, 323.

5) Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. *Nature reviews. Molecular cell biology* **2014**, *15*, 384.

6) Royer, C. A. *Chem. Rev.* **2006**, *106*, 1769.

7) Giepmans, B. N.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217.

8) Loving, G. S.; Sainlos, M.; Imperiali, B. *Trends Biotechnol.* **2010**, *28*, 73.

9) Forster, T. *Discuss. Faraday Soc.* **1959**, *No. 27*, 7.

10) Speiser, S. *Chem. Rev.* **1996**, *96*, 1953.

11) Doose, S.; Neuweiler, H.; Sauer, M. *ChemPhysChem* **2009**, *10*, 1389.

12) Rehm, D.; Weller, A. *Isr. J. Chem.* **1970**, *8*, 259.

13) Piston, D. W.; Kremers, G.-J. *Trends Biochem. Sci.* **2007**, *32*, 407.

14) Orevi, T.; Lerner, E.; Rahamim, G.; Amir, D.; Haas, E. In *Fluorescence Spectroscopy and Microscopy*; Engelborghs, Y., Visser, A. J. W. G., Eds.; Humana Press: 2014; Vol. 1076, p 113.

15) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269.

16) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707.

17) Tanaka, T.; Yamamoto, T.; Tsukiji, S.; Nagamune, T. *ChemBioChem* **2008**, *9*, 802.

18) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat Meth* **2005**, *2*, 99.

19) Fernandez-Suarez, M.; Baruah, H.; Martinez-Hernandez, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotech.* **2007**, *25*, 1483.

20) Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786.

21) Seo, M.-H.; Lee, T.-S.; Kim, E.; Cho, Y. L.; Park, H.-S.; Yoon, T.-Y.; Kim, H.-S. *Analytical Chemistry* **2011**, *83*, 8849.

22) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.

23) Wang, L.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 34.

24) Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764.

25) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 6974.

26) Young, D. D.; Young, T. S.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. G. *Biochemistry* **2011**, *50*, 1894.

27) Schultz, K. C.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 13984.

28) Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. *J. Am. Chem. Soc.* **2008**, *130*, 17664.

29) Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36*, 1249.

30) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 4052.

31) Nguyen, D. P.; Elliott, T.; Holt, M.; Muir, T. W.; Chin, J. W. *J. Am. Chem. Soc.* **2011**, *133*, 11418.

32) Wan, W.; Huang, Y.; Wang, Z. Y.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 3211.

33) Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R. *ChemBioChem* **2012**, *13*, 1405.

34) Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272.

35) Wang, K.; Sachdeva, A.; Cox, D. J.; Wilf, N. W.; Lang, K.; Wallace, S.; Mehl, R. A.; Chin, J. W. *Nat. Chem.* **2014**, *6*, 393.

36) Magliery, T. J.; Anderson, J. C.; Schultz, P. G. *J. Mol. Biol.* **2001**, *307*, 755.

37) Wang, K.; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Biotech.* **2007**, *25*, 770.

38) Neumann, H.; Wang, K. H.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441.

39) Wang, K.; Sachdeva, A.; Cox, D. J.; Wilf, N. W.; Lang, K.; Wallace, S.; Mehl, R. A.; Chin, J. W. *Nat Chem* **2014**, *6*, 393.

40) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518.

41) Knall, A.-C.; Slugovc, C. *Chem. Soc. Rev.* **2013**, *42*, 5131.

42) Selvaraj, R.; Fox, J. M. *Curr. Opin. Chem. Biol.* **2013**, *17*, 753.

43) Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struc. Mol. Biol.* **1995**, *2*, 758.

44) Sachdeva, A.; Wang, K.; Elliott, T.; Chin, J. W. *J. Am. Chem. Soc.* **2014**, *136*, 7785.

45) Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W. *Nat Biotechnol* **2014**, *32*, 465.

46) Chen, S.; Fahmi, N. E.; Wang, L.; Bhattacharya, C.; Benkovic, S. J.; Hecht, S. M. *J. Am. Chem. Soc.* **2013**, *135*, 12924.

47) Speight, L. C.; Muthusamy, A. K.; Goldberg, J. M.; Warner, J. B.; Wissner, R. F.; Willi, T. S.; Woodman, B. F.; Mehl, R. A.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 18806.

48) Miyake-Stoner, S. J.; Miller, A. M.; Hammill, J. T.; Peeler, J. C.; Hess, K. R.; Mehl, R. A.; Brewer, S. H. *Biochemistry* **2009**, *48*, 5953.

49) Gelb, D. J.; Oliver, E.; Gilman, S. *Archives of Neurology* **1999**, *56*, 33.

50) Galvin, J. E.; Lee, V. M.; Trojanowski, J. Q. *Archives of Neurology* **2001**, *58*, 186.

51) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M. Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. *Nature* **1997**, *388*, 839.

52) Esposito, G.; Ana Clara, F.; Verstreken, P. *Dev. Neurobiol.* **2012**, *72*, 134.

53) Auluck, P. K.; Caraveo, G.; Lindquist, S. *Annu. Rev. Cell Dev. Biol.* **2010**, *26*, 211.

54) Feany, M. B.; Bender, W. W. *Nature* **2000**, *404*, 394.

55) Fernagut, P.-O.; Chesselet, M.-F. *Neurobiol. Dis.* **2004**, *17*, 123.

56) Luk, K. C.; Kehm, V.; Carroll, J.; Zhang, B.; O'Brien, P.; Trojanowski, J. Q.; Lee, V. M. *Science (New York, N.Y.)* **2012**, *338*, 949.

57) Bartels, T.; Choi, J. G.; Selkoe, D. J. *Nature* **2011**, *477*, 107.

58) Uversky, V. N. *J. Biomol. Struct. Dyn.* **2003**, *21*, 211.

59) Conway, K. A.; Harper, J. D.; Lansbury, P. T. *Nat. Med.* **1998**, *4*, 1318.

60) Uversky, V. N.; Li, J.; Fink, A. L. *J. Biol. Chem.* **2001**, *276*, 10737.

61) Uversky, V. N.; Eliezer, D. *Curr. Protein Pept. Sci.* **2009**, *10*, 483.

62) Alderson, T. R.; Markley, J. L. *Intrinsically Disord. Proteins* **2013**, *1*, e26255.

63) Wang, W.; Perovic, I.; Chittuluru, J.; Kaganovich, A.; Nguyen, L. T.; Liao, J.; Auclair, J. R.; Johnson, D.; Landeru, A.; Simorellis, A. K.; Ju, S.; Cookson, M. R.; Asturias, F. J.; Agar, J. N.; Webb, B. N.; Kang, C.; Ringe, D.; Petsko, G. A.; Pochapsky, T. C.; Hoang, Q. Q. *Proc. Natl. Acad. Sci.* **2011**, *108*, 17797.

64) Binolfi, A.; Theillet, F. X.; Selenko, P. *Biochem. Soc. T.* **2012**, *40*, 950.

65) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 476.

66) Bertoncini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. *Proc. Natl. Acad. Sci.* **2005**, *102*, 1430.

67) Murray, I. V. J.; Giasson, B. I.; Quinn, S. M.; Koppaka, V.; Axelsen, P. H.; Ischiropoulos, H.; Trojanowski, J. Q.; Lee, V. M. Y. *Biochemistry* **2003**, *42*, 8530.

68) Lee, J. C.; Langen, R.; Hummel, P. A.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci.* **2004**, *101*, 16466.

69) Trexler, A. J.; Rhoades, E. *Mol. Neurobiol.* **2013**, *47*, 622.

70) Nath, A.; Sammalkorpi, M.; DeWitt, D. C.; Trexler, A. J.; Elbaum-Garfinkle, S.; O'Hern, C. S.; Rhoades, E. *Biophysical Journal* **2012**, *103*, 1940.

71) Trexler, A. J.; Rhoades, E. *Biophysical Journal* **2010**, *99*, 3048.

72) Chandra, S.; Chen, X. C.; Rizo, J.; Jahn, R.; Sudhof, T. C. *J. Biol. Chem.* **2003**, *278*, 15313.

73) Trexler, A. J.; Rhoades, E. *Biochemistry* **2009**, *48*, 2304.

74) Baskakov, I.; Bolen, D. W. *J. Biol. Chem.* **1998**, *273*, 4831.

75) Baskakov, I. V.; Kumar, R.; Srinivasan, G.; Ji, Y. S.; Wayne Bolen, D.; Thompson, E. B. *J. Biol. Chem.* **1999**, *274*, 10693.

76) Graziano, G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17689.

77) Bandyopadhyay, A.; Saxena, K.; Kasturia, N.; Dalal, V.; Bhatt, N.; Rajkumar, A.; Maity, S.; Sengupta, S.; Chakraborty, K. *Nat. Chem. Biol.* **2012**, *8*, 238.

78) Zou, Q.; Bennion, B. J.; Daggett, V.; Murphy, K. P. *J. Am. Chem. Soc.* **2002**, *124*, 1192.

79) Ma, J.; Pazos, I. M.; Gai, F. *Proc. Natl. Acad. Sci.* **2014**, *111*, 8476.

80) Uversky, V. N.; Li, J.; Fink, A. L. *Febs. Lett.* **2001**, *509*, 31.

81) Ferreon, A. C.; Moosa, M. M.; Gambin, Y.; Deniz, A. A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 17826.

82) Moosa, M. M.; Ferreon, A. C.; Deniz, A. A. *ChemPhysChem* **2015**, *16*, 90.

83) Giasson, B. I.; Uryu, K.; Trojanowski, J. Q.; Lee, V. M. *J. Biol. Chem.* **1999**, *274*, 7619.

84) Serpell, L. C.; Berriman, J.; Jakes, R.; Goedert, M.; Crowther, R. A. *Proc. Natl. Acad. Sci.* **2000**, *97*, 4897.

85) Vilar, M.; Chou, H. T.; Luhrs, T.; Maji, S. K.; Riek-Loher, D.; Verel, R.; Manning, G.; Stahlberg, H.; Riek, R. *Proc. Natl. Acad. Sci.* **2008**, *105*, 8637.

86) Der-Sarkissian, A.; Jao, C. C.; Chen, J.; Langen, R. *J. Biol. Chem.* **2003**, *278*, 37530.

87) Miake, H.; Mizusawa, H.; Iwatsubo, T.; Hasegawa, M. *J. Biol. Chem.* **2002**, *277*, 19213.

88) Chen, M.; Margittai, M.; Chen, J.; Langen, R. *J. Biol. Chem.* **2007**, *282*, 24970.

89) Heise, H.; Hoyer, W.; Becker, S.; Andronesi, O. C.; Riedel, D.; Baldus, M. *Proc. Natl. Acad. Sci.* **2005**, *102*, 15871.

90) Comellas, G.; Lemkau, L. R.; Nieuwkoop, A. J.; Kloepper, K. D.; Ladror, D. T.; Ebisu, R.; Woods, W. S.; Lipton, A. S.; George, J. M.; Rienstra, C. M. *J. Mol. Biol.* **2011**, *411*, 881.

91) Gath, J.; Habenstein, B.; Bousset, L.; Melki, R.; Meier, B.; Böckmann, A. *Biomol. NMR Assign.* **2012**, *6*, 51.

92) Pornsuwan, S.; Giller, K.; Riedel, D.; Becker, S.; Griesinger, C.; Bennati, M. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 10290.

93) Apetri, M. M.; Maiti, N. C.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. *J. Mol. Biol.* **2006**, *355*, 63.

94) van Rooijen, B. D.; van Leijenhorst-Groener, K. A.; Claessens, M. M.; Subramaniam, V. *J. Mol. Biol.* **2009**, *394*, 826.

95) Mysling, S.; Betzer, C.; Jensen, P. H.; Jorgensen, T. J. *Biochemistry* **2013**, *52*, 9097.

96) Drescher, M.; Huber, M.; Subramaniam, V. *ChemBioChem* **2012**, *13*, 761.

97) Wolfe, J. P.; Wagaw, S.; Marcoux, J. F.; Buchwald, S. L. *Acc. Chem. Res.* **1998**, *31*, 805.

98) Dusa, A.; Kaylor, J.; Edridge, S.; Bodner, N.; Hong, D. P.; Fink, A. L. *Biochemistry* **2006**, *45*, 2752.

99) Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M. *J. Mol. Biol.* **2008**, *378*, 1064.

100) Yushchenko, D. A.; Fauerbach, J. A.; Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M. *J. Am. Chem. Soc.* **2010**, *132*, 7860.

101) Yap, T. L.; Pfefferkorn, C. M.; Lee, J. C. *Biochemistry* **2011**, *50*, 1963.

102) Kaylor, J.; Bodner, N.; Edridge, S.; Yamin, G.; Hong, D. P.; Fink, A. L. *J. Mol. Biol.* **2005**, *353*, 357.

103) Nath, S.; Meuvis, J.; Hendrix, J.; Carl, S. A.; Engelborghs, Y. *Biophysical journal* **2010**, *98*, 1302.

104) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

105) Truter, M. R. *J. Am. Chem. Soc.* **1960**, 997.

106) Sifferlen, T.; Rueping, M.; Gademann, K.; Jaun, B.; Seebach, D. *Helv. Chim. Acta.* **1999**, *82*, 2067.

107) Wiberg, K. B.; Rush, D. J. *J. Am. Chem. Soc.* **2001**, *123*, 2038.

108) Dudek, E. P.; Dudek, G. *J. Org. Chem.* **1967**, *32*, 823.

109) Hollosi, M.; Majer, Z.; Zewdu, M.; Ruff, F.; Kajtar, M.; Kover, K. E. *Tetrahedron* **1988**, *44*, 195.

110) Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T. *Journal of the American Chemical Society* **2013**, *135*, 7843.

111) Basch, H.; Robin, M. B.; Kuebler, N. A. *J. Chem. Phys.* **1967**, *47*, 1201.

112) Hosoya, H.; Tanaka, J.; Nagakura, S. *Bulletin of the Chemical Society of Japan* **1960**, *33*, 850.

113) Goldberg, J. M.; Wissner, R. F.; Klein, A. M.; Petersson, E. J. *Chemical Communications* **2012**, *48*, 1550.

114) Shalaby, M. A.; Grote, C. W.; Rapoport, H. *J. Org. Chem.* **1996**, *61*, 9045.

115) Goldberg, J. M.; Batjargal, S.; Petersson, E. J. *J. Am. Chem. Soc.* **2010**, *132*, 14718.

116) Batjargal, S.; Wang, Y. J.; Goldberg, J. M.; Wissner, R. F.; Petersson, E. J. *J. Am. Chem. Soc.* **2012**, *134*, 9172.

117) Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 6529.

118) Batjargal, S.; Huang, Y.; Wang, Y.; Petersson, E. J. *J. Pept. Sci.* **In Press**.

119) Lacour, T. F. M. *Int. J. Pept. Protein Res.* **1987**, *30*, 564.

120) Miwa, J. H.; Pallivathucal, L.; Gowda, S.; Lee, K. E. *Org. Lett.* **2002**, *4*, 4655.

121) Reiner, A.; Wildemann, D.; Fischer, G.; Kiefhaber, T. *J. Am. Chem. Soc.* **2008**, *130*, 8079.

122) Miwa, J. H.; Patel, A. K.; Vivatrat, N.; Popek, S. M.; Meyer, A. M. *Org. Lett.* **2001**, *3*, 3373.

123) Culik, R. M.; Jo, H.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2012**, *134*, 8026.

124) Kahnt, J.; Buchenau, B.; Mahlert, F.; Kruger, M.; Shima, S.; Thauer, R. K. *Febs J.* **2007**, *274*, 4913.

125) Choudhary, A.; Raines, R. T. *ChemBioChem* **2012**, *12*, 1801.

126) Bachmann, A.; Wildemann, D.; Praetorius, F.; Fischer, G.; Kiefhaber, T. *Proc. Natl. Acad. Sci.* **2011**, *108*, 3952.

127) Wildemann, D.; Schiene-Fischer, C.; Aumueller, T.; Bachmann, A.; Kiefhaber, T.; Luecke, C.; Fischer, G. *J. Am. Chem. Soc.* **2007**, *129*, 4910.

128) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045.

129) McKnight, C. J.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. *J. Mol. Biol.* **1996**, *260*, 126.

130) Wiczk, W. M.; Gryczynski, I.; Szmacinski, H.; Johnson, M. L.; Kruszynski, M.; Zboinska, J. *Biophys. Chem.* **1988**, *32*, 43.

131) Rownicka-Zubik, J.; Sulkowska, A.; Bojko, B.; Maciazek-Jurczyk, M.; Pozycka, J.; Pentak, D.; Sulkowski, W. W. *J. Photochem. Photobiol., B* **2009**, *97*, 54.

132) Urbauer, J. L.; Short, J. H.; Dow, L. K.; Wand, A. J. *Biochemistry* **1995**, *34*, 8099.

133) Contessa, G. M.; Orsale, M.; Melino, S.; Torre, V.; Paci, M.; Desideri, A.; Cicero, D. O. *J. Biomol. NMR* **2005**, *31*, 185.

134) Chen, Y.; Barkley, M. D. *Biochemistry* **1998**, *37*, 9976.

135) Chen, Y.; Liu, B.; Yu, H. T.; Barkley, M. D. *J. Am. Chem. Soc.* **1996**, *118*, 9271.

136) Qiu, W. H.; Li, T. P.; Zhang, L. Y.; Yang, Y.; Kao, Y. T.; Wang, L. J.; Zhong, D. P. *Chem. Phys. Lett.* **2008**, *350*, 154.

137) Kilhoffer, M. C.; Demaille, J. G.; Gerard, D. *Biochemistry* **1981**, *20*, 4407.

138) Pundak, S.; Roche, R. S. *Biochemistry* **1984**, *23*, 1549.

139) Haiech, J.; Kilhoffer, M.-C. *Top. Fluoresc. Spectrosc.* **2000**, *6*, 175.

140) Kilhoffer, M. C.; Roberts, D. M.; Adibi, A.; Watterson, D. M.; Haiech, J. *Biochemistry* **1989**, *28*, 6086.

141) Klee, C. B. *Biochemistry* **1977**, *16*, 1017.

142) Bordwell, F. G.; Algrim, D. J.; Harrelson, J. A. *J. Am. Chem. Soc.* **1988**, *110*, 5903.

143) Dawson, P. E.; Muir, T. W.; Clarklewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776.

144) van Ham, T. J.; Esposito, A.; Kumita, J. R.; Hsu, S. T. D.; Schierle, G. S. K.; Kaminsk, C. F.; Dobson, C. M.; Nollen, E. A. A.; Bertoncini, C. W. *J. Mol. Biol.* **2010**, *395*, 627.

145) Jenny, R. J.; Mann, K. G.; Lundblad, R. L. *Protein Express. Purif.* **2003**, *31*, 1.

146) Goldberg, J. M.; Wissner, R. F.; Klein, A. M.; Petersson, E. J. *Chem. Comm.* **2012**, *48*, 1550.

147) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283.

148) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

149) Brzeska, H.; Venyaminov, S. V.; Grabarek, Z.; Drabikowski, W. *Febs. Lett.* **1983**, *153*, 169.

150) Goldberg, J. M.; Wissner, R. F.; Klein, A. M.; Petersson, E. J. *Chemical Communications* **2012**, *48*, 1550.

151) Goldberg, J. M.; Batjargal, S.; Chen, B. S.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 18651.

152) Alewood, P.; Engelhard, M.; Kent, S. B. H. *J. Pept. Sci.* **2010**, *16*, 513.

153) Muralidharan, V.; Muir, T. W. *Nat. Methods* **2006**, *3*, 429.

154) Wood, S. J.; Wypych, J.; Steavenson, S.; Louis, J. C.; Citron, M.; Biere, A. L. *J. Biol. Chem.* **1999**, *274*, 19509.

155) Li, X.; Fekner, T.; Ottesen, J. J.; Chan, M. K. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 9184.

156) Valiyaveetil, F. I.; Sekedat, M.; MacKinnon, R.; Muir, T. W. *Proc. Natl. Acad. Sci.* **2004**, *101*, 17045.

157) Fu, Y. W.; Bieschke, J.; Kelly, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 15366.

158) Fu, Y. W.; Gao, J. M.; Bieschke, J.; Dendle, M. A.; Kelly, J. W. *J. Am. Chem. Soc.* **2006**, *128*, 15948.

159) Peeler, J.; Mehl, R. In *Unnatural Amino Acids*; Pollegioni, L., Servi, S., Eds.; Humana Press: 2012; Vol. 794, p 125.

160) Lakowicz, J. R. *Principles of fluorescence spectroscopy*; Third ed.; Springer: New York, NY, 2006.

161) Chen, H.; Ahsan, S. S.; Santiago-Berrios, M. E. B.; AbrunÌƒa, H. D.; Webb, W. W. *J. Am. Chem. Soc.* **2010**, *132*, 7244.

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

163) Hejjaoui, M.; Haj-Yahya, M.; Kumar, K. S. A.; Brik, A.; Lashuel, H. A. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 405.

164) Gentle, I. E.; De Souza, D. P.; Baca, M. *Bioconjugate Chem.* **2004**, *15*, 658.

165) Flemer, S. *J. Pept. Sci.* **2009**, *15*, 693.

166) Camarero, J. A.; Mitchell, A. R. *Protein Pept. Lett.* **2005**, *12*, 723.

167) Nagalingam, A. C.; Radford, S. E.; Warriner, S. L. *Synlett* **2007**, 2517.

168) Kang, J.; Macmillan, D. *Org. Biol. Chem.* **2010**, *8*, 1993.

169) Kawakami, T.; Aimoto, S. *Tetrahedron* **2009**, *65*, 3871.

170) Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J. Mol. Biol.* **2010**, *395*, 361.

171) Desplats, P.; Lee, H. J.; Bae, E. J.; Patrick, C.; Rockenstein, E.; Crews, L.; Spencer, B.; Masliah, E.; Lee, S. J. *Proc. Natl. Acad. Sci.* **2009**, *106*, 13010.

172) Volpicelli-Daley, L. A.; Luk, K. C.; Patel, T. P.; Tanik, S. A.; Riddle, D. M.; Stieber, A.; Meaney, D. F.; Trojanowski, J. Q.; Lee, V. M. Y. *Neuron* **2011**, *72*, 57.

173) Winner, B.; Jappelli, R.; Maji, S. K.; Desplats, P. A.; Boyer, L.; Aigner, S.; Hetzer, C.; Loher, T.; Vilar, M.; Campionic, S.; Tzitzilonis, C.; Soragni, A.; Jessberger, S.; Mira, H.; Consiglio, A.; Pham, E.; Masliah, E.; Gage, F. H.; Riek, R. *Proc. Natl. Acad. Sci.* **2011**, *108*, 4194.

174) Conway, K. A.; Harper, J. D.; Lansbury, P. T. *Biochemistry* **2000**, *39*, 2552.

175) Hejjaoui, M.; Butterfield, S.; Fauvet, B.; Vercruysse, F.; Cui, J.; Dikiy, I.; Prudent, M.; Olschewski, D.; Zhang, Y.; Eliezer, D.; Lashuel, H. A. *J. Am. Chem. Soc.* **2012**, *134*, 5196.

176) Fauvet, B.; Fares, M. B.; Samuel, F.; Dikiy, I.; Tandon, A.; Eliezer, D.; Lashuel, H. A. *J. Biol. Chem.* **2012**.

177) Allison, J. R.; Varnai, P.; Dobson, C. M.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 18314.

178) Celej, M. S.; Sarroukh, R.; Goormaghtigh, E.; Fidelio, G. D.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2012**, *443*, 719.

179) Karyagina, I.; Becker, S.; Giller, K.; Riedel, D.; Jovin, T. M.; Griesinger, C.; Bennati, M. *Biophysical Journal* **2011**, *101*, L1.

180) Ramakrishnan, M.; Jensen, P. H.; Marsh, D. *Biochemistry* **2006**, *45*, 3386.

181) Wu, K. P.; Baum, J. *J. Am. Chem. Soc.* **2010**, *132*, 5546.

182) Taskent-Sezgin, H.; Chung, J.; Patsalo, V.; Miyake-Stoner, S. J.; Miller, A. M.; Brewer, S. H.; Mehl, R. A.; Green, D. F.; Raleigh, D. P.; Carrico, I. *Biochemistry* **2009**, *48*, 9040.

183) Taskent-Sezgin, H.; Marek, P.; Thomas, R.; Goldberg, D.; Chung, J.; Carrico, I.; Raleigh, D. P. *Biochemistry* **2010**, *49*, 6290.

184) Ferreon, A. C. M.; Gambin, Y.; Lemke, E. A.; Deniz, A. A. *Proc. Natl. Acad. Sci.* **2009**, *106*, 5645.

185) Ferreon, A. C. M.; Moosa, M. M.; Gambin, Y.; Deniz, A. A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 17826.

186) Lee, J. C.; Gray, H. B.; Winkler, J. R. *J. Am. Chem. Soc.* **2005**, *127*, 16388.

187) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T. J.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proc. Natl. Acad. Sci.* **1999**, *96*, 3670.

188) Tucker, M. J.; Oyola, R.; Gai, F. *J. Phys. Chem. B* **2005**, *109*, 4788.

189) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. *J. Mol. Biol.* **2005**, *352*, 672.

190) Wang, J. Y.; Xie, J. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 8738.

191) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026.

192) Speight, L. C.; Goldberg, J. M.; Warner, J. B.; Mehl, R. A.; Petersson, E. J. **Unpublished results**.

193) Huang, C.; Yan, S. J.; Li, Y. M.; Huang, R.; Lin, J. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4665.

194) Pivato, M.; De Franceschi, G.; Tosatto, L.; Frare, E.; Kumar, D.; Aioanei, D.; Brucale, M.; Tessari, I.; Bisaglia, M.; Samori, B.; de Laureto, P. P.; Bubacco, L. *PLoS One* **2012**, *7*, e50027.

195) Zhou, W.; Freed, C. R. *J. Biol. Chem.* **2004**, *279*, 10128.

196) Hendrickson, T. L.; Imperiali, B. *Biochemistry* **1995**, *34*, 9444.

197) Tam, J. P.; Yu, Q. T. *Biopolymers* **1998**, *46*, 319.

198) Saporito, A.; Marasco, D.; Chambery, A.; Botti, P.; Monti, S. M.; Pedone, C.; Ruvo, M. *Biopolymers* **2006**, *83*, 508.

199) Aussedat, B.; Fasching, B.; Johnston, E.; Sane, N.; Nagorny, P.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 3532.

200) Pachamuthu, K.; Schmidt, R. R. *Synlett* **2003**, 659.

201) Tanaka, T.; Wagner, A. M.; Warner, J. B.; Wang, Y. J.; Petersson, E. J. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 6210.

202) Suto, K.; Shimizu, Y.; Watanabe, K.; Ueda, T.; Fukai, S.; Nureki, O.; Tomita, K. *Embo J.* **2006**, *25*, 5942.

203) Watanabe, K.; Toh, Y.; Suto, K.; Shimizu, Y.; Oka, N.; Wada, T.; Tomita, K. *Nature* **2007**, *449*, 867.

204) Link, A. J.; Vink, M. K. S.; Agard, N. J.; Prescher, J. A.; Bertozzi, C. R.; Tirrell, D. A. *Proc. Natl. Acad. Sci.* **2006**, *103*, 10180.

205) Batjargal, S.; Wang, Y. J.; Goldberg, J. M.; Wissner, R. F.; Petersson, E. J. *J. Am. Chem. Soc.* **2012**, *134*, 9172.

206) Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 6529.

207) Hamada, H.; Kameshima, N.; Szymanska, A.; Wegner, K.; Lankiewicz, L.; Shinohara, H.; Taki, M.; Sisido, M. *Bioorg. Med. Chem. Lett.* **2005**, *13*, 3379.

208) Szymanska, A.; Wegner, K.; Lankiewicz, L. *Helv. Chim. Acta.* **2003**, *86*, 3326.

209) Peeler, J. C.; Woodman, B. F.; Averick, S.; Miyake-Stoner, S. J.; Stokes, A. L.; Hess, K. R.; Matyjaszewski, K.; Mehl, R. A. *J. Am. Chem. Soc.* **2010**, *132*, 13575.

210) Tanaka, T.; Wagner, A. M.; Warner, J. B.; Wang, Y. X. J.; Petersson, E. J. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 6210.

211) Hanahan, D.; Jessee, J.; Bloom, F. R. *Methods Enzymol.* **1991**, *204*, 63.

212) Stevens-Truss, R.; Marletta, M. A. *Biochemistry* **1995**, *34*, 15638.

213) Tucker, M. J.; Oyola, R.; Gai, F. *Biopolymers* **2006**, *83*, 571.

214) Ferreon, A. C. M.; Moran, C. R.; Gambin, Y.; Deniz, A. A. In *Methods Enzymol.*; Nils, G. W., Ed.; Academic Press: 2010; Vol. Volume 472, p 179.

215) Haas, E. In *Intrinsically Disordered Protein Analysis*; Uversky, V. N., Dunker, A. K., Eds.; Humana Press: 2012; Vol. 895, p 467.

216) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angewandte Chemie* **2002**, *114*, 2708.

217) Chalker, J. M.; Bernardes, G. J.; Lin, Y. A.; Davis, B. G. *Chem. Asian. J.* **2009**, *4*, 630.

218) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441.

219) Huang, Y.; Russell, W. K.; Wan, W.; Pai, P. J.; Russell, D. H.; Liu, W. S. *Molecular Biosystems* **2010**, *6*, 683.

220) Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 3211.

221) Johnson, D. B. F.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L. *Nat. Chem. Biol.* **2011**, *7*, 779.

222) Xiao, H.; Chatterjee, A.; Choi, S.-h.; Bajjuri, K. M.; Sinha, S. C.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 14080.

223) Dougherty, M. J.; Kothakota, S.; Mason, T. L.; Tirrell, D. A.; Fournier, M. J. *Macromolecules* **1993**, *26*, 1779.

224) Yoshikawa, E.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Macromolecules* **1994**, *27*, 5471.

225) Chalker, J. M.; Bernardes, G. J.; Davis, B. G. *Acc Chem Res* **2011**, *44*, 730.

226) Varshavsky, A. *Genes to Cells* **1997**, *2*, 13.

227) Wagner, A. M.; Fegley, M. W.; Warner, J. B.; Grindley, C. L. J.; Marotta, N. P.; Petersson, E. J. *J. Am. Chem. Soc.* **2011**, *133*, 15139.

228) Roberti, M. J.; Jovin, T. M.; Jares-Erijman, E. *PLoS One* **2011**, *6*, e23338.

229) Grupi, A.; Haas, E. *J. Mol. Biol.* **2011**, *411*, 234.

230) Schmid, A. W.; Fauvet, B.; Moniatte, M.; Lashuel, H. A. *Molecular & Cellular Proteomics : MCP* **2013**, *12*, 3543.

231) Li, C.; Lutz, E. A.; Slade, K. M.; Ruf, R. A. S.; Wang, G.-F.; Pielak, G. J. *Biochemistry* **2009**, *48*, 8578.

232) Warner, J. B., University of Pennsylvania, 2014.

233) Nehring, S.; Budisa, N.; Wiltschi, B. *PLoS One* **2012**, *7*, e31992.

234) Auluck, P. K.; Caraveo, G.; Lindquist, S. *Annual Review of Cell and Developmental Biology, Vol 26* **2010**, *26*, 211.

235) LI, J.; ZHU, M.; MANNING-BOG, A. B.; DI MONTE, D. A.; FINK, A. L. *FASEB J.* **2004**, *18*, 962.

236) Singh, P. K.; Kotia, V.; Ghosh, D.; Mohite, G. M.; Kumar, A.; Maji, S. K. *ACS. Chem. Neurosci.* **2013**, *4*, 393.

237) Bieschke, J.; Russ, J.; Friedrich, R. P.; Ehrnhoefer, D. E.; Wobst, H.; Neugebauer, K.; Wanker, E. E. *Proc. Natl. Acad. Sci.* **2010**, *107*, 7710.

238) Prabhudesai, S.; Sinha, S.; Attar, A.; Kotagiri, A.; Fitzmaurice, A.; Lakshmanan, R.; Ivanova, M.; Loo, J.; Klärner, F.-G.; Schrader, T.; Stahl, M.; Bitan, G.; Bronstein, J. *Neurotherapeutics* **2012**, *9*, 464.

239) Li, J.; Zhu, M.; Rajamani, S.; Uversky, V. N.; Fink, A. L. *Chem. Biol.* **2004**, *11*, 1513.

240) Gautam, S.; Karmakar, S.; Bose, A.; Chowdhury, P. K. *Biochemistry* **2014**, *53*, 4081.

241) Caruana, M.; Högen, T.; Levin, J.; Hillmer, A.; Giese, A.; Vassallo, N. *Febs. Lett.* **2011**, *585*, 1113.

242) Lea, W. A.; Simeonov, A. *Expert Opin Drug Discov.* **2011**, *6*, 17.

243) Luk, K. C.; Hyde, E. G.; Trojanowski, J. Q.; Lee, V. M. Y. *Biochemistry* **2007**, *46*, 12522.

244) Brinkley, M. *Bioconjugate Chem.* **1992**, *3*, 2.

245) Tramier, M.; Coppey-Moisan, M. In *Methods in Cell Biology*; Kevin, F. S., Ed.; Academic Press: 2008; Vol. Volume 85, p 395.

246) Eliezer, D.; Kutluay, E.; Bussell Jr, R.; Browne, G. *J. Mol. Biol.* **2001**, *307*, 1061.

247) Ulmer, T. S.; Bax, A.; Cole, N. B.; Nussbaum, R. L. *J. Biol. Chem.* **2005**, *280*, 9595.

248) Conway, K. A.; Rochet, J.-C.; Bieganski, R. M.; Lansbury, P. T. *Science* **2001**, *294*, 1346.

249) Norris, E. H.; Giasson, B. I.; Hodara, R.; Xu, S.; Trojanowski, J. Q.; Ischiropoulos, H.; Lee, V. M.-Y. *J. Biol. Chem.* **2005**, *280*, 21212.

250) Rekas, A.; Knott, R.; Sokolova, A.; Barnham, K.; Perez, K.; Masters, C.; Drew, S.; Cappai, R.; Curtain, C.; Pham, C. L. *Eur Biophys J* **2010**, *39*, 1407.

251) Planchard, M. S.; Exley, S. E.; Morgan, S. E.; Rangachari, V. *Protein Science* **2014**, *23*, 1369.

252) Rochet, J.-C.; Fleming Outeiro, T.; Conway, K.; Ding, T.; Volles, M.; Lashuel, H.; Bieganski, R.; Lindquist, S.; Lansbury, P. *J Mol Neurosci* **2004**, *23*, 23.

253) Lorenzen, N.; Nielsen, S. B.; Yoshimura, Y.; Vad, B. S.; Andersen, C. B.; Betzer, C.; Kaspersen, J. D.; Christiansen, G.; Pedersen, J. S.; Jensen, P. H.; Mulder, F. A. A.; Otzen, D. E. *J. Biol. Chem.* **2014**, *289*, 21299.

254) Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E. *Nat. Struc. Mol. Biol.* **2008**, *15*, 558.