1-1-2015

Chemical Modification Methods for Protein Misfolding Studies

Yanxin Wang
University of Pennsylvania, wangyanxin9003@gmail.com

Chemical Modification Methods for Protein Misfolding Studies

Abstract
Protein misfolding is the basis of various human diseases, including Parkinson's disease, Alzheimer's disease and Type 2 diabetes. When a protein misfolds, it adopts the wrong three dimensional structures that are dysfunctional and sometime pathological. Little structural details are known about this misfolding phenomenon due to the lack of characterization tools. Our group previously demonstrated that a thioamide, a single atom substitution of the peptide bond, could serve as a minimalist fluorescence quencher. In the current study, we showed the development of protein semi-synthesis strategies for the incorporation of thioamides into full-length proteins for misfolding studies.

We adopted the native chemical ligation (NCL) method between a C-terminal thioester fragment and an N-terminal Cys fragment. We first devised strategies for the synthesis of thioamide-containing peptide thioesters as NCL substrates, and demonstrated their applications in generating a thioamide/Trp-dually labeled α-synuclein (αS), which was subsequently used in a proof-of-concept misfolding study. To remove the constraint of a Cys at the ligation site, we explored traceless ligation methods that desulfurized Cys into Ala, or β- and γ- thiol analogs into native amino acids after ligation in the presence of thioamides. We further demonstrated that selective deselenization could be achieved in the presence of both Cys residues and thioamides, expanding the scope of thioamide incorporation through traceless ligation to proteins with native Cys. Finally, we showed that hemiselenide protected selenocysteines (Sec) can be incorporated onto the protein N-terminus through chemoenzymatic modification by aminoacyl transferase (AaT) as ligation handles. Further developments are underway in our laboratory to expand the AaT substrate scope for β- and γ- thiol amino acid analogs. In summary, we developed a set of methods that allowed the incorporation of thioamide probes into full-length protein, which enabled the application of this minimalist probe in protein misfolding studies.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemistry

First Advisor
E. James Petersson

Keywords
Chemical biology, Parkinson's disease, Protein engineering, Protein semi-synthesis, Synuclein, Thioamide

Subject Categories
Biochemistry | Chemistry

CHEMICAL MODIFICATION METHODS FOR PROTEIN MISFOLDING STUDIES

Yanxin Wang

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

Associate Professor of Chemistry

Graduate Group Chairperson

_____

Gary A. Molander, Hirschmann-Makineni Professor of Chemistry

Dissertation Committee

Tobias Baumgart, Associate Professor of Chemistry

Donna Huryn, Adjunct Professor of Chemistry

Jeffery G. Saven, Professor of Chemistry

CHEMICAL MODIFICATION METHODS FOR PROTEIN MISFOLDING STUDIES

For my parents and mentors

ACKNOWLEDGMENT

First and foremost, I would like to thank Professor E. James Petersson for allowing me to join his laboratory and for his generous guidance throughout my graduate career. He has been a very knowledgeable and supportive mentor, who introduced me to the field of chemical biology and helped me develop my professional skills as a scientist. He gave me much freedom in intellectual exploration, while always making himself available for feedback and discussions. I would not have been where I am without his support. I would also like to thank my dissertation committee members − Professors Tobias Baumgart, Donna Huryn and Jeffery G. Saven – for consistently providing useful suggestions and fresh perspectives that helped me improve my research.

I am fortunate to have worked with the many members and alumni of the Petersson group, who have been great colleagues, collaborators, and friends: Dr. Tomohiro Tanaka, Dr. Yun Huang, Dr. Conor Haney, Dr. Jacob Goldberg, Dr. Anne Wagner, Dr. John Warner, Dr. Lee Speight, Dr. Solongo Batjargal, Dr. Rebecca Wissner, Mark Fegley, Christopher Walters, X. Stella Chen, John J. Ferrie, D. Miklos Szantai-Kis, Itthipol Sung-weinwong, Christina Cleveland, J. M. Vicky Jun, Eileen Moison, Colin Fadzen, Anand Muthusamy, E. Keith Keenan, Eileen Hoang, and Jimin Yoon. We worked together to keep the instruments functional, the supplies well-stocked, and the everyday life exciting. I am particularly appreciative of Dr. Jacob Goldberg, Dr. Anne Wagner and Dr. Solongo Batjargal for their invaluable inputs on my research, and for getting me started in the lab.

I owe my gratitude for the many people in and outside the department for helping me obtain good quality data and for maintaining a care-free research environment: Dr. Rakesh Kohli for mass spectrometry, Dr. George Furst and Dr. Jun Gu for NMR spectroscopy, Dr. Patrick Carroll for X-ray crystallography, Dr. Christopher Lanci for biological chemistry instrumentation, Dr. Dewight Williams for electron microscopy, Dr. Dustin Covell for cell culture, Judith N. Currano for the numerous times that she helped me find references and information, Mandy Swope and Kristen Muscat for watching my back on administrative business, and all other staff of the chemistry department.

Finally, I would like to express my sincere appreciation for my family and friends. With me being the only child, my parents made tremendous sacrifices by allowing me to study here half way across the globe for five years. I owe my achievements and growth to their continuous inspiration, cultivation and unconditional support. I am also thankful for having a group of caring and supportive friends; we had much fun together and helped each other maintain our sanity in graduate school.

ABSTRACT

CHEMICAL MODIFICATION METHODS FOR PROTEIN MISFOLDING STUDIES

Yanxin Wang

E. James Petersson

Protein misfolding is the basis of various human diseases, including Parkinson's disease, Alzheimer's disease and Type 2 diabetes. When a protein misfolds, it adopts the wrong three dimensional structures that are dysfunctional and sometime pathological. Little structural details are known about this misfolding phenomenon due to the lack of characterization tools. Our group previously demonstrated that a thioamide, a single atom substitution of the peptide bond, could serve as a minimalist fluorescence quencher. In the current study, we showed the development of protein semi-synthesis strategies for the incorporation of thioamides into full-length proteins for misfolding studies.

We adopted the native chemical ligation (NCL) method between a C-terminal thioester fragment and an N-terminal Cys fragment. We first devised strategies for the synthesis of thioamide-containing peptide thioesters as NCL substrates, and demonstrated their applications in generating a thioamide/Trp-dually labeled α-synuclein (αS), which was subsequently used in a proof-of-concept misfolding study. To remove the constraint of a Cys at the ligation site, we explored traceless ligation methods that desulfurized Cys into Ala, or β- and γ- thiol analogs into native amino acids after ligation in the presence of thioamides. We further demonstrated that selective deselenization could be achieved in

the presence of both Cys residues and thioamides, expanding the scope of thioamide incorporation through traceless ligation to proteins with native Cys. Finally, we showed that hemiselenide protected selenocysteines (Sec) can be incorporated onto the protein N-terminus through chemoenzymatic modification by aminoacyl transferase (AaT) as ligation handles. Further developments are underway in our laboratory to expand the AaT substrate scope for β- and γ- thiol amino acid analogs. In summary, we developed a set of methods that allowed the incorporation of thioamide probes into full-length protein, which enabled the application of this minimalist probe in protein misfolding studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter 1 . Introduction

Part of the content in this chapter was originally published in *Organic & Biomolecular Chemistry*. It is adapted here with permission from the publisher:

1.1 Protein Misfolding and Human Diseases

Protein Misfolding Proteins are an important class of biological macromolecules, and key building blocks of life[1]. They are responsible for the majority of cellular activities in living organisms, from maintaining cell morphology and catalyzing metabolic reactions, to signaling and homeostasis[2]. All proteins are constructed from a simple scaffold of linear chains of amino acids that are joined together through peptide bonds[3]. The classic view − as represented by the "central dogma"[4] and the Anfinsen hypothesis[5] − states that the structures and functions of proteins are largely encoded by their amino acid sequences; a defined sequence is transcribed from genetic information stored in DNA into RNA, which is in turn translated into a protein that folds into a defined three-dimensional structure to serve a set of defined biological functions.

With the emergence of the "energy landscape" theory for protein folding[6] and empirical characterization of protein dynamics[7,8,9], we now understand that proteins may assume a variety of conformations and that the amino acid sequence *per se* is not sufficient to determine its three-dimensional structure. For many proteins, their biologically active conformations are not the "global energy minimum", but rather a "local energy minimum" that is either stabilized by a binding partner or trapped by kinetic barriers to the next available conformation (Figure 1-1)[10]. As proteins sample various energy minima on their folding energy landscapes, they can also "misfold" into wrong conformations that are dysfunctional and sometimes pathological. In our laboratory, we strive to construct chemical tools for the characterization of this misfolding phenomenon.

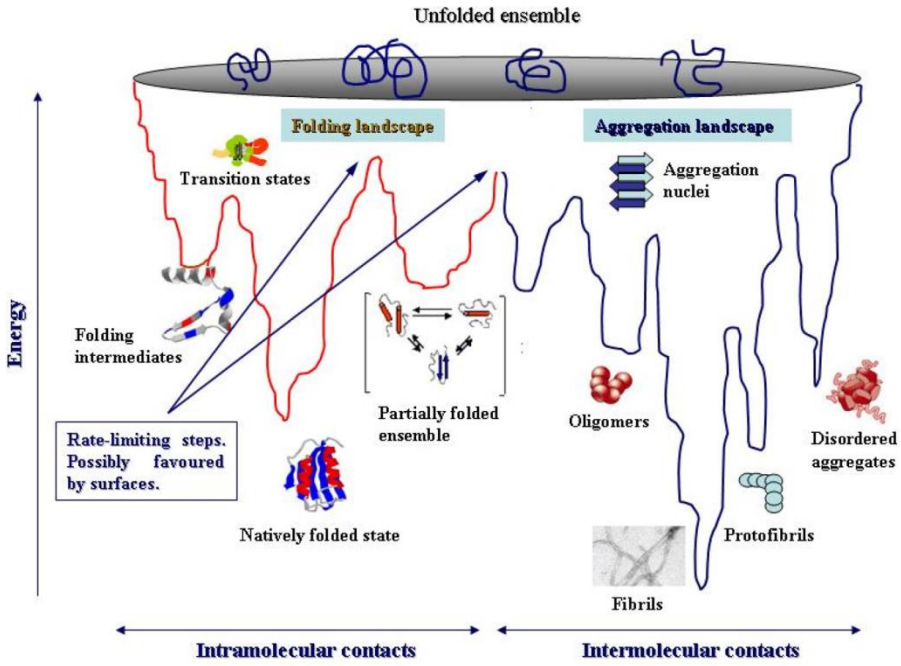


Figure 1-1. A Unified View of Protein Folding and Misfolding.

Various folding states on a protein energy landscape. Graphics adapted from Stefani[10].

Protein Misfolding Diseases Protein misfolding is of great significance to human health, particularly in the context of aging-associated diseases. The majority of "degenerative diseases" (including Alzheimer's disease (AD), Parkinson's disease (PD) and Type 2 diabetes) involve toxic loss-of-function or gain-of-function in particular tissues, which at the molecular level, can be attributed to protein misfolding[11,12,13]. Table 1-1 summarizes some of the most well-studied protein misfolding diseases[14,15]; it is worth-noting that many of these diseases are sporadic rather than hereditary – the misfolded protein originates from a normally folded, functional protein that becomes pathological (or accumulates to a critical amount that overwhelms the cell's ability to clear misfolded proteins) at the onset of the disease[16]. The accumulation of misfolded proteins leads to a

variety of consequences such as oxidative stress and cell deformation, which ultimately results in cell death and thus the "degenerative" symptoms (Figure 1-2)[14].

Table 1-1. Selected Protein Misfolding Diseases in Human.

Human Disease	Associated Protein(s)	Affected Tissue
Neurodegenerative Diseases		
Alzheimer's disease (AD)	Amyloid β, tau	Brain
Parkinson's disease (PD)	α-Synuclein, tau	Brain
Lewy-body dementia (DLB)	α-Synuclein	Brain
Huntington's disease (HD)	Hungtintin	Brain
Frontotemporal dementia (FTD)	Tau	Brain
Amyotrophic lateral sclerosis (ALS)	Superoxide dismutase 1	Brain
Other Aging-Related Diseases		
Primary systemic amyloidosis (PSA)	Ig light chain	Systemic
Senile systemic amyloidosis (SSA)	Transthyretin	Microvasculature
Atherosclerosis	Low-density lipoprotein (LDL)	Arteries
Cataract	Crystallin	Eye
Type 2 diabetes (T2D)	Amylin	Pancreas

* Adapted from Christensen et al.[14] and Gebbink et al.[15]

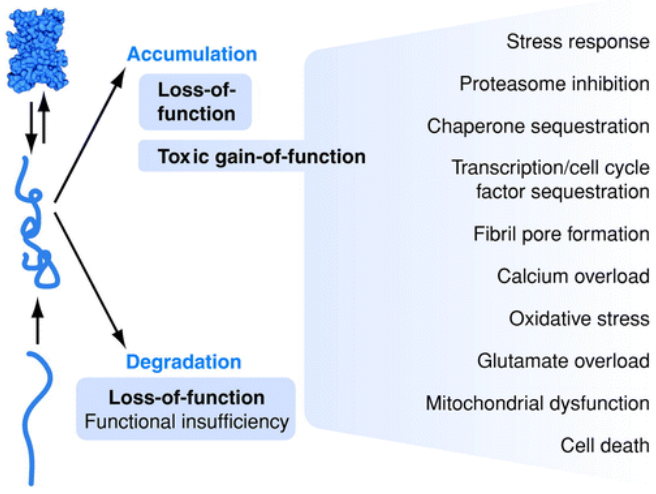


Figure 1-2. Cellular Consequences of Pathological Protein Misfolding. Graphics adapted from Christensen *et al.*[14]

At the molecular level, the most common type of pathological protein misfolding is aggregation – in particular, the transformation of a natively folded or unfolded protein into β-sheet rich aggregates known as "amyloids"[17]. Due to the nature of the β-sheet secondary structure, these aggregates have unsatisfied backbone hydrogen bond donors and acceptors on either ends of the β-sheets, and have a tendency to propagate by recruiting additional proteins and templating their folding into β-sheets (Figure 1-3)[15,18]. Once a small population of misfolded monomers and/or oligomers is formed, a vicious self-propagation cycle is initiated[18]. Additionally, amyloids are more resistant to proteolytic degradation than their natively folded or unfolded counterparts[19], making it difficult to reverse the misfolding. This amyloidosis phenomenon is the basis for many of the most infamous aging-related diseases, including AD and PD[12,18].

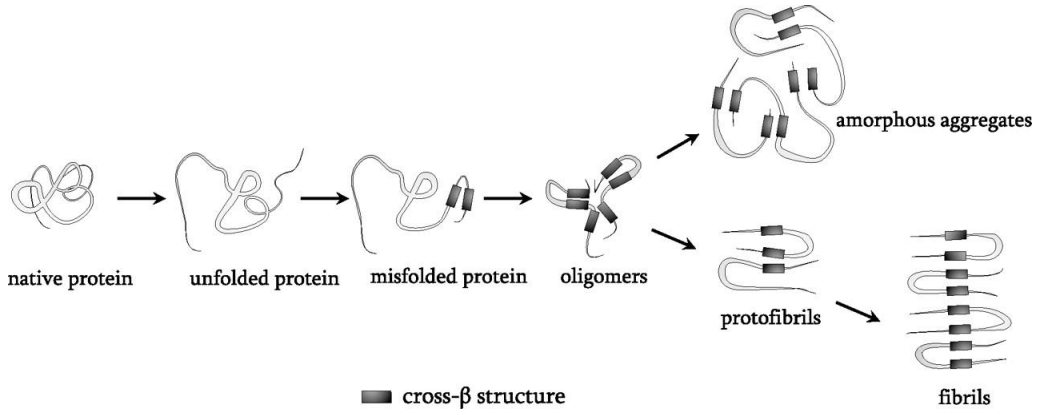


Figure 1-3. Generalized Scheme: the Molecular Basis of Amyloid Formation. Graphics adapted from Gebbink *et al.*[15]

α-Synuclein (αS) Misfolding and Parkinson's Disease PD is the second most common neurodegenerative disorder that affects approximately 7 million people worldwide[20]. It was first described by British doctor James Parkinson in 1817 as a

"shaking palsy"[21], and later fully characterized as a progressive degenerative disorder with motor impairments as the initial symptoms[22]. The physiological basis of PD was discovered in the 1960s, where loss of dopaminergic neurons in the *substantia nigra* were observed by post-mortem analyses of PD patients, leading to the introduction of levodopa as a dopamine supplement (which is still the "gold standard" for PD treatment today)[23]. In 1912, Lewy body inclusions were found as the "pathological hallmark" of PD[24]; in 1997, αS was characterized as the major component of Lewy body inclusions, and the first familial PD mutation A53T was identified.[25] Since its initial discovery, the central role of αS in PD pathogenesis has been extensively studied – currently, the extent of "synucleopathy" is commonly recognized as the most important parameter for PD diagnosis and disease stage classification[26,27].

αS is a 140 residue protein that is abundantly expressed in pre-synaptic terminals, and is estimated to "comprise 1% of total cytosolic proteins in the brain"[28]. Its physiological function has not been fully characterized, but current evidences suggest that it is involved in synaptic transmission[29] – in one study, fluorescently labeled αS was observed to dissociate from synaptic vesicles upon neuronal firing, and then gradually re-associate with newly formed vesicles during resting[30]. The sequence of αS consists of three different regions (Figure 1-4) – an N-terminal region (residues 1−60) that contains seven imperfect repeats of lipid-binding KTKEGV motifs and confers α-helical propensity, a central hydrophobic region (residues 61−95) known as the non-amyloid β component (NAC) domain that has β-sheet formation potential, and a highly acidic C-terminal region

(residues 96-140) that is largely disordered under physiological pH[31]. Known mutations that result in early onset of PD include A30P[32], E46K[33], H50Q[34], G51D[35] and A53T[25].

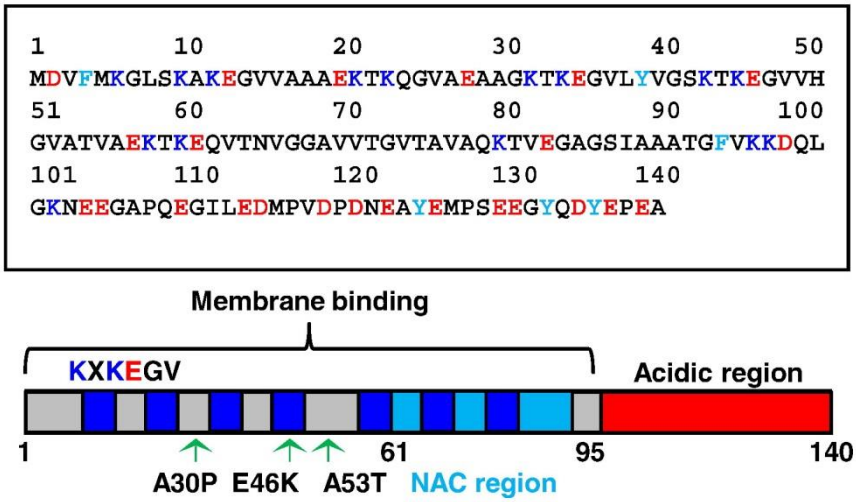


Figure 1-4. Amino Acid Sequence of Human αS.

Graphics adapted from Lee *et al.*[31]

αS can adopt three distinct types of structures depending on its environment − it is α-helical when bound to micelles or vesicles[36], disordered when in aqueous solution[37], and β-sheet rich during *in vitro* or *in vivo* aggregation[38]. In its micelle bound form, αS adopts a well-defined helical structure that has been characterized by both nuclear magnetic resonance (NMR) spectroscopy and electron paramagnetic resonance (EPR) spectrscopy, where two helices (residues 1−41 and 45−94) and a disordered C-terminal region were observed (Figure 1-5)[36,39]. When bound of unilamellar vesicles, the two helices has been shown to stretch into an extended helical conformation in response to membrane curvature by single molecule Förster resonance energy transfer (smFRET) studies from

the Rhoades group[40]. Additional studies using truncated αS fragments determined that the first 25 residues in the N-terminal domain were most crucial for membrane binding[41].

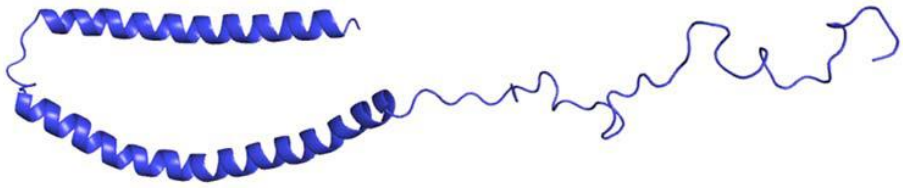


Figure 1-5. Membrane Bound Structure of αS as Determined by NMR (PDB 1XQ8)[36].

In comparison, the structures of lipid-free αS monomer and fibrils are not nearly as defined. αS monomers are largely unstructured as determined by circular dichroism spectroscopy (CD) and Fourier transform infrared spectroscopy (FT-IR)[37]. However, αS does exhibit some level of compaction as evidenced by the smaller radius of gyration than that expected of a fully unfolded conformation (40 vs. 55 Å)[42]. Other studies also suggested the potential for transient tertiary contacts between the N- and C-terminal regions, which were postulated to be protective against αS aggregation.[43,44,45]

The overall morphology of αS fibrils has been well-documented, but their molecular details are poorly understood. Transmission electron microscopy (TEM) studies showed that both patient-derived and *in vitro* aggregated αS fibrils exhibited long, unbranched morphology with a uniform width of 5−10 nm[46,38]; two fibril strands may further assemble into "mature fibrils" with a straight or twisted orientation[47]. At the molecular level, various secondary structure assignments have been proposed by different group

using NMR methods, but a clear three-dimensional structure model has yet to be constructed (Figure 1-6)[47,48,49,50]. Recently, Eisenberg *et al.* used an innovative micro-electron diffraction method and obtained structures of fibrillized αS-derived peptides.[51]

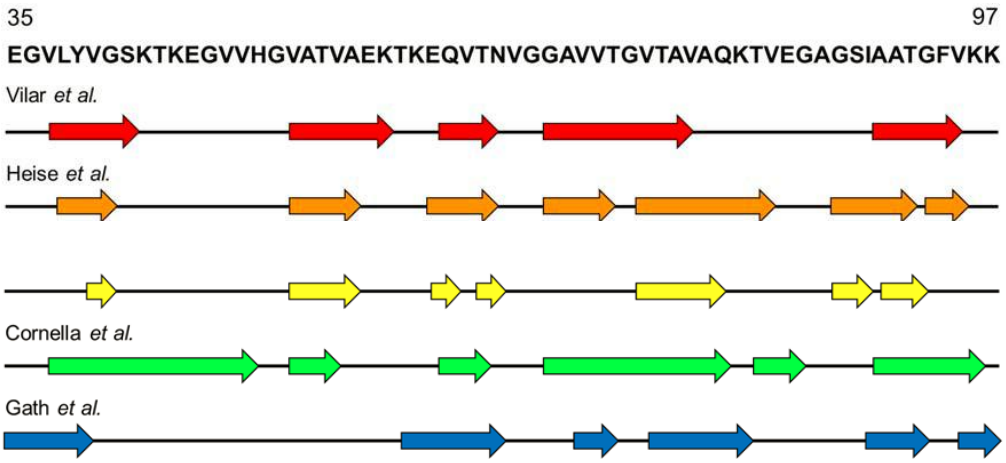


Figure 1-6. Proposed Secondary Structures for αS Fibrils.
Graphics adapted from Bennati *et al.*[52]

In terms of the αS aggregation process where monomers convert into pathological fibrils, even fewer structural details are known. Most studies of αS aggregation were based on global monitoring of the "extent of aggregation", where fibril-binding dyes or band intensities on SDS-PAGE gels were used to quantify monomer/fibril concentrations[46,53]. The most commonly used dye is thioflavin T (ThT) – it consists of two aromatic rings, and exhibits turn-on fluorescence when the relative rotations of the two rings are restricted, preventing, internal charge transfer, as it binds to hydrophobic surfaces on amyloid fibrils[54]. Data from these crude global monitoring methods revealed that αS aggregation was a nucleation-elongation process: it exhibits a lag phase that

corresponds to the formation of nuclei, followed by a rapid growth phase where β-sheet propagation takes place (Figure 1-7)[55,28]. Additional studies showed the significance of seeding and fibril fragmentation, which could significantly accelerate aggregation[56,57]. The aggregation process is also dependent on salts and buffer conditions[58], and most intriguingly, it was found that dopamine itself may interact with αS oligomers/fibrils and contribute to their cytotoxicity[59,60,61] – the current standard of care with levodopa may actually exacerbate PD progression, while temporarily compensating for the diminished neuronal function[62].

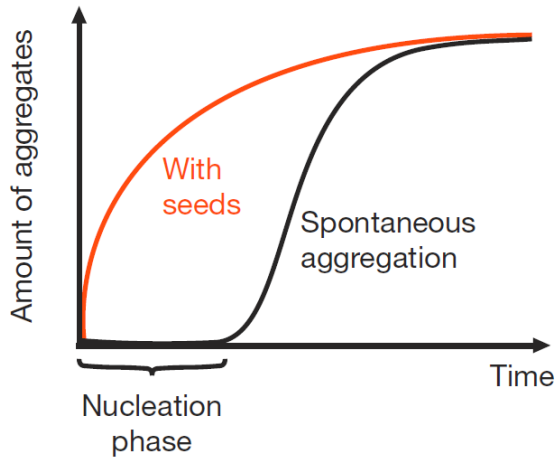


Figure 1-7. αS Aggregation Process as Characterized by Thioflavin T (ThT) Monitoring. Graphics adapted from Jucker *et al.*[18]

Last but not least, αS fibrils were found to transmit from cell to cell, and propagate in the brain in a "spatiotemporal manner"[63,64,65]. In particular, two conformational "strains" of αS fibrils were identified, where strain B evolved from strain A after 6~7 rounds of fibril propagation and gained the ability to cross-seed *tau* aggregation, without exhibiting

much difference in conventional CD or FT-IR characterizations[66]. To date, neither the structures nor the formation processes of these strains are known.

Challenges in Protein Misfolding Studies Obtaining clear structural information on protein misfolding is extremely challenging because of the polymorphism, or heterogeneity, in amyloid structures and formation processes[48,67]. As summarized in Eisenberg *et al.*[53], there are eight possible structure elements in β-sheet rich aggregates, depending on the parallel/antiparallel nature within one β-sheet and the relative orientations of different layers of β-sheets (Figure 1-8). For most misfolded proteins such as αS, there are multiple β-strand segments in the sequences, giving rise to various possible combinations of the structural elements, and thus different fibril conformations (see Figure 1-6). Therefore, it is very difficult to determine fibril structures by conventional structural biology methods such as X-ray crystallography and NMR spectroscopy, which frequently require the preparation of a homogenous material[68].

In comparison, fluorescence spectroscopy, particularly single molecule fluorescence spectroscopy, is a much more powerful tool in the context of protein misfolding studies[69]: 1) it has nanosecond temporal resolution, and can be conducted in real time to capture the dynamic conformational changes in the aggregation process; 2) using Förster resonance energy transfer (FRET, see next section) between a fluorescence donor and an acceptor/quencher, the distance between the two labels can be accurately extrapolated, yielding structural information about aggregation intermediates and the final fibrils; 3) in addition to the most common steady state fluorescence intensity measurements, one could also

utilize florescence life-time measurements, fluorescence correlation spectroscopy (FCS), fluorescence anisotropy, and fluorescence imaging techniques to obtain *in vitro* and *in vivo* properties of the target protein at various stages of aggregation; 4) it is amenable to single molecule measurements, where a heterogeneous ensemble of conformations can be each accounted for as distinct populations in fluorescence spectroscopy. In fact, some of the most useful information about αS monomers and oligomers were obtained through single molecule fluorescence studies[70,71].

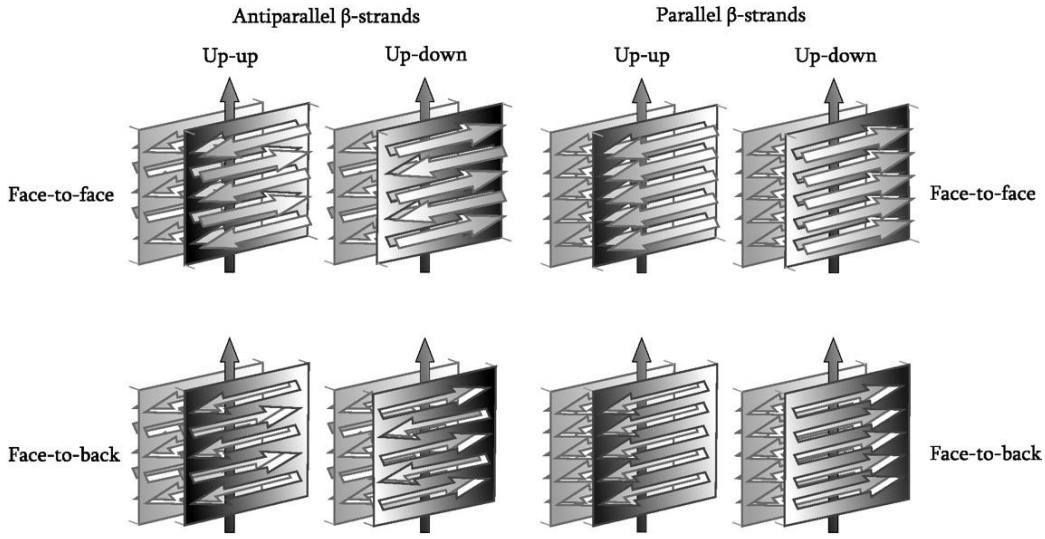


Figure 1-8. Structural Elements of β-Sheet Rich Fibrils.
Adapted from Eisenberg *et al.*[53] and Gebbink *et al.*[15]

Our challenge was to identify and incorporate the appropriate chromophores to realize our vision of characterizing the molecular details behind the protein misfolding phenomenon using fluorescence spectroscopy (we use the term "chromophore" to include both fluorophores and quenchers). In addition to the photophysics and stability require-

ments, our chosen chromophores must also be amenable to positional scanning, i.e. it should be minimally perturbing when incorporated at any desired position, including the β-sheet rich fibril core. Traditional chromophores such as fluorescent proteins (FPs) and organic fluorophores/quenchers are too bulky for such applications[72]. Therefore, our group set out to investigate a novel fluorescence quencher, a thioamide.

1.2 Thioamides as a Minimalist Probes

Properties of the Thioamide Bond A thioamide is a single-atom substitution of the standard amide functional group (herein referred to as "oxoamide"), where the carbonyl oxygen is replaced with a sulfur. It is nearly isosteric to the amide group, while exhibiting different physical and chemical properties (Table 1-2)[73,74]. The sulfur atom is slightly larger than the oxygen atom (van der Waals radii 1.85 Å *vs.* 1.40 Å)[75], and the C=S bond is slightly longer than the C=O bond (1.71 Å *vs.* 1.23 Å)[76]. Due to the additional electron shell in sulfur *vs.* oxygen, the overlap is less strong with carbon $2sp^2$ and 2p orbitals, therefore the C=S bond is weaker than the C=O bond (average bond energy 130 kcal/mol *vs.* 170 kcal/mol)[77] and more polar (5.07 D *vs.* 3.79 D)[78]. The rotation energy barrier for thioamide *cis-trans* isomerization is greater that of oxoamides (22−25 kcal/mol *vs.* 20 kcal/mol) due to "a considerable transfer of charge density from N to S" that elevates energies of the transition states as characterized by Wiberg,[79] which is also reflective of the less electronegative nature of sulfur than oxygen (2.58 vs 3.44)[80].

Thioamides are stronger hydrogen bond donors than oxoamides, but weaker acceptors. Based on experimental studies[81,82] and theoretical modeling[83] from various groups, the C=S ··· H−N hydrogen bond has been characterized as 0.6 Å longer and 2 kcal/mol weaker than its oxoamide counterpart, due to the larger van der Waals radius of sulfur, as well as the less effective orbital overlap between the sulfur $3sp^2$ orbital (as compared to the $2sp^2$ orbital of oxygen) and the hydrogen 1s orbital. In contrast, the CSN−H ··· O=C bond is approximately 1 kcal/mol more favorable, due to more polarized

N−H bond in the thioamide, which is also reflected in its lower pK_a (18.5 vs. 25.5).[84] In addition, the presence of a sulfur atom gives the thioamide higher nucleophilicity[85] as well as greater affinity for soft metals[86] in terms of chemical reactivity.

Of particular interest to our group is the π–π* absorption and oxidation potential of thioamides. The thioamide absorption is red-shifted as compared to that of an oxoamide (270 nm *vs.* 200 nm)[87], allowing the selective excitation of thioamide in a background of oxoamides. Similarly, a thioamide is easier to oxidize than an oxoamide (1.21 V *vs.* 3.29 V)[84], allowing the specific oxidation of thioamide when the reduction potential of the electron acceptor is carefully controlled. The exact photophysical applications of these properties will be discussed in later sections. In early studies of thioamide-containing peptides, the stability of the thioamide group in aqueous buffer was well-established.[88,89] Overall, thioamides are excellent isosteric replacements for the natural oxoamides in peptides and proteins, where one could leverage their unique chemical and physical properties without major concerns on steric perturbation or stability.

Table 1-2. Selected Properties of Thioamides.

Property	Oxo	Thio
Van der Waals radius (Å)	1.40	1.85
C=X bond length (Å)	1.23	1.71
Electronegativity	3.44	2.58
C=X···H–N bond dissociation energy (kcal mol^{-1})	6.1	4.8
π–π* absorption (nm)	200	270
E_{Ox} (V *vs.* S.H.E.)	3.29	1.21

* S.H.E. = Standard hydrogen electrode

Thioamides in Nature and in Medicinal Chemistry In nature, thioamides are found in small molecules and at least one folded protein. To our knowledge, there are five natural products with thioamide bonds, namely cycasthioamide from the seeds of *Cycas revoluta*[90], (4-methoxyphenyl)-N-methyl-2-oxothioacetamide from *Polycarpa aurata*[91], apo-methanobactin from *Methylosinus trichosporium*[92], thioviridamide from *Streptomyces olivoviridis*[93], and closthioamide from *Clostridium cellulolyticum*[94]. Among these compounds, closthioamide has been most thoroughly characterized (Figure 1-9)[94,95] – it has six thioamide bonds and exhibits strong antibiotic activity against *S. aureus*; the compound was fully inactive when all thioamides were substituted with oxoamides, and a significant reduction in potency was observed even with a single thio-to-oxo substitution. A recent study showed that the biological activity of closthioamide was related to its copper binding ability, which explained the importance of the thioamides[96]. The only known protein with a thioamide bond is methyl-coenzyme M reductase, an enzyme involved in archeal methane formation[97]. It was shown to possess a thioglycine residue near its active site by X-ray crystallography and mass spectrometry, which presumably facilitates the oxidation and reduction of cofactor and substrate.

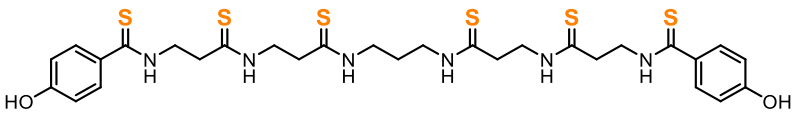
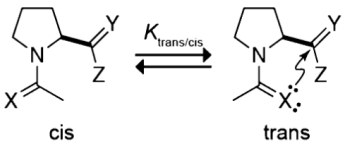


Figure 1-9. Chemical Structure of Closthioamide.

In medicinal chemistry, thioamides are frequently used as an amide bond derivative for structure-activity relationship (SAR) studies of natural products; they may also be subsequently derivatized into thiazoles, thiazolines, thiazines, or amidines to expand the candidate library[98]. Representative examples include thioamide derivatization of oxytocin by du Vigneaud and coworkers[99], of leucine enkephalin by Lawesson and colleagues[100], and of vancomycin by the Boger group[101]. Interestingly, in the case of leucine enkephalin, thioamide substitution at Tyr1 completely eliminated its biological activity, while substitution at Gly2 enhanced its activity by 3−14 fold[100]. There are several approved drugs that contain thioamides, the most notable of which is 6-*n*-propyl-2-thiouracil (PTU) that derives its anti-hyperthyroidism activity from the thioamide moiety[102].

Early Applications of Thioamides in Peptides and Proteins An early example for applying thioamides to studying peptides and proteins was a carboxypeptidase A kinetic analysis, where thioamide-modified short peptide substrates were evaluated for their enzymatic activity.[103] With a thioalanine or thiophenylalanine (Ala^S or Phe^S; we use a superscript "S" to denote thioamide-containing residues) at the scissile bond, the authors demonstrated that thioamides can function as effective substrates. More interestingly, utilizing the affinity of sulfur to soft metals, they were able to identify cadmium as a more effective metal cofactor as compared to the native zinc, where the k_{cat}/K_m was 2.4–9.7 fold higher for the Cd(II) carboxypeptidase than for the Zn(II) carboxypeptidase when thioamide-containing peptides were used as substrates.

More recently, Raines *et al.* have utilized the differences in the electronic properties of C=S and C=O bonds to study the n-π* interactions between two adjacent peptide bonds.[104] The n-π* interaction is characterized as the donation of one of the non-bonding lone pairs (n) of a carbonyl into the antibonding orbital (π*) of an adjacent carbonyl. Since sulfur is a better electron donor, the n-π* interaction is proved to be stronger for thioamides (Figure 1-10). Based on these data, the authors argued that n-π* interactions were a protein stabilizing force that were complimentary to hydrogen bonds[105].



Compound	X	Y	Z
1	O	O	OMe
2	S	O	OMe
3	O	O	NMe_2
4	S	O	NMe_2
5	O	S	NMe_2
6	S	S	NMe_2

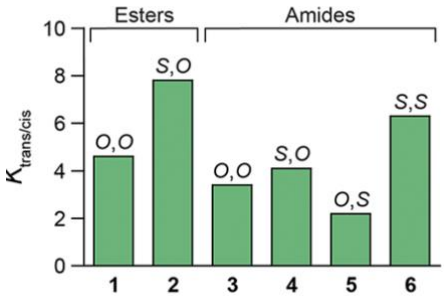


Figure 1-10. Strength of n-π* Interactions as Measured in *cis-trans* Equilibrium[106].

Several direct studies on the effects of thioamides in protein secondary structures have also been conducted, where thioamides were incorporated into different model peptides for conformational studies. Reiner *et al.* incorporated a thioamide near the N-terminus of an α-helical peptide, and demonstrated its structural similarity to the oxoamide peptide using NMR spectroscopy, as well as a 1.6 kcal/mol stabilization to the oxopeptide[88]. Miwa *et al.* conducted a similar study with a thioamide near the C-terminus of a model α-helical peptide, and reached the same conclusion that the thiopeptide

exhibited a nearly identical structure to the oxopeptide[89]. The same group also tested the compatibility of thioamides with β-sheets, by incorporating a thioglycine (Gly^S) at the β-turn in a β-hairpin model peptide, and showed structural compatibility[107]. As a follow-up, Culik *et al.* incorporated thioamides into other positions in a β-hairpin tryptophan zipper, and showed that the substitution was well-tolerated except when it was placed directly next to the β-turn[108]. Most recently, Raines *et al.* incorporated thioamides into the collagen triple helix and observed minimal perturbation to its structural stability; in fact, the authors noted that the thioamide was "the first in the collagen backbone that does not compromise thermostability"[109]. These studies showed that while thioamides may subtly affect folding pathways, they are generally well-tolerated in these secondary structures.

Thioamide as a Minimalist Fluorescence Quencher Thioamides are particularly appealing modifications for our vision of using a "minimalist" chromophore to monitor protein misfolding. With the abundance of peptide bonds and side chain oxoamides in peptides and proteins, one could envision placing thioamides at any residue along the protein sequence as probes for mechanistic and/or structural studies. Earlier experiments showed that thioamides exhibit unique circular dichroism signature around 272 nm[89] and can be selectively photo-isomerized[110] in a background of oxoamides. In our own group, we have extensively explored the utility of thioamides as fluorescence quenchers through FRET or photo-induced electron transfer (PET).

FRET is a non-radiative energy transfer phenomenon, where the excited state of a donor fluorophore transfers its energy to a ground state acceptor fluorophore; as a result

of this process, the donor will relax back to ground state, while the acceptor enters an excited state and subsequently relaxes back either through release of a photon or through thermal or collisional relaxation (see Figure 1-11)[69]. The energy transfer efficiency (E_Q) is distance dependent, and can be described by the following equations:

$$E_Q = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} = 1 - \frac{F}{F_0}$$

$$R_0^6 = \frac{9000(ln10)\kappa^2 \Phi_D J}{128\pi^5 n^4 N_A}$$

where R is the distance between acceptor and donor fluorophores, and R_0 is the Förster distance at which the transfer efficiency is 50%. R_0 can be experimentally determined from E_Q or theoretically calculated – experimentally, E_Q can be measured from donor fluorescence in the presence (F) or absence (F_0) of the acceptor chromophore, and fitted to the Förster equation; theoretically, the parameters needed include the spectral overlap integral between donor emission and acceptor absorption (J), an orientation factor to describe the interaction of the transition dipoles of the donor and the acceptor (κ^2), the quantum yield of the donor (Φ_D), index of refraction of the solvent (n), and Avogadro's number (N_A).

UV wavelength dyes, such as *p*-cyanophenylalanine (Cnf) and the natural amino acid Tyr, are quenched by thioamides through a FRET mechanism, where the emission spectra of these chromophores exhibit overlap with the absorption spectra of thioamides. [111,112] Using rigid "polyproline rulers", we determined the empirical Föster distances (R_0) to be 16.5 Å for Cnf, and 16.2 Å for Tyr (Figure 1-11). These R_0 values translate for an

effective monitoring distance of 10−25 Å (where E_Q is approximately between 10% and 90%), a very useful scale to monitor protein conformational changes at the secondary or tertiary structure level, as well as tight protein-protein binding interactions. As proof-of-concept examples, we also demonstrated the application of the Cnf-thioamide fluorophore-quencher pair to monitoring thermal denaturation of a 35-residue model peptide through intramolecular quenching[111], and the utilization of the Tyr-thioamide pair in quantifying protein-peptide binding through intermolecular quenching[112].

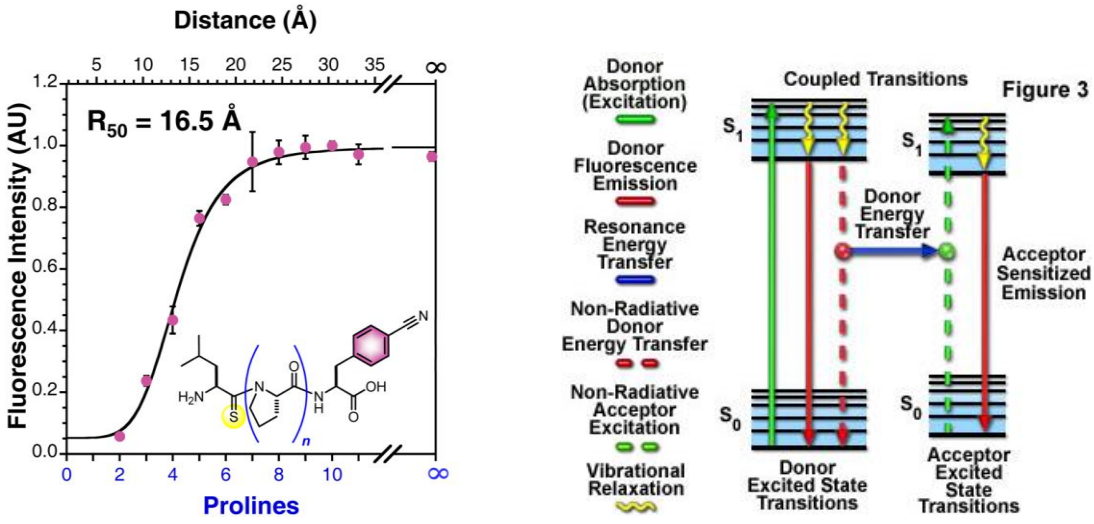


Figure 1-11. Thioamide Quenching of Fluorescence through FRET.

Left: distance dependent quenching as measured by "proline ruler". Right: Jablonski diagram showing the FRET process. Graphics adapted from *Goldberg et al.*[111-112] and Olympus Inc.

PET is also a distance-dependent non-radiative energy transfer phenomenon, but is achieved through electron transfers between the excited state donor and the ground state acceptor chromophore[69]. For two chromophores to be a PET pair, they must have matching oxidation/reduction potentials, and be present in sufficient proximity to enable

the physical transfer of an electron. PET favorability can be determined by the Gibbs free energy in the Rehm-Weller model[113] with the following equation:

$$\Delta G_{ET} = F\left[E_{ox}(D) - E_{red}(A) - E_{0,0}\right] + C$$

where F is the Faraday constant; E_{ox}(D) and E_{red}(A) are the fround state oxidation and reduction potentials of the electron donor and acceptor (the fluorescence donor can be either the electron donor or the electron acceptor), respectively; $E_{0,0}$ is the zero vibrational electronic excitation energy of the fluorophore, calculated as the average energy of the absorption and emission; and C is a term for Coulombic interactions that are typically assumed to be negligible in aqueous solutions. In the case of fluorescence quenching by a thioamide, the fluorophore is typically the electron acceptor, and the thioamide is the electron donor. With an oxidation potential of 1.21 V (*vs.* S.H.E.), thioamides were found to quench a number of near UV and visible wavelength fluorophores (Figure 1-12), including commonly used dyes such as Alexa Fluor 488 (ΔG_{ET} = - 0.86 eV) and BODIPY FL (ΔG_{ET} = - 0.37 eV)[114,115]. As a proof-of-concept example, we synthesized short peptides that were dually labeled with a fluorophore and a thioamide, and showed that they were effective substrates for proteases such as calpain both *in vitro* and in cell lysate[116]. Work is underway in our group to further pursue this strategy as a tool to construct minimally-perturbing protease sensors.

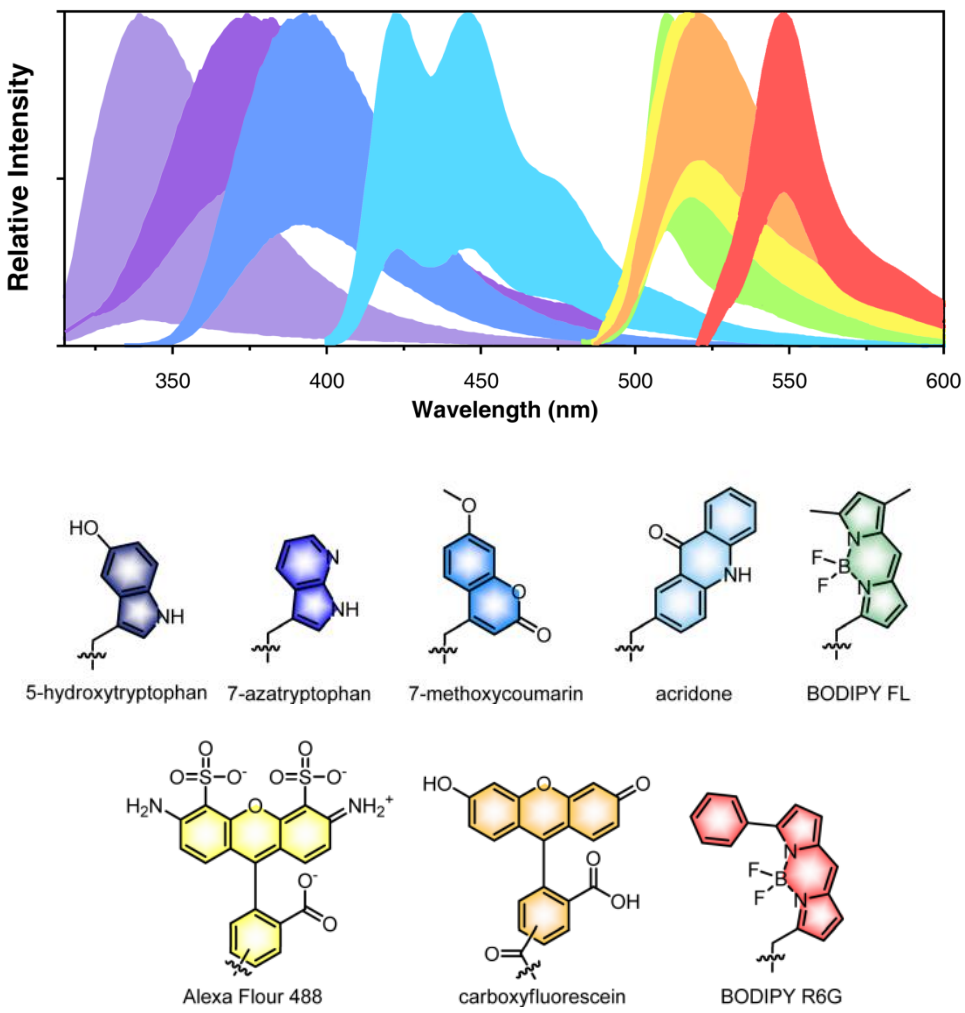


Figure 1-12. Thioamide Quenching of Fluorescence through PET.

Graphic adapted from Goldberg *et al.*[114,115,116] and Petersson *et al.*[74]

In the context of protein misfolding, we envision using thioamide as a minimalist label for "protein motion capture" (Figure 1-13). The target protein can be labeled with fluorophore-thioamide pairs at various positions, and then fluorescence changes can be monitored throughout the aggregation process and in the mature fibrils. The quenching efficiencies observed will then be converted into distances, and by combining results from various combinations of labeling positions, we can reconstruct the conformational

changes that underlie the misfolding process. The fact that the thioamide is a nearly isosteric substitution of the native amide bond is extremely important to the successful application of this strategy − we can move the thioamide across the peptide backbone in a "positional scanning" manner while keeping the fluorophore at a relatively non-perturbing position, which would not be possible with any other organic fluorophores or fluorescent proteins (FPs).

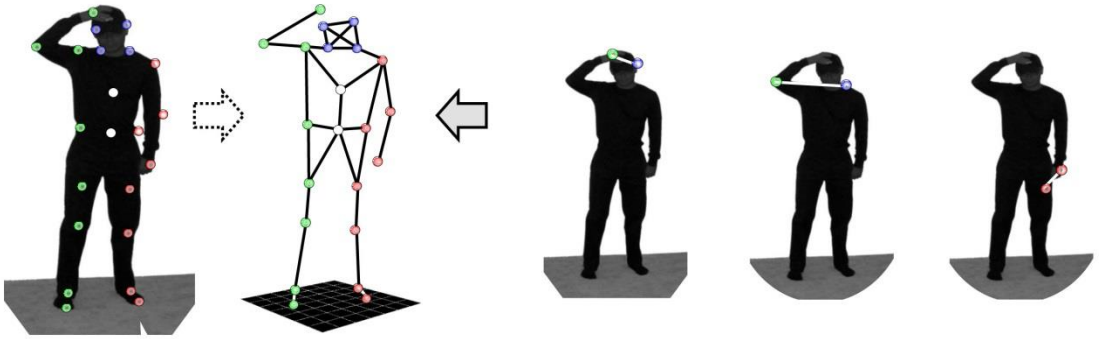


Figure 1-13. A Conceptual Representation of "Protein Motion Capture".

Analogous to motion capture in movie making, proteins can be tagged with fluorophore-quencher pairs at various positions, and then subject to *in vitro* or *in vivo* misfolding. The distances between the fluorophore and the quencher can be monitored throughout the process by fluorescence quenching; results from all labeling positions can be combined to recapitulate the "motions", or conformational changes, involved in the misfolding process.

Thioamide Incorporation into Small Molecules and Peptides In order to utilize the thioamide as a minimalist chromophore, we must first be able to introduce it into the peptide or protein of interest. Prior to our work (as shown in Chapters 2 to 4), thioamides could only be incorporated into small molecules by solution phase thionation, or into short peptides by solid phase peptide synthesis (SPPS).

To our knowledge, the first synthetic thioamide-containing compound was prepared by Gay-Lussac in 1815 from cyanogen and hydrogen sulfide[117]. In modern organic synthesis, thioamides can be prepared from a variety of precursors including amides, aldehydes, ketones and isothiocyanates, which has been extensively reviewed by several authors[118,119]. The most applicable type of transformation for our studies is the direct conversion of an oxoamide into thioamide. The two commonly used reagents are phosphorous pentasulfide (P_2S_5 or P_4S_{10})[120,121] and Lawesson's reagent[122,123] – both reactions involve a concerted sulfur transfer in a Wittig-type intermediate (Figure 1-14). P_2S_5 can be used at lower temperature than Lawesson's reagent (which typically requires reflux for activation), but is also more difficult to remove by flash chromatography[119].

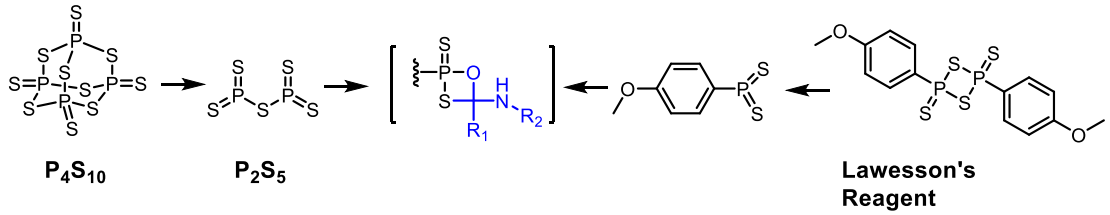


Figure 1-14. Structures and Mechanisms of Common Thionating Reagents. Adapted from Ozturk *et al.*[120] and Jesberger *et al.*[123]

While thionation by P_2S_5 or Lawesson's reagent is applicable to short peptides, it will indiscriminately convert all oxoamides into thioamides (or in some cases, result in a mixture of peptides with various degrees of thionation)[119]. To site-selectively incorporate thioamides into peptides, one would need to adopt SPPS, where a preactivated thioamide

precursor is first synthesized and then coupled to the growing peptide chain through carbonyl substitution reactions (Figure 1-15). Rapoport *et al.* pioneered the synthesis of *N*-Boc (Boc = *t*- butyloxycarbonyl) protected thioamide precursors, where 19 out of the 20 natural amino acids (except Gly) were successfully prepared[124]. More recently, Chatterjee *et al.* systematically evaluated the preparation of *N*-Fmoc (Fmoc = fluorenylmethyloxy-carbonyl) protected thioamide precursors and their coupling conditions, where they synthesized 12 thoiamide precursors for *D*- and *L*-amino acids and identified CH_2Cl_2 as the preferred solvent for thioamide coupling in SPPS[125]. In our own group, we have successfully prepared 12 thioamide precursors using a similar strategy.

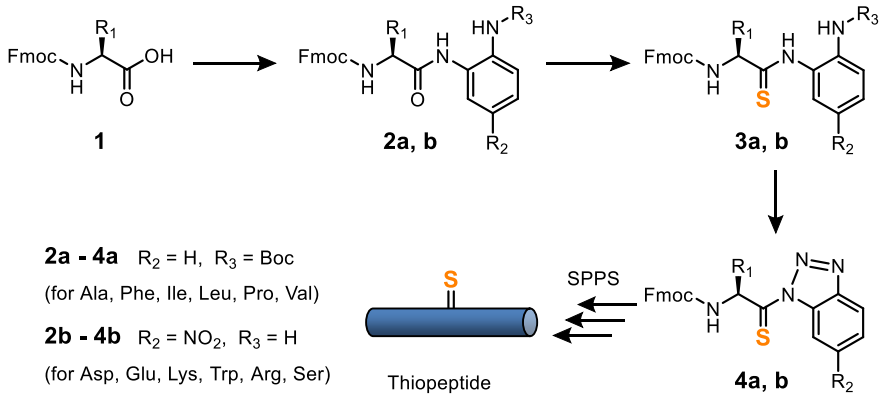


Figure 1-15. General Scheme for Thioamide Incorporation through SPPS.

1.3 Protein Semi-Synthesis Methods

Protein semi-synthesis is the chemoselective condensation of two or more peptides/ proteins fragments. It has been a tremendously useful tool in incorporating synthetic moieties that would otherwise not be accessible through cellular expression or single-chain SPPS, for the structural and functional studies of long peptides and full-length proteins[126]. In the context of thioamide-containing protein preparation, we will focus on methods that result in an amide bond at the ligation junction in the following discussions. Overview of other methods − including thioether ligation, oxime ligation, and azide-alkyne cycloaddition – can be found in a detailed review by Verzele and Madder[127].

Native Chemical Ligation (NCL) NCL was first introduced by Kent *et al.*[128] in 1994, and quickly become the most widely adopted protein semi-synthesis method in the last two decades. It utilized two unprotected peptide fragments in aqueous buffer − one with a C-terminal thioester, the other with an N-terminal Cys. The thiol side chain of Cys first undergoes a transthioesterification with the thioester fragment, and then an S-to-N acyl shift takes place to form a native amide bond at the junction (Figure 1-16). The mild reaction conditions are compatible with chemically sensitive groups such as glycosylated, phosphorylated, or ubiquitinated side chains[129]. The most common applications for NCL are[130]: 1) accessing long peptide sequences that are not possible to construct in a single-chain SPPS; 2) incorporation of unnatural modification into peptides and proteins, that would otherwise not be accessible through cellular expression; 3) studying the properties and functions of peptides/proteins with natural or unnatural substitutions at specific sites,

which would otherwise be obtained as a mixture either in cell extract or through other derivatization methods; 4) preparation of cyclized peptides and peptide analogs through head-to-tail NCL reactions for activity screening. Even for short peptides that can be synthesized directly through SPPS, NCL still offers two levels of benefits: 1) for chemically sensitive groups, one could reduce the exposure of the peptide to harsh synthesis conditions (e.g. trifluoroacetic acid, or TFA) by synthesizing shorter fragments that can be rapidly deprotected; 2) for peptides that are functionalized at more than one site, one could use combinatorial synthesis for the efficient generation of peptide libraries. A notable example of NCL applications is the total chemical synthesis of αS by Lashuel *et al.* − after establishing the feasibility of αS total synthesis by NCL[131], they further elucidated the effects of site-specific αS phosphorylation[132], ubiquitination[133], and nitration[134] on αS aggregation by synthesizing the corresponding "post-translationally" modified αS. Another exemplary work was from Muir *et al.*, where they synthesized large libraries of modified histones *via* NCL and studied their biological functions[135].

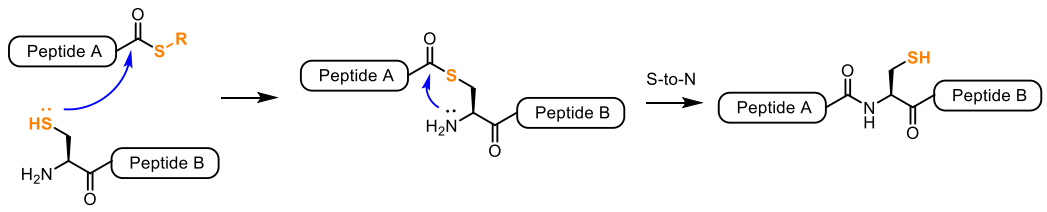


Figure 1-16. Mechanism of Native Chemical Ligation (NCL).

In practice, NCL is frequently catalyzed by thiol additives. A systematic study by Kent *et al.*[136] characterized the reactivity of various thiols, and identified that aromatic thiols such as 4-mercaptophenylacetic acid (MPAA) as the most effective catalyst.

Extended NCL Strategies Major extensions of the NCL method include tandem ligation, expressed protein ligation (EPL) and reverse native chemical ligation. Tandem ligation was developed by Kent *et al.*[137] using thiazolidine (Thz) as a masked Cys. The middle fragment of the target protein was synthesized with an N-terminal Thz and C-terminal thioester to prevent self-ligation, and then joined to the C-terminal fragment. After the first ligation, the Thz group was deprotected using $MeONH_2 \cdot HCl$, and then subjected to another round of ligation with the N-terminal fragment. The tandem ligation can be repeated to combine more than three fragments, and is frequently used in the synthesis of large proteins[135].

Expressed protein ligation (EPL) takes advantage of naturally-existing inteins to generate protein thioesters, and greatly expands the ligation scope − prior to EPL, ligation sites were largely restricted to the N-terminal region, as the thioester fragment could only be chemically synthesized[130]. An intein is a protein fragment that exhibits self-splicing properties – in nature, a protein sequence can be expressed with exteins and inteins (similar to the exons and introns of nucleic acids), after which the inteins would self-excise, leaving just the exteins in the final protein[138]. The process starts with an N-to-S acyl transfer at the N-terminal Cys of the intein fragment, followed by a transthio-esterification that joins the N-terminal extein with the C-terminal extein; the final step is the formation of a succinimide at the C-terminus of intein to release the α-nitrogen of the C-terminal extein (Figure 1-17)[139]. Muir *et al.* first reported the use of intein-mediated EPL in 1998; without a C-terminal extein, the intein would remain as a thioester and

serve as a substrate for subsequent ligation to a synthetic peptide[140]. Additional developments using naturally existing[141] or evolved split inteins[142] allowed the extension of EPL into *in vivo* systems; for example, a cell surface receptor can be expressed with one intein fragment, and then treated with a fluorescent protein fused to the other intein fragment, to achieve selective labeling on the cell surface (in comparison, traditional fusion protein expression will result in high background fluorescence from mis-trafficked proteins in the cytosol)[143].

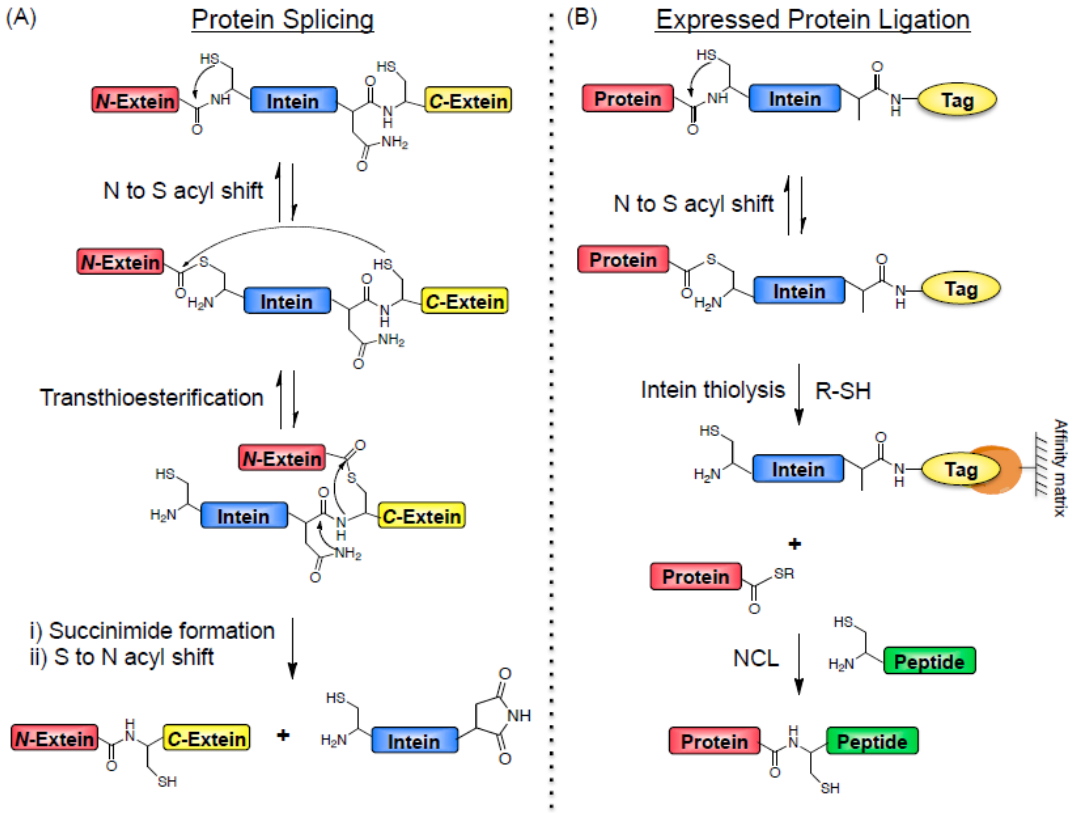


Figure 1-17. Intein-Mediate Protein Splicing and Expressed Protein Ligation. Picture courtesy of Solongo Batjargal.

Finally, reverse native chemical ligation is a similar technique to NCL, but uses a thioacid and a side chain alkylhalide as the ligation handles instead. In the first step, a thioester is formed by nucleophilic substitution on the alkylhalide, which subsequently rearranges through standard S-to-N acyl shift to generate the amide bond, leaving a Cys at the junction. Since both ligation handles are not accessible through cellular expression, this method is primarily used for conjugating peptides[127].

Traceless Staudinger Ligation While NCL is limited to an N-terminal Cys (or thiol surrogates as discussed in later sections) as the nucleophile, traceless Staudinger ligation has no inherent limit on the N-terminal residue of the C-terminal fragment. Developed by Raines and Bertozzi[144,145,146], it utilizes the well-known Staudinger reaction where an azide reacts with a phosphine to form an iminophosphorane. In the first step, a phosphinothioester is generated *in situ* from a thioester; the azide then reacts with the phosphine moiety to generate an iminophosphorane, which then undergoes intramolecular carbonyl substitution to form the final product (Figure 1-18). While

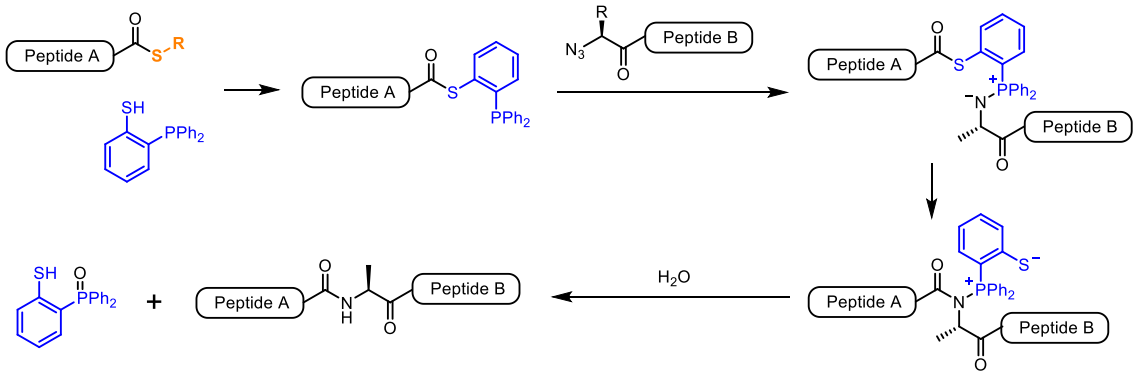


Figure 1-18. Traceless Staudinger Ligation.

chemically elegant, this method suffers from several drawbacks including facile air oxidation of the phosphine and the necessity to generate the azide-containing fragment through chemical synthesis.

Thioacid-Azide Ligation A thioacid can directly react with an azide to form an amide bond through a thiatriazoline intermediate (Figure 1-19)[147]. Recent studies showed that this reaction can be performed in various organic solvents as well as in water, and can be successfully applied to the synthesis of β-glycosylamides from glycosyl azides[148,149]. While no applications in peptides have been shown, it has been proposed as a phosphine-free alternative to traceless Staudinger ligation[130]. Interestingly, this reaction can also be utilized in the synthesis of thioamides; however, the reaction conditions needs to be carefully controlled to suppress competing oxoamide formation [150].

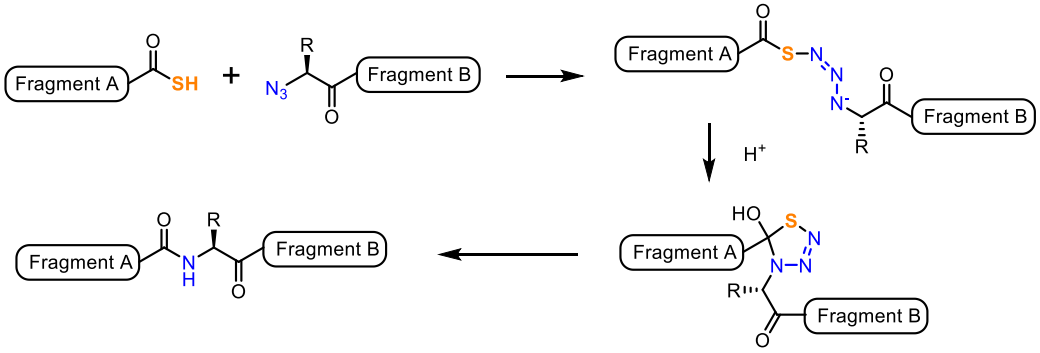


Figure 1-19. Mechanism of Thioacid-Azide Ligation.

α-Ketoacid–Hydroxylamine Amide-Forming (KAHA) Ligation Bode *et al.* devised another amide-bond-forming ligation method by conjugating an α-ketoacid to an *N*-alkylhydroxylamine[151]. Mechanistic studies revealed that the reaction started with a

nucleophilic attack of the *N*-alkylhydroxylamine nitrogen on the α-ketoacid α-carbonyl, followed by a transfer of its hydroxyl (or benzylhydroxy) group on to the carbonyl carbon. The formation of the amide bond was then achieved through collapse of the tetrahedral intermediate with release of a CO_2 (Figure 1-20)[152]. KAHA reactions can be classified into two types – Type I utilizes an unsubstituted hydroxylamine and is conducted in organic solvents, while Type II uses an *O*-benzyl hydroxylamine and can be performed under aqueous conditions[151]. Recent studies from the Bode group also showed the feasibility of using 5-oxaproline and oxazetidine as the source of the hydroxylamine, which would result in a homoserine or serine, respectively, at the ligation site[153,154]. Similar to the traceless Staudinger ligation, this method has limited application in full-length proteins because the ligation handles need to be chemically incorporated.

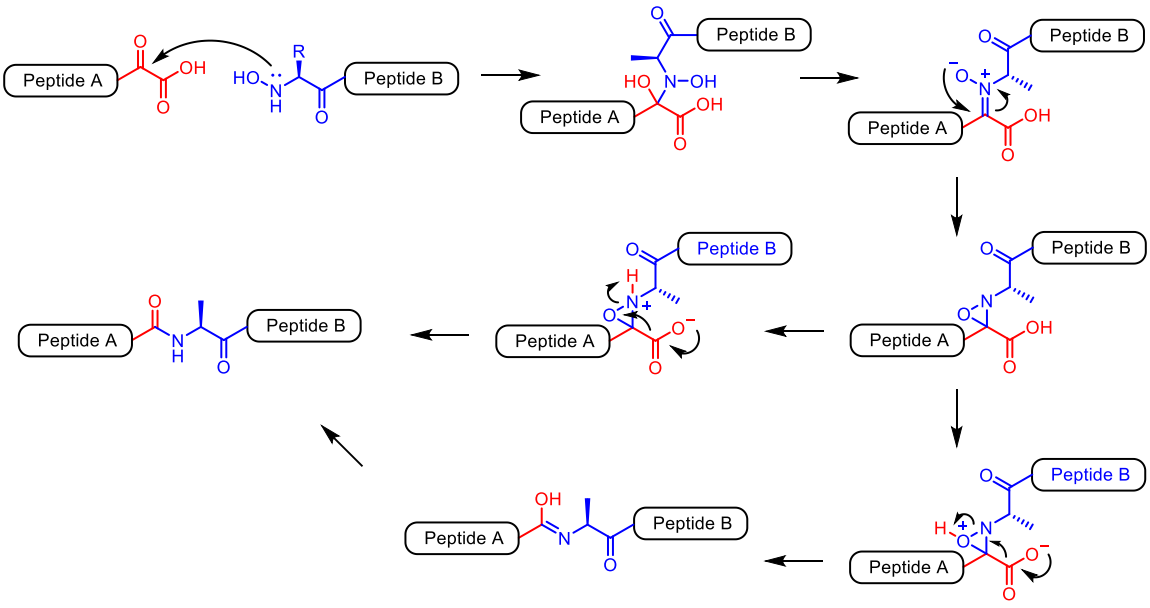


Figure 1-20. Mechanism for α-Ketoacid–Hydroxylamine Amide-Forming (KAHA) Ligtaion.

1.4 Methods for Peptide Thioester Synthesis

Thioester Synthesis by PyBOP Activation Thioesters are typically synthesized by carbonyl substitution reactions from a more reactive precursor, such as an activated amide, ester or anhydride[155]. In peptide chemistry, this can be achieved through activating the C-terminal carboxylic acid *in situ* with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), and then performing nucleophilic substitution with a free thiol (Figure 1-21).[156] Due to the harsh reaction conditions, the reaction needs to be conducted on a fully protected peptide – in particular, all carboxylic acid moieties and nucleophiles must be fully protected to avoid side reactions. In a typical procedure, a peptide is synthesized on 2-chlorotritylchloride resin and then cleaved under mild acidic conditions with 10% acetic acid or 1% TFA to preserve all the protecting groups. The resulting protected peptide is very hydrophobic and poorly soluble, which greatly limits the reaction efficiency and hinders its application to long sequences. In addition, this activation method frequently results in epimerization – the activation group also serves as a good leaving group for intramolecular cyclization, which subsequently tautomerizes and results in steric scrambling of the α-carbon. Although PyBOP has been identified as the least epimerizing reagent (as compared to other coupling reagents such as HBTU, HCTU)[157], it still exhibits 5% epimerization in short peptides, which can be exacerbated with longer reaction time or difficult sequences.

SPPS Methods using Unmasked Thioesters Early on-resin syntheses of peptide thioesters was achieved using Boc-based SPPS, where successive cycles of coupling and

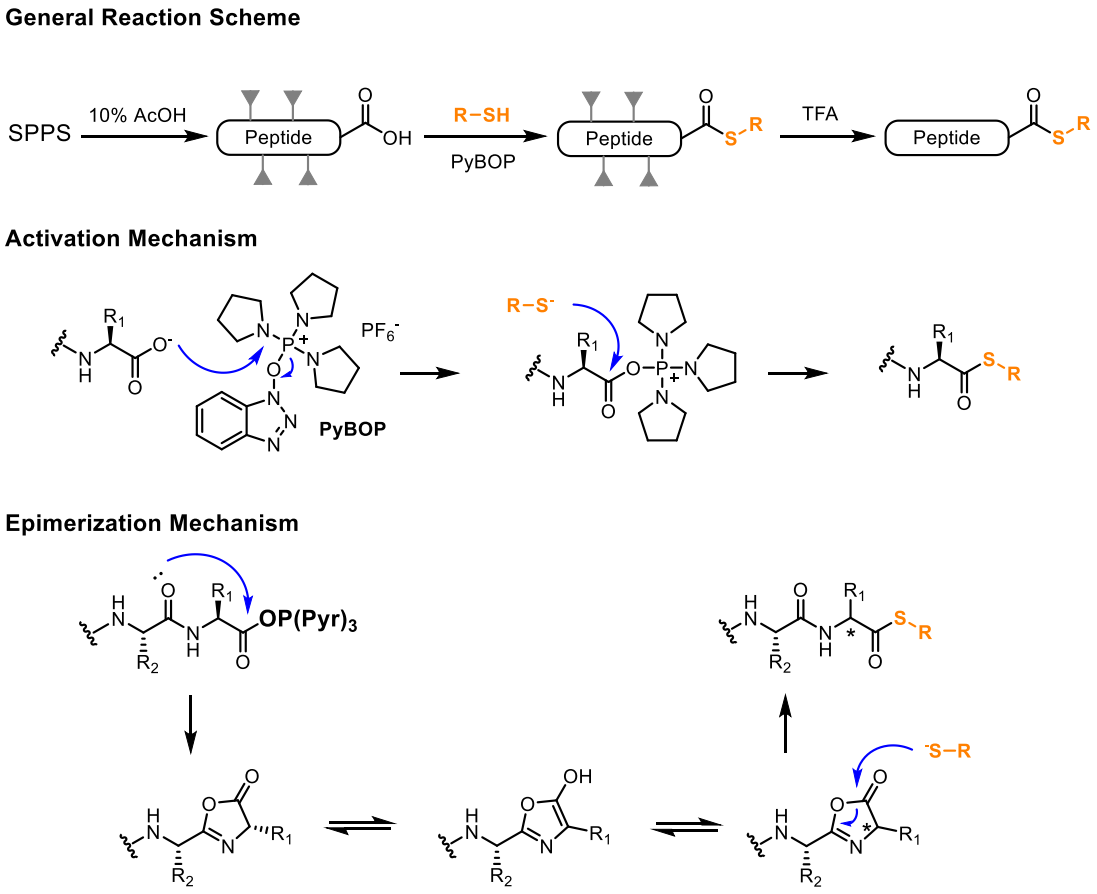


Figure 1-21. Peptide Thioester Synthesis by PyBOP Activation.

Adapted from Nagalingam *et al.*[157] and von Eggelkraut-Gottanka *et al.*[156]

TFA Boc removal were used to elongate the peptide. The thioester moiety is sufficiently stable under acid conditions to allow the direct elongation of peptide on a thioester linkage (Figure 1-22A)[158]. However, the requirement of harsh hydrofluoric acid cleavages at the end of the synthesis precludes its application to chemically sensitive groups such as thioamides. On the other hand, thioesters are not stable under the basic conditions used in Fmoc-based SPPS, particularly in the presence of nucleophilic bases such as piperidine used in the Fmoc deprotection steps. Recently, Rademann *et al.*

attempted to use a tertiary thiol linker, hypothesizing that the steric constraints around the thioester bond would reduce the rate of its degradation (Figure 1-22B)[159]. With a $t_{1/2}$ of 6.5 days for the thioester in this method, it was proven successful on model peptides, but did not fully resolve the degradation issue.

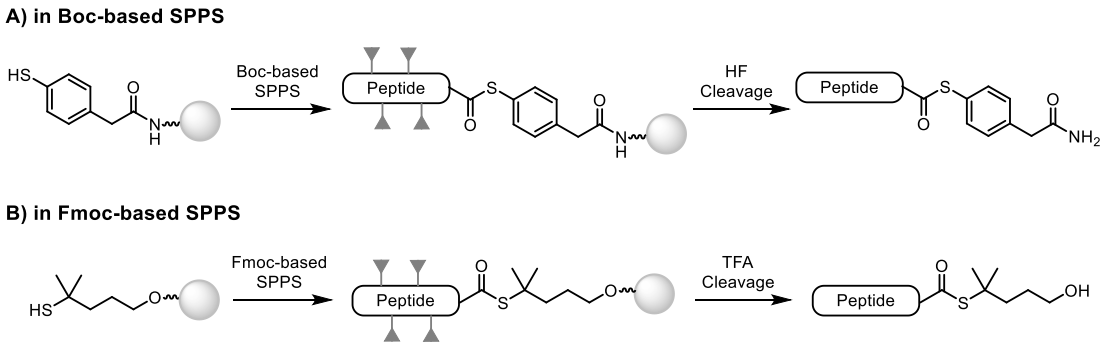


Figure 1-22. SPPS Methods using Unmasked Thioester Linkages.
Adapted from Kent *et al.*[158] and Radmann *et al.*[159]

"Safety-Catch" Type Linkers The "safety-catch" method was first introduced by Pessi *et al.* for their strategy of anchoring the growing peptide chain to the resin *via* a non-cleavable linkage such as a sulfonamide, and then activating the linkage at the end of synthesis to form a good leaving group, which can subsequently undergo thiolysis to yield a thioester [160]. Various types of linkers have been developed under this principle as summarized in Figure 1-23. In Pessi *et al.*, a sulfonamide linkage was used, which was subsequently alkylated by trimethylsilyldiazomethane or iodoacetonitrile to yield an *N*-alkyl-peptidylsulfonamide as a good leaving group for thiolysis[160]. Jensen *et al.* used an orthogonally protected Glu residue, which was selectively deprotected at the end of

synthesis and cyclized to the adjacent amide using harsh PyBrOP activation[161]. Dawson *et al.* introduced 3,4-diaminobenzoic acid (Dbz)[162] – only the *meta*-amine was acylated for peptide elongation under standard SPPS conditions; the *para*-amine was suppressed due to the electron-withdrawing effect of the amide group on the aromatic ring. Upon completion of SPPS, the *para*-amine was acylated by *p*-nitrocholoroformate, leading to a spontaneous ring closure to form an *N*-acyl benzimidazolinone that was labile towards thiolysis. Two mdified versions of Dbz were also prepared – one by Mahto *et al.*, where the *para*-amine was protected by an allyl orthogonal protecting group that can be selectively removed prior to the final acylation[163], the other by Dawson *et al.* where methylation was used to lower the *para*-amine nucleophilicity[164]. Finally, Ficht *et al.* took a slightly different approach and utilized side-chain anchoring instead of post-SPPS thiolysis. Rather than being anchored to the resin through the peptide C-terminus, the desired peptide was attached to the resin through a Glu side chain. At the end of synthesis, the C-terminal carboxylic acid was selectively deprotected, thioesterified by PyBOP, and then globally cleaved to yield the peptide thioester[165,166]. While it is possible to attach additional amino acids/thioesters to the C-terminus, a Glu residue is still needed either as the C-terminal residue or one residue away from the C-terminus.

All these strategies are clever in by-passing the thioester stability problem, but they invariably suffer from low yields due to reactive site accessibility[167] – since the activation is performed at the end of SPPS, reagents have to diffuse through a "crowded" resin decorated with protected peptide in order to reach the C-terminus. Due to the harsh conditions used for linker activation, these methods also require the N-terminus to be

acetylated to prevent side reactions, which is not always desirable since the N-terminal acetylation of proteins has been shown to affect their properties[168]. In the context of thioamide incorporation, since the harsh linker activation chemistry will take place after installation of the nucleophilic thioamide moiety into the peptide, these methods are not as desirable as the latent thioester strategies that will be discussed next.

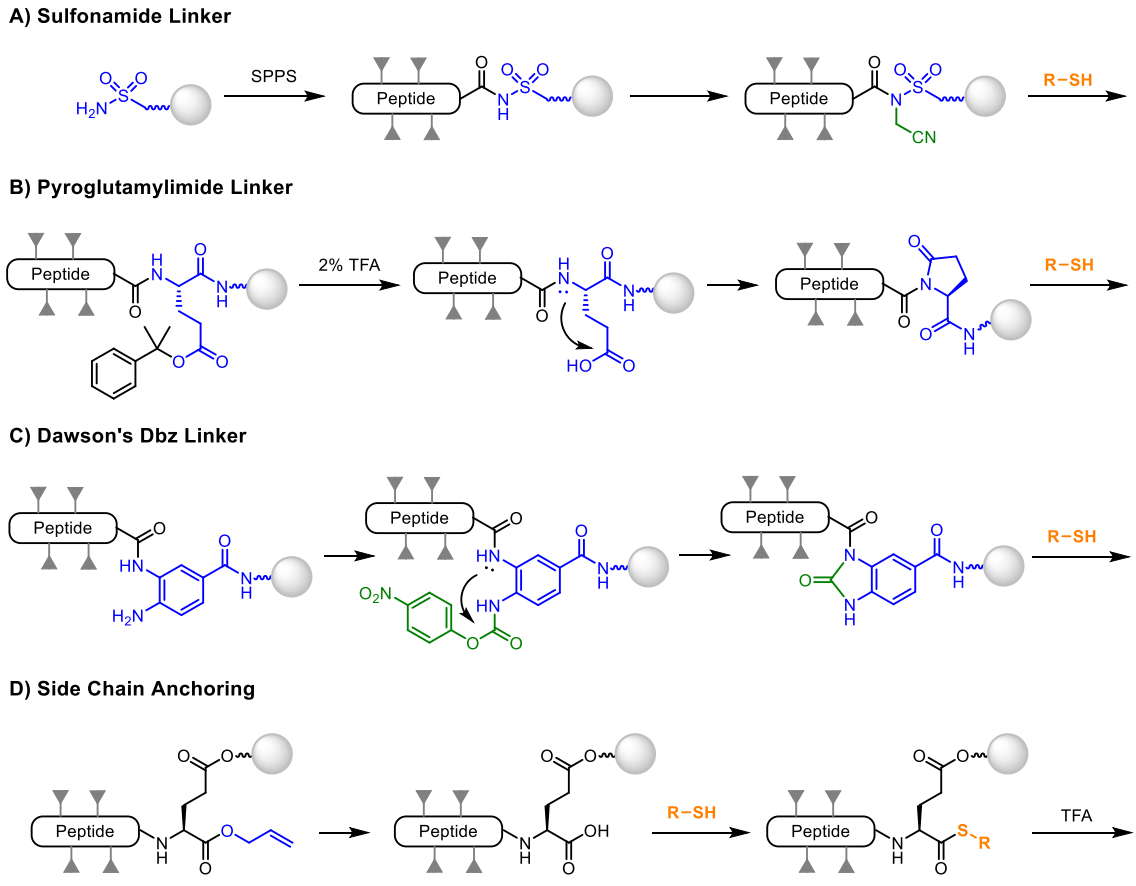


Figure 1-23. Selected "Safety Catch" Type Linkers.

N-to-S Latent Thioester Linkers Taking inspiration from intein-mediated thioester formation in nature, various groups have designed linkers that utilize an amide

bond as the resin anchor, which is then activated *in situ* through a spontaneous or assisted N-to-S acyl transfer for conversion into a thioester (Figure 1-24A). Due to the reversible nature of N-to-S acyl shift, various auxiliaries are frequently necessary to promote the formation of thioester (Figure 1-25).

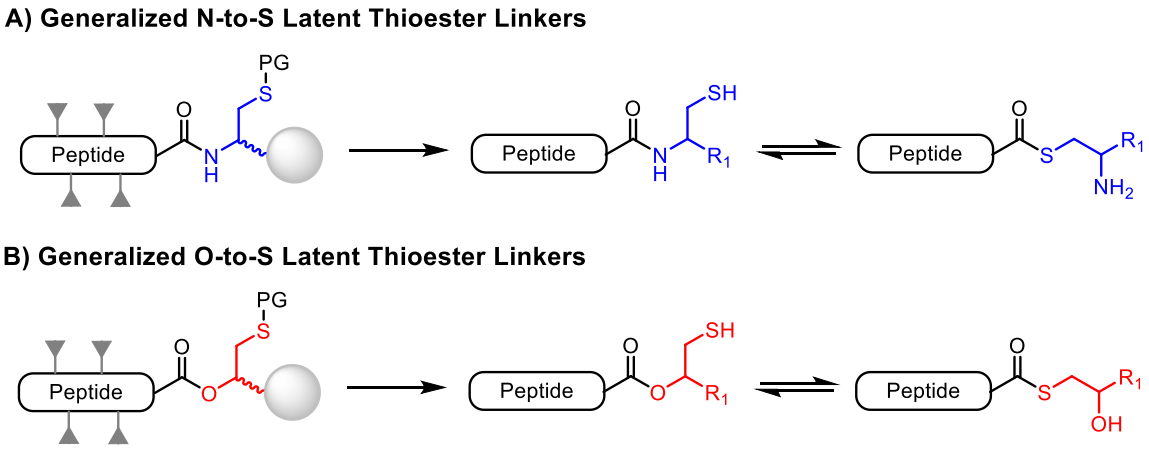


Figure 1-24. General Mechanism for N-to-S and O-to-S Latent Thioester Linkers. Protecting groups are designated as "PG" or grey triangles.

Aimoto *et al.* developed a 2-mercapto-4,5-dimethoxybenzyl (Dmmb) auxiliary, where the tertiary amide nitrogen served as the leaving group, and an adjacent aromatic thiol was the attacking thiol[169]. It has been applied in several systems, where the longest thioester made was a 41 residue thioester[170]. The same group also developed a Cys-Pro ester (CPE) linker, where the N-to-S acyl shift is promoted by the irreversible formation of a diketopiperazine (DKP) moiety[171,172,173]. Hojo *et al.* designed a mercaptoprolyl-prolyl ester motif and successfully synthesized a 25 residue thioester; the method was originally designed to promote DKP formation, but was found to proceed through a

simple N-to-S acyl shift instead[174]. As a follow-up, they were able to remove the second Pro, and obtained comparable N-to-S acyl shift efficiency; notably, they also demonstrated the compatibility of this particular linker with microwave assisted synthesis[174]. Nakahara *et al.* showed the use of a simple *N*-alkylated Cys as a latent thioester, and identified that an ethyl substitution was most effective in promoting the thioester formation (34%)[175]. Otaka *et al.* identified acyl oxazolidinone as a potential linker[176], and later optimized it into an *N*-substituted aniline linker, where the aniline served as an even better leaving group[177,178]. Using the second method, a short nine reisude phosphopeptide was synthesized in 67% yield[177].

Melnyk *et al.* explored bis(2-sulfanylethyl)amido (SEA) as a tunable linker, where the N-to-S acyl shift can be turned off and on simply by the oxidation or reduction of the di-thiol moiety[179]. Offer et al. showed the utility of α-methylcysteine in accelerating the N-to-S acyl shift; the extra methyl group on the Cys α-carbon favored the necessary cyclic transition state through a Thorpe-Ingold effect[180], enriching the population of the molecule that was in the Burgi-Dunitz trajectory[181]. Liu *et al.* further synthesized alkene analogs of Cys, and showed that it was also a viable linker[182]. Recently, Aucagne *et al.* constructed a self-catalyzing linker with a side chain hydroxyl group that mimicked the natural intein; this linker was also shown to be compatible with microwave assisted SPPS[183]. Finally, Macmillan *et al.* demonstrated the similar application of selenocysteine in the N-to-Se acyl shift, which was more effective at lower temperature due to the higher reactivity of the selenol[184]. For our initial study in thioamide-containing peptide thioester synthesis, we chose the CPE linker[171] because of its synthetic accessibility.

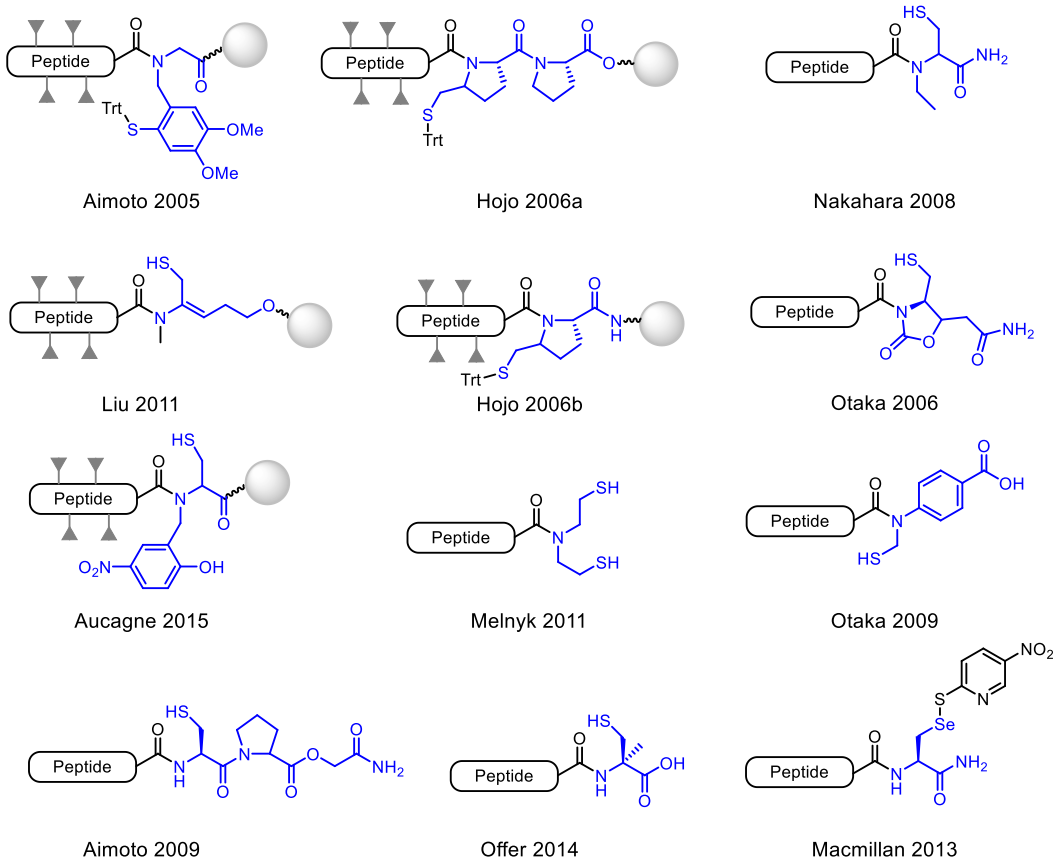


Figure 1-25. Overview of N-to-S Latent Thioester Linkers.

An N-to-Se linker from Macmillan *et al.*[184] is also included for comparison.

O-to-S Latent Thioester Linkers Similar to their N-to-S counterparts, O-to-S latent thioester linkers utilize an ester bond to anchor the growing peptide to the resin, which is activated *in situ* through O-to-S acyl transfer (Figure 1-24B). The ester bond is less stable than the amide bond in N-to-S linkers; therefore, O-to-S linkers are less robust for long sequences that require repeated exposure to nucleophilic base in SPPS, but are also faster in generating an active thioester through O-to-S acyl shift.

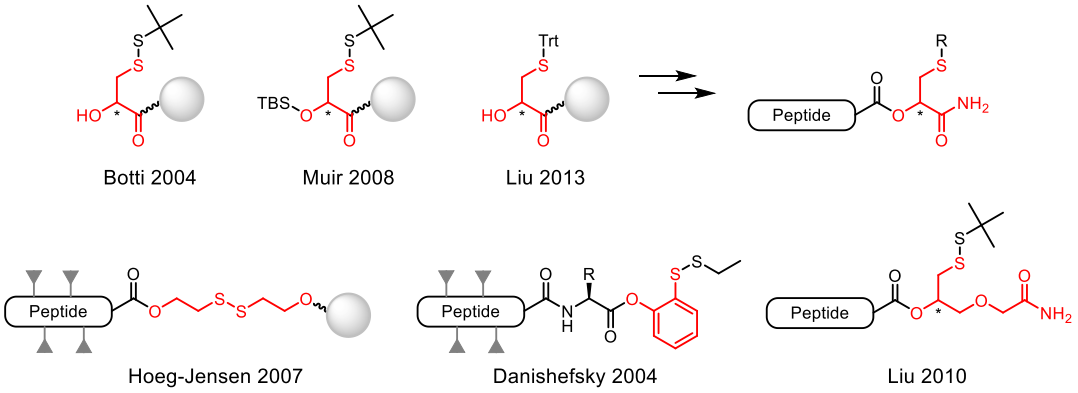


Figure 1-26. Overview of O-to-S Latent Thioester Linkers.

There are four O-to-S scaffolds introduced to date, namely α-hydroxycysteine, β-mercaptoethanol, thioglycerol and hydroxythiophenol (Figure 1-26). Botti *et al.* first synthesized the α-hydroxycysteine linker by hydroxylation of cysteine on resin, and successfully ligated two model peptide fragments using this method[185]; while it was a great proof-of-concept demonstration, the method required a water-compatible PEG-based resin, which was fragile and difficult to handle for routine SPPS. *Muir et al.* subsequently improved the synthesis and introduced a protected precursor (which we termed TBS-ChB-OH), where the thiol was capped with a *t*-butylthio group and the α-hydroxyl was protected by a silyl group[186]. Liu *et al.* used a different route and prepared trityl protected α-hydroxycysteine monomer through ring opening of an epoxide[187]. The β-mercaptoethanol scaffold was explored by Hoeg-Jensen *et al.*, where it was introduced as its dimer, dithiodiethanol, and activated *in situ* with reducing agents[188]. The thioglycerol moiety was identified by Liu et al. after screening a variety of O-to-S linkers for rate of NCL over hydrolysis[189]. The hydroxythiophenol scaffold was introduced by

Danishefsky *et al.*, where it was conjugated to the last residue in its protected form, and then attached to the rest of a 5-residue peptide through solution phase coupling[190]. While this linker was not directly conjugated to the resin in this particular case, one could easily foresee its adaptation to SPPS by introducing a carboxylic handle on the phenyl ring as a resin anchor. For our initial trial, we chose to adapt the α-hydroxycysteine scaffold with the protected precursor approach by Muir *et al.*[186]

Hydrazide as a Thioester Precursor In 2011, Liu *et al.* introduced a different approach to synthesizing latent thioesters by using a hydrazide linkage[191]. The hydrazide linkage can be conveniently introduced by treating 2-chlorotrityl chloride resin with hydrazine, after which the peptide can be elongated using standard SPPS procedures. After isolating the purified hydrazide peptide, it is activated *in situ* through $NaNO_2$ treatment and immediately thiolyzed to generate an active thioester (Figure 1-27). This method has been shown to be compatible with the majority of natural amino acids (except Asn, Asp, and Gln) as the C-terminal reisdue, and has been successfully applied to the semi-synthesis of αS[192].

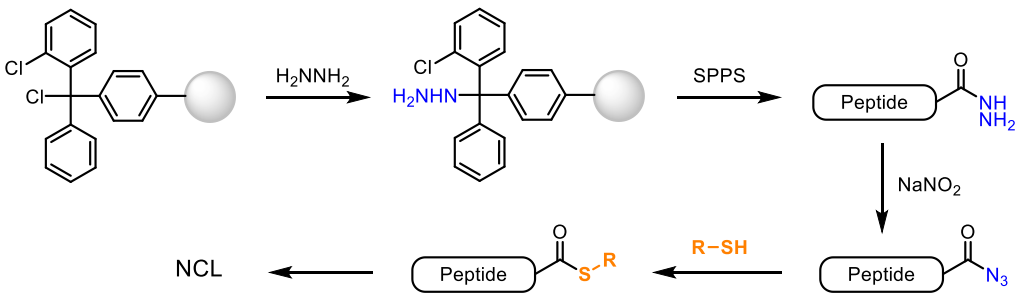


Figure 1-27. Hydrazides as Thioester Precursors.

1.5 Traceless Ligation Methods

Masking Methods The residual Cys after NCL can be further derivatized into analogs of Lys, Glu and Gln through nucleophilic substitution (Figure 1-28). Treatment of Cys with aziridine or 2-bromoethylamine results in *S*-alkylation to form an isosteric Lys analog.[193] Similarly, treatment with iodoacetamide[194] or iodoacetic acid[195] yielded a Gln or Glu analog, respectively, with thioether insertion in the side chains. While these methods are straight-forward to implement (i.e. they do not require synthetic ligation handles), they result in modified side chains with thioether linkages, and are not truly "traceless". One of the first synthetic ligation handles to result in a native amino acid is homocysteine (Hcs), which yields a native Met upon alkylation with methyl iodide or methyl *p*-nitrobenzenesulfonate[196]. Interestingly, Roelfes and Hilvert applied the same strategy to selenohomocysteine, and successfully generated a selenomethionine in a peptide hormone analog through NCL followed by alkylation[197].

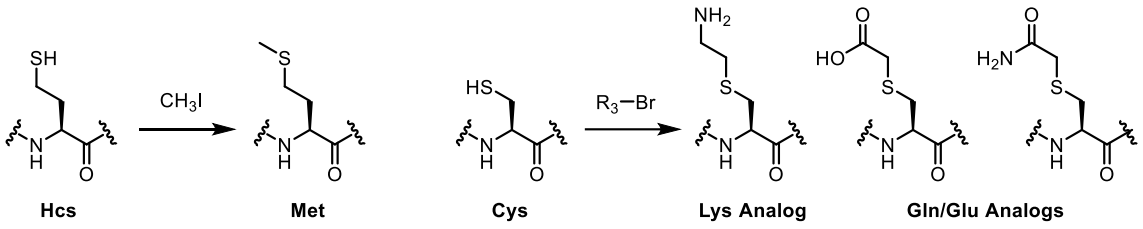


Figure 1-28. Masking Methods for Traceless Ligation.

Desulfurization Methods An on-going paradigm shift in NCL is the introduction of thiol-bearing amino acid analogs as ligation handles[198]. The thiol serves as the

nucleophile for NCL, which is removed after ligation to generate a native amino acid. The first desulfurization was reported by Dawson *et al.* in 2001 for the conversion of Cys into Ala using Raney nickel[199]. Using this approach, they were able to synthesize a 110-amino acid ribonuclease, barnase. However, the Raney nickel method *per se* is not translatable to other amino acids due to the lack of selectivity (for example, hydrogenation of the Trp indole and demethythiolation of Met are two known complications with Raney nickel)[198]. In 2007, Danishefsky *et al.* used radical initiated desulfurization by a water soluble organic initiator VA-044 as an alternative approach.[200] In the same year, Crich *et al.* synthesized the first β-thiol amino acid analog, β-mercaptophenylalanine, and successfully demonstrated its desulfurization[201]. Since then, a number of thiol analogs have been introduced for 12 out of all 20 natural amino acids (Figure 1-29).

In 2008, Danishefky *et al.* and Seitz *et al.* independently prepared two types of valine analogs, γ-mercaptovaline[202] and β-mercaptovaline[203]. Two types of Lys analogs were also synthesized by two different groups, namely γ-mercaptolysine by Liu *et al.* in 2007[204], and δ-mercaptolysine by Brik *et al.* in 2009[205]. Interestingly, these Lys analogs were designed with orthogonal protecting groups for the side chain amine so that they can be used for multiple ligations to elongate the backbone and introduce ubiquitin on the side chain − the N-terminal amine was first ligated to one thioester fragment, after which the side chain amine was revealed through selective deprotection and then conjugated to another thioester; finally the thiol auxiliary was removed through desulfurization to yield a native Lys at the branching site. In 2010, three additional analogs were introduced by the Danishefsky group, namely β-mercaptoleucine[206], γ-mercaptothreonine[207], and γ-

mercaptoproline[208]. In 2012, Brik *et al.* followed up with a report on the synthesis of γ-mercaptoglutamine[209]. Most recently, the Payne group has made much contribution to the analog repertoire by adding five additional analogs, namely β-mercaptoarginine[210], β-mercaptoglutamic acid[211], β-mercaptoaspartic acid[212], β-mercaptoasparagine[213], and a Trp analog with thiol on its indole ring[214]. With the success of these 14 analogs for 12 amino acids, the β- and γ-thiol strategy can conceivably be extended to 19 out of all 20 amino acids (except for Gly that does not have a β-carbon, which would require an α-thiol and four-membered ring transition state in NCL), which is corroborated by a computational study on the desulfurization reaction[215].

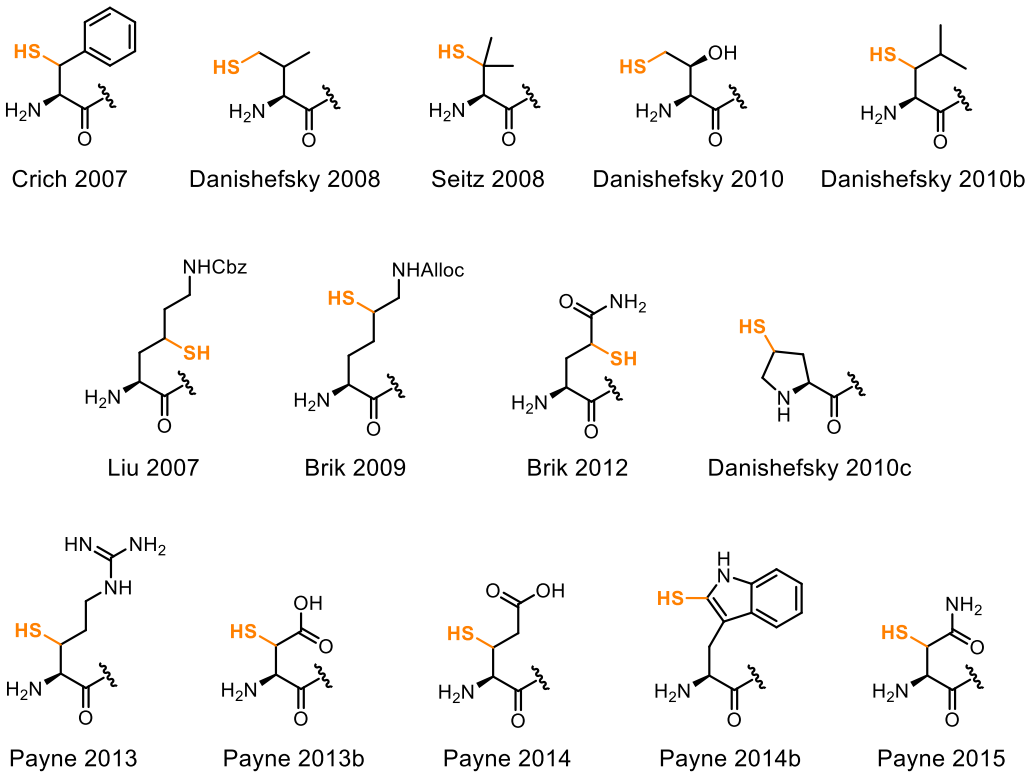


Figure 1-29. Summary of Thiol Analogs of Natural Amino Acids.

Deselenization Methods Similar to thiols, selenols may undergo radical initiated deselenization reactions, which can be utilized in traceless NCL methods with β-selenol amino acid analogs. The selenium-carbon bond is weaker than the sulfur-carbon bond due to less effective orbital overlap, which makes the deselenization reaction easier. In the first demonstration of traceless NCL through deselenization, Dawson *et al.* showed that deselenization could be initiated by spontaneous homolytic bond cleavage of a diselenide or hemiselenide, which then proceeded to severe the C−Se bond with the assistance of tris(2-carboxyethyl)phosphine (TCEP)[216]. Under these mild conditions, Cys residues in the peptide sequence remained intact, allowing the traceless ligation of peptides and proteins that contain native Cys residues in their sequence. In addition to the naturally existing selenocysteine, there are two synthetic selenol amino acids analogs reported thus far (Figure 1-30), namely γ-selenoproline by Danishefsky *et al.*[217] and β-selenophenyl-alanine by Payne *et al.*[218]. Due to the reactive nature of the selenol, peptides containing these analogs were typically isolated as dimers with intermolecular Se−Se bond, and then activated *in situ* by reduction[219].

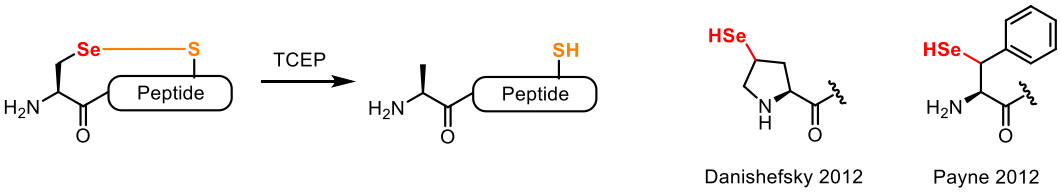


Figure 1-30. Post-ligation Deselenization of Selenocysteine (Sec) and Selenol Analogs.

A unique property of selenols is that they can be directly converted into hydroxyls under oxidative conditions[220]. This allows the post-ligation conversion of Sec into Ser,

adding one additional amino acid as potential ligation sites. The first demonstration of its application to peptide ligation was recently conducted by Payne *et al.*, where they constructed a glycol-peptide by Sec ligation and subsequent conversion to Ser[221]. A recent mechanistic study also confirmed the "dual personality" of selenol – when treated with TCEP under anaerobic conditions, Sec was predominantly converted into Ala; when oxygen or ozone was supplemented, however, the alaninyl radical could be quenched by oxygen instead, generating Ser-containing peptide as the major product (Figure 1-31)[222].

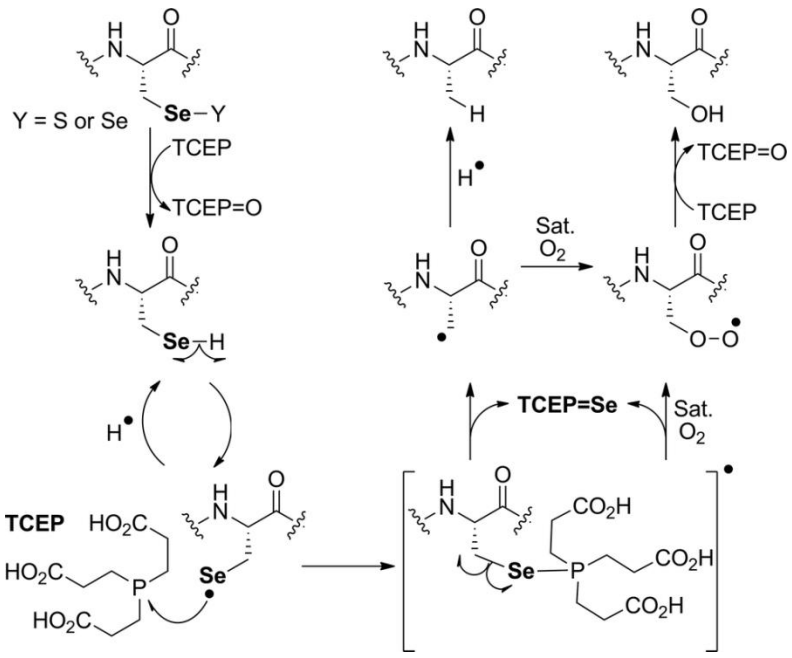


Figure 1-31. Proposed Mechanism of Oxidative Conversion of Sec into Ser. Graphics adapted from Metanis *et al.*[222]

1.6 Selenocysteine

Properties of Selenocysteine Similar to the relationship between thioamides and oxoamides, selenocysteine (Sec) is nearly isosteric to Cys, but exhibits very different properties (Table 1-3)[223]. With an additional shell of electrons, selenium is slightly larger than sulfur (1.90 Å *vs.* 1.85 Å);[75] the C−Se bond is slightly longer (1.96 Å *vs.* 1.80 Å)[224] and weaker (bond dissociation energy 56 kcal mol^{-1} vs. 65 kcal mol^{-1})[77] than the C−S bond. The Sec side chain pK_a is much lower than that of Cys (5.2 *vs.* 8.3)[225] – at physiological pH, the majority of selenol will exist in the ionized form of selenoate. The diselenide/hemeselenide exchange reaction is also faster[226] – the reaction rate constant of RS−SR exchange with thiol is 3.6 M^{-1} s^{-1}, while that of RS−SeR is three times faster at 11 M^{-1} s^{-1}; the corresponding reaction for RSe−SeR is 10^5 faster at 1.3×10^5 M^{-1} s^{-1}. Finally, although selenium and sulfur have similar electronegativity (2.58 *vs.* 2.55)[80], diselenides are much easier to reduce than disulfides (E_{red} -0.233 V *vs.* -0.488 V)[227].

Table 1-3. Selected Properties of Cysteine and Selenocysteine.

Property	**Cys** (X = S)	**Sec** (X = Se)
Van der Waals radius of X (Å)	1.85	1.90
C−X bond length (Å)	1.80	1.96
C−X bond dissociation energy (kcal·mol^{-1})	65	56
−XH pKa	8.3	5.2
RS−XR exchange rate constant (M^{-1} S^{-1})	3.6	11
E_{red} for RX–XR (V *vs.* S.H.E.)	-0.488	-0.233

* S.H.E. = Standard hydrogen electrode

With these properties, Sec is a much more reactive amino acid than Cys; as exemplified in the mutagenesis study of thioredoxin reductase, when the native Cys-Sec catalytic motif of the enzyme was replaced with a Cys-Cys motif, a 20-fold reduction in catalytic efficiency was observed (k_{cat}/K_m 1.3 × 10^7 *vs.* 6.0 × 10^5 M^{-1} s^{-1})[228]. While the majority of the reactions that Sec participates in are analogous to those of Cys, Sec undergoes a unique reaction that is termed "glutathione peroxidase (GPx) like activity" – when a selenol is oxidized into hydroxyselenol, it can be reversed simply by the addition of a thiol (Figure 1-32)[220]. This reactivity has been observed both in selenol-containing small molecules and in Sec-containing enzymes[229]; when handling Sec-containing peptides, we typically add excess thiol to protect the selenol side chain against oxidation.

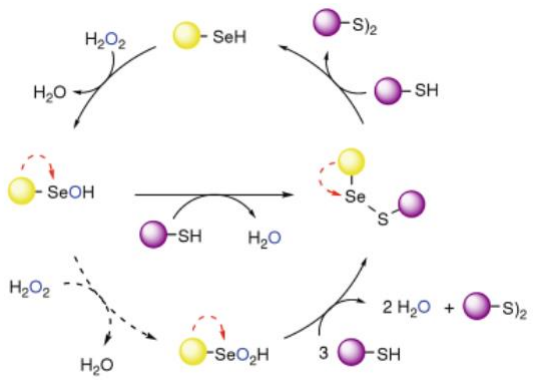


Figure 1-32. Glutathione Peroxidase (GPx) Like Activity of Selenols. Graphics adapted from Derek *et al.*[220]

Selected Applications of Selenocysteine Selenocysteine is a useful tool in protein mechanistic studies and in biotechnology[230]. First of all, Sec is a unique reaction handle both for NCL and for other conjugation chemistry. Its reactivity in NCL was rigorously

characterized by Raines *et al.* in 2001, where they discovered that due to the low pK_a of Sec, the ligation reaction can be conducted at pH ranges between 5 and 8, as compared to pH 7−8 for Cys NCL[231]. They also found that the reaction rates were generally 3 times faster, with a second order reaction rate of 9.5×10^2 $M^{-1}s^{-1}$ for Sec as compared 3.7×10^2 M^{-1} s^{-1} for Cys at neutral pH. Sec has been incorporated into a variety of target proteins through NCL, including into full-length proteins by EPL[232]. A most representative example is the semi-synthetic azurin studies by van der Donk *et al.*,[233] where they incorporated Sec as a replacement of the native Cys112 through EPL, and used it as a probe to elucidate the active site copper coordination geometry.

With its distinct redox properties compared to natural amino acids, Sec can also be used as a site-specific label to study oxidative folding[234,227]. Moroder *et al.* used Sec to elucidate the structure of apamin, an 18 residue bee venom toxin[235]. With four Cys in its sequence, apamin can adopt a variety of conformations with different disulfide bond combinations; the authors systematically replaced Cys pairs with Sec pairs to isolate each conformation (Sec will preferentially form diselenide with another Sec before forming a hemiselenide with Cys), and solved the NMR structure of each disulfide/diselenide combination. In another example, Bulaj *et al.* replaced a disulfide bonded Cys pair with a diselenide bonded Sec pair in a background of six Cys on conotoxin; the other four Cys then correctly folded into a single, functional conformation rather than forming a mixture of various disulfides due to the templating effect of the first diselenide[236].

Another widely explored application of Sec is ^{77}Se NMR spectroscopy. Selenium has six stable natural isotopes, the most abundant of which are the NMR inactive ^{78}Se and ^{80}Se. The NMR active ^{77}Se has a natural abundance of 7.63%, and can be enriched for NMR spectroscopy[220]. Early studies used ^{77}Se-6,6'-diselenobis(3-nitrobenzoic acid), the selenium version of Ellman's reagent, to covalently label Cys residues of the desired proteins[237,238]. Hilvert *et al.* prepared semi-synthetic ^{77}Se-enriched subtilisin, and directly observed the formation of selenide (−Se$^-$), hydroxyselenol (−Se−OH) and selenoic acid (−SeOOH) as intermediates of its catalytic cycle[239]. Recently, Rozovsky *et al.* developed a metabolic replacement method, where a target Cys containing protein was expressed in *E. coli* in media with minimal sulfates and ^{77}Se selenite supplements[240]. Up to 80% ^{77}Se enrichment can be achieved with this method.

Selenocysteine Incorporation in SPPS Sec-containing peptides can be synthesized by SPPS from protected Sec amino acids. The most commonly used protecting groups are phenyl, *p*-methoxybenzyl (Pmb), xanthyl and trityl groups.[241,242] Due to the reactive nature of selenols, cleavage from the resin is typically performed in the presence of thiols, most commonly 2,2′-dithiobis(5-nitropyridine) (DTNP), to form a hemiselenide bond as a side chain protection for selenol[243]. The hemiselenide can then be reduced *in situ* by treatment with dithiothreitol (DTT).

Selenoprotein Biosynthesis Sec is known as "the 21st amino acid" and can be co-translationally inserted through special cellular machinery[244]. Sec is encoded by the stop codon, UGA, which is repurposed for Sec using a set of special mRNA recognition

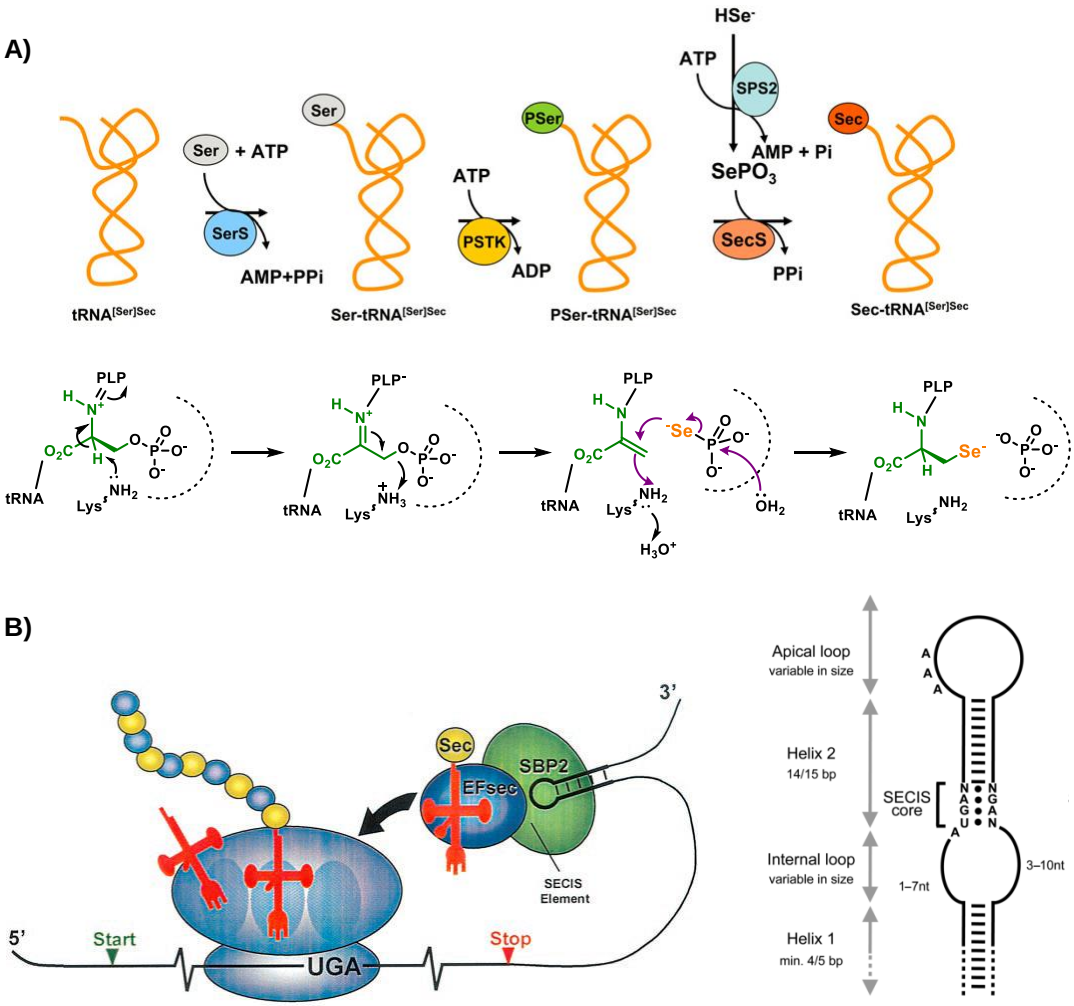


Figure 1-33. Translational Machinery for Selenocysteine Incoporation.

A) Biosynthesis of Sec-tRNA^{UCA}; B) Machinery for co-translational insertion of Sec and structure of SECIS element. SerS = seryl-tRNA synthetase; PSTK = *O*-phosphoseryl-tRNA kinase; PSer = phosphoserine; SecS = selenocysteine synthetase; SPS2 = selenophosphate synthetase 2; SECIS = selenocysteine insertion sequence; SBP2 = SECIS-binding protein 2; EFsec = special elongation factor. Graphics adapted from Hatfield *et al.*[245]

elements, elongation factors and $tRNA_{UCA}$ (Figure 1-33)[245]. The Sec-tRNA is synthesized from Ser-tRNA; the Ser is first phosphorylated, and then the phosphate is β-eliminated to form dehydroalanine, which is subsequently attacked by selenophosphate to generate selenocysteine on the tRNA[246]. The Sec-tRNA is then transported to the ribosome by a

special elongation factor EFsec[247]. The decoding of the UGA codon is achieved through a selenocysteine insertion sequence (SECIS) in the non-coding region near the 3'-end of peptide coding sequence; SECIS binding protein 2 (SBP2) binds to the SECIS element and then recruits the Sec-$tRNA_{UCA}$/EFsec complex to insert the Sec into the growing peptide chain[248]. The process is in competition with chain termination through release factor (RF) binding; therefore, the UGA codon and SECIS element must be in spatial proximity and able to bend into a certain geometry to allow the efficient binding of EFsec and SBP2 for the incorporation of Sec[249]. The human proteome contains 25 naturally existing selenoproteins, most of which are redox enzymes[250,251]; due to the reactive nature of the selenol side chain, all these proteins contain a Cys near the Sec residue, to sequester the Sec as a intramolecular hemiselenide.

Although recombinant expression of proteins with a Sec insertion has been achieved, it is at a very early stage of development. Gladyshev *et al.* were able to express the Sec mutant of a mammalian glutathione peroxidase in *E coli.* by grafting a UGA codon and a SECIS element into the protein-coding plasmid, but predominantly observed Trp mis-incorporation (the Trp codon is UGG, which is similarly to the UGA codon) instead of the desired Sec-containing protein[252]. A handful of other examples have been reported[253,254,255]; in all cases, Sec incorporation was limited to the C-terminal region of the target protein due to the spatial proximity requirement of the UGA codon and SECIS element. To our knowledge, there has not been any demonstration on Sec incorporation into proteins other than redox enzymes, which contain native Cys-Cys motifs that can be mutated into Sec-Cys.

1.7 Aminoacyl Transferase (AaT)

Structure and Biological Function of AaT AaT is a key component of the N-end degradation pathway in *E. coli*[256]. The N-end pathway is a cellular mechanism to regulate protein half-life through N-terminal "destabilizing" amino acids known as "degrons". In *E. coli*, the primary degron is an Arg or Lys (these residues are rarely exposed at the N-terminus of a well-folded protein). The positively charged Arg or Lys is recognized by AaT, which installs a secondary degron of Leu or Phe onto the N-terminus of the protein. The L/F-R/K motif is then recognized by a "carrier protein" ClpS, which transports the tagged protein to a proteolytic complex, ClpA/P, for degradation (Figure 1-34).

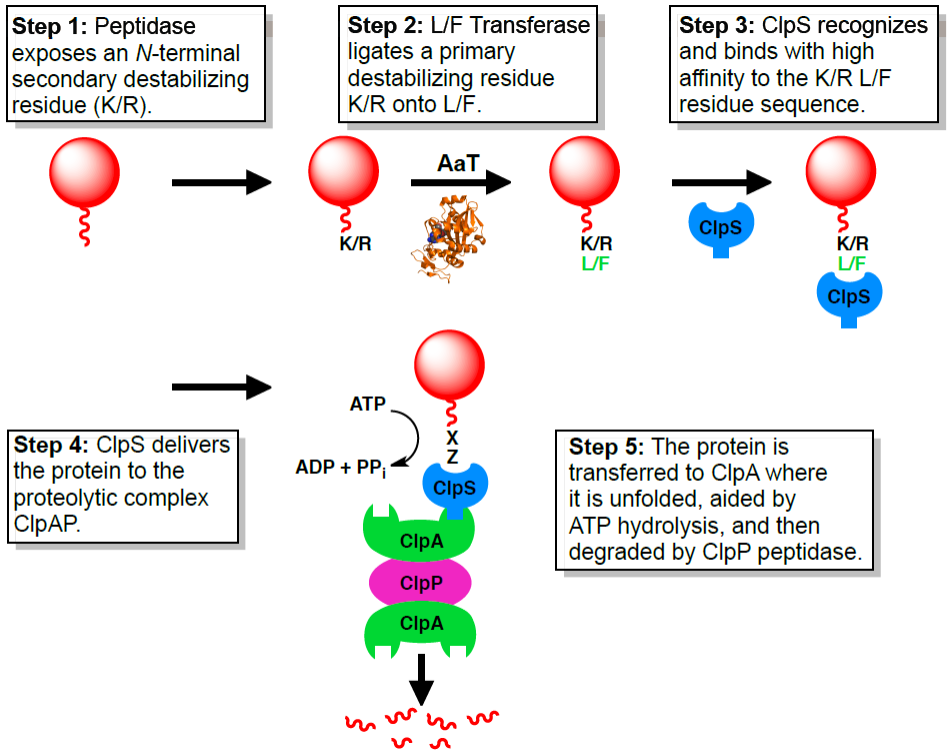


Figure 1-34. Aminoacyl Transferase (AaT) and *E. coli* N-End Degradation Pathway. Adapted from Bukau *et al.*[256] Picture courtesy of Anne M. Wagner.

The first observation of AaT activity was made in 1965 by Kaji *et al.*, where the incorporation of radiolabeled leucine and phenylalanine was observed in cell extracts of *E. coli*[257]. The AaT enzyme was later isolated in 1970 by Leibowitz and Soffer[258]. They also characterized the basic functional properties of AaT, including the equal preference for its natural substrates Leu and Phe, the requirement of a monovalent cation source such as KCl, and the optimal pH of 8. Subsequent research established that the substrates of AaT are protein N-termini nearing a Lys or Arg, and linked its biological function to the N-end degradation pathway[259,260]. A crystal structure was solved in 2007 by Suto *et al.*[261]; it revealed two distinct binding pockets – a hydrophobic amino acid binding pocket that consists of Met144, Met158, Leu170 and Ile185, and a negatively charged Lys/Arg binding pocket that is formed by Tyr42, Tyr120, Glu156 and Gln188 (Figure 1-35). The catalytic mechanism seems to be a simple "induced fit" model[262]; Gln188 was proposed to be catalytic in initial studies[261], but was later proved to be unnecessary for the catalytic process as an Ala mutation can be made at this site without loss of activity[262].

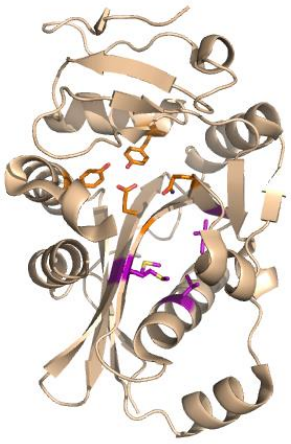


Figure 1-35. Crystal Structure of E. coli Aminoacyl Transferase (AaT).
Adatped from Suto *et al.* (PDB 2Z3N)[261]. Binding pockets highlighted in orange and purple.

AaT as a Protein N-Terminal Modification Tool In 1996, Abramochkin *et al.* determined that the aminoacyl-tRNA anticodon was not necessary for AaT recognition using mutant tRNAs[263]. It was later confirmed in the crystal structure that only the last adenosine on the acceptor stem was involved in AaT binding[261]. Pioneering studies by Sisido and Tirrell established the feasibility of using AaT for protein N-terminal modification. Tirrell *et al.* took a fully enzymatic approach, where mutant synthetases were used to charge their corresponding tRNAs with unnatural amino acids; the aminoacyl-tRNAs were then subjected to activity screening with AaT[264]. Sisido *et al.* showed that truncated nucleotides with from 2 to 20 bases were also tolerated as AaT substrates[265,266,267]. In our own group, we were able to further minimize the substrate to aminoacylated adenosine donors[268], and used this method to identify novel AaT substrates such as disulfide protected homocysteines (Figure 1-36)[269]. Known substrates of AaT from these studies are summarized in Figure 1-37.

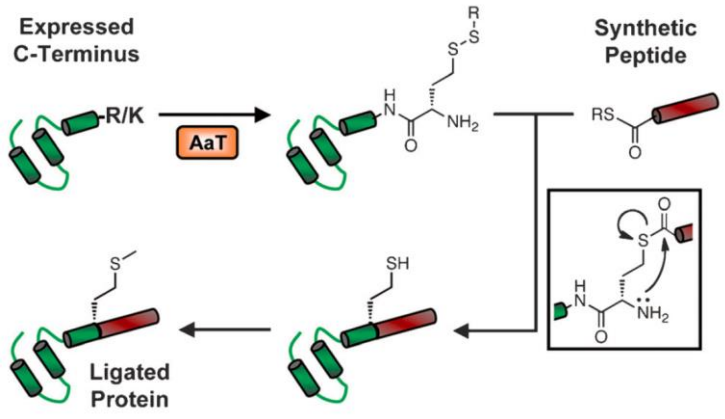
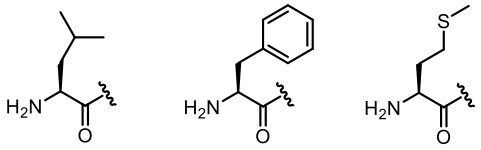
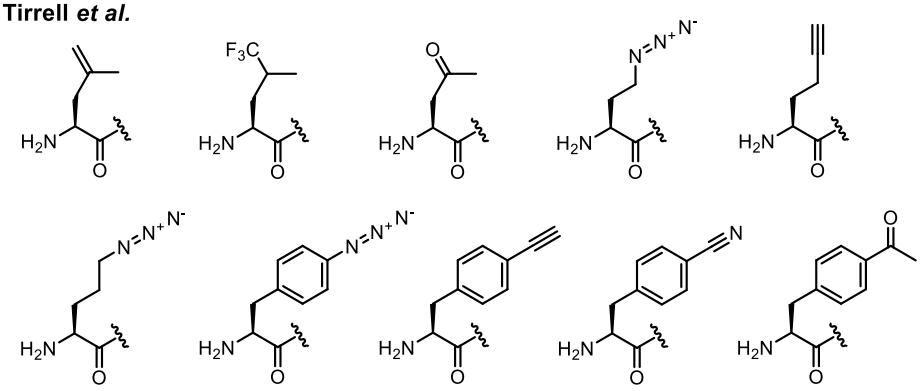


Figure 1-36. Aminoacyl Transferase as a Tool to Incorporate Homocysteine. Graphics adapted from Tanaka *et al.*[269]

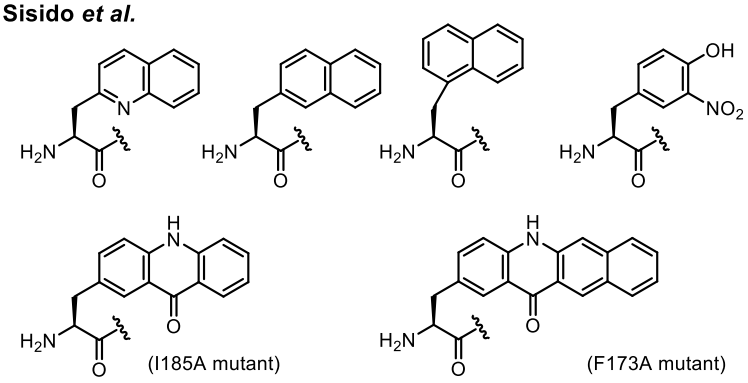
Natural Substrates



Tirrell *et al.*



Sisido *et al.*



Our Group

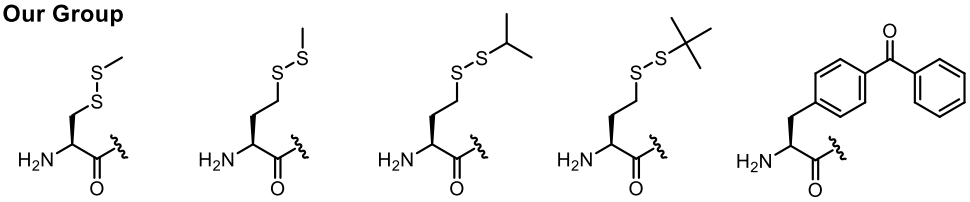


Figure 1-37. Known Substrates of Aminoacyl Transferase.

1.8 Summary

Protein misfolding is the molecular basis of various human diseases, including AD, PD and Type 2 diabetes. However, little information is known about either the structural details of the misfolded species or the conformational changes that are involved in the misfolding process, which greatly restricts our ability to develop effective therapies for these diseases. We propose to study the misfolding phenomenon using thioamides as minimalist fluorescence quenchers. Previous work in our group has established the photophysics of fluorescence quenching by thioamides. In the current work, we will explore methods to incorporate thioamides into full-length proteins, a prerequisite for using fluorophore/thioamide dually labeled proteins for misfolding studies.

Chapter 2 . Latent Peptide Thioester Strategies for Incorporation of Thioamides into Full-length Proteins via Native Chemical Ligation

The C^bPG_o linker and full-length αS studies in this chapter was originally published in the *Journal of the American Chemical Society*. It is adapted here with permission from the publisher:

2.1 Introduction

Having demonstrated the utility of the thioamide as a minimalist fluorescence quencher in previous work from our group[115], we were faced with a major chemistry challenge: as a backbone modification, thioamides can only be installed onto small molecules through solution phase thionation, or short peptides by Fmoc-based (Fmoc = fluorenylmethyloxycarbonyl) solid phase peptide synthesis (Figure 2-1). To the best of our knowledge, the longest thioamide-containing peptides previously synthesized were two 35-residue peptides, one by Miwa *et al.*[89] and one by our group[115]. In order to utilize thioamide as a probe for misfolding studies, we needed to first devise a method to incorporate it into full-length proteins reliably and site-specifically.

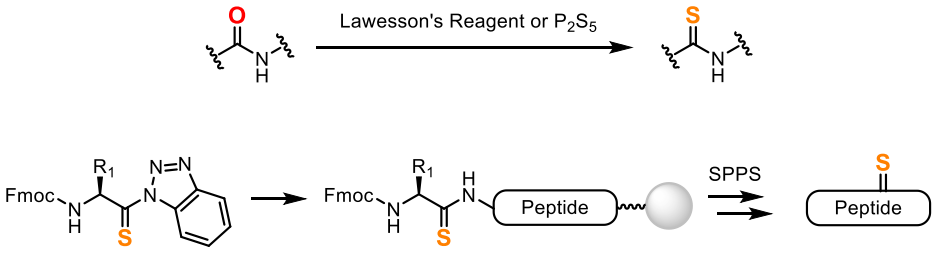


Figure 2-1. Thioamide Incorporation into Small Molecules and Peptides.

For small molecules and dipeptides, solution phase thionation can be used to convert an oxo-amide precrusor into thioamide; this would not be applicable to peptides with more than two residues due to the lack of site-specificity. Solid phase peptide synthesis (SPPS) using activated thioacyl benzotriazole precursors can be used for site-specific thioamide installation; however, SPPS has an inherent length limit of ~70 residues[270], beyond which cleavage and purification would become extremely difficult.

There are four general strategies to incorporate synthetic moieties into proteins: post-translational chemical derivatization[271], unnatural amino acid (Uaa) mutagenesis[272], *in*

vitro translation from chemically synthesized tRNA[273], and protein semi-synthesis[130]. The first two are primarily used for side chain derivatization, and would be difficult to adapt for backbone thioamide incorporation. With chemically synthesized tRNA, it is possible to incorporate a dipeptide substrate (a direct thioacid linkage to tRNA would be unstable); however, extensive synthetic efforts and directed evolution of the translational machinery are necessary for each Xaa-Yaa dipeptide combination[274], which is not desirable for our goal of thioamide incorporation at any position. Among available strategies, protein semi-synthesis, particularly native chemical ligation (NCL), is most promising for our applications – a small thioamide-containing peptide fragment can be chemically synthesized *via* SPPS, and then joined to the rest of the target protein produced by cellular expression, yielding a full-length target protein with a site-specific thioamide label.

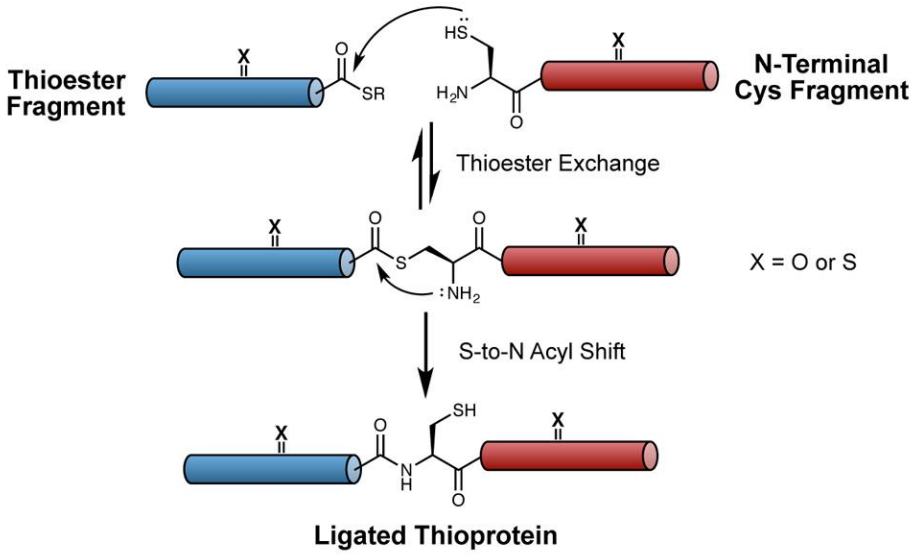


Figure 2-2. Native Chemical Ligation (NCL) as a Strategy for Thioamide Incorporation.

Native chemical ligation (NCL), as pioneered by Kent *et al.*[128], utilizes the trans-thioesterification between two unprotected peptide fragments − one with a C-terminal thioester and another with an N-terminal Cys – and generates a native amide bond after S-to-N acyl shift (Figure 2-2). The reaction takes place under mild aqueous conditions, and has been successfully applied to the semi-synthesis of chemically sensitive peptide derivatives such as glycopeptides[275]. In our preliminary trials, we were delighted to find that the thioamide was compatible with the NCL method; the real challenge, however, was to synthesize the peptide fragments, particularly the thioester fragment, for ligation.

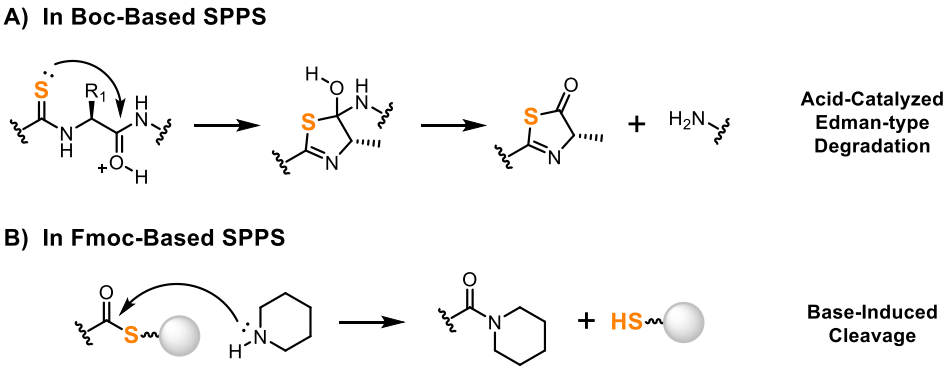


Figure 2-3. Challenges in the Synthesis of Thioamide-Containing Peptide Thioesters.

In standard SPPS methods, thioesters are synthesized using Boc-based SPPS (Boc = *t*-butyloxycarbonyl), as they are not stable towards the basic conditions in Fmoc-based SPPS; on the contrary, the thioamide is not stable towards the highly acidic conditions in Boc-based SPPS, and requires the Fmoc method for successful incorporation (Figure 2-3). In fact, this apparent mismatch has led to the speculation that NCL of thioamide-containing peptides was "not suitable because the presence of [thioamide] bonds is not

compatible with the subsequent synthesis of the thioester moiety".[110] We hypothesized that we could find a compromise by adopting Fmoc-based SPPS for the synthesis of thioamide-containing peptides, and then identify a viable method to either generate the thioester off-resin or synthesize it on-resin as a latent thioester that is activated *in situ*.

Upon a survey of the literature, we identified three main strategies for thioester synthesis in the context of Fmoc-based SPPS (Figure 2-4). Solution phase PyBOP activation[130] is performed on the free carboxylic acid terminus of a peptide; due to the harsh reaction conditions, it requires side chains of the peptide to be fully protected, which greatly affects its solubility and thus lowers reaction efficiency. Nbz and other "safety-catch" type linkers[276,277] utilize non-labile amide linkages to anchor the growing peptide to the resin; the linkers are activated at the end of the synthesis (typically by *N*-acylation) and converted to thioesters by thiolysis. Latent thioester linkers[278,171] are synthesized and purified as ester or amide using standard SPPS; they are only activated *in situ* through O-to-S or N-to-S acyl shift to generate an active thioester in aqueous buffer.

We chose to investigate the latent thioester strategy for thioamide incorporation because of its various advantages: 1) while the other methods require additional derivatization after thioamide installation, all harsh reactions in the latent thioester strategy are completed at the linker synthesis stage, ensuring maximal compatibility with chemically sensitive functional groups such as thioamide; 2) from a practical point of view, latent thioesters are purified and stored as esters or amides until the final NCL, which are much more stable and easier to handle than an actual thioester.

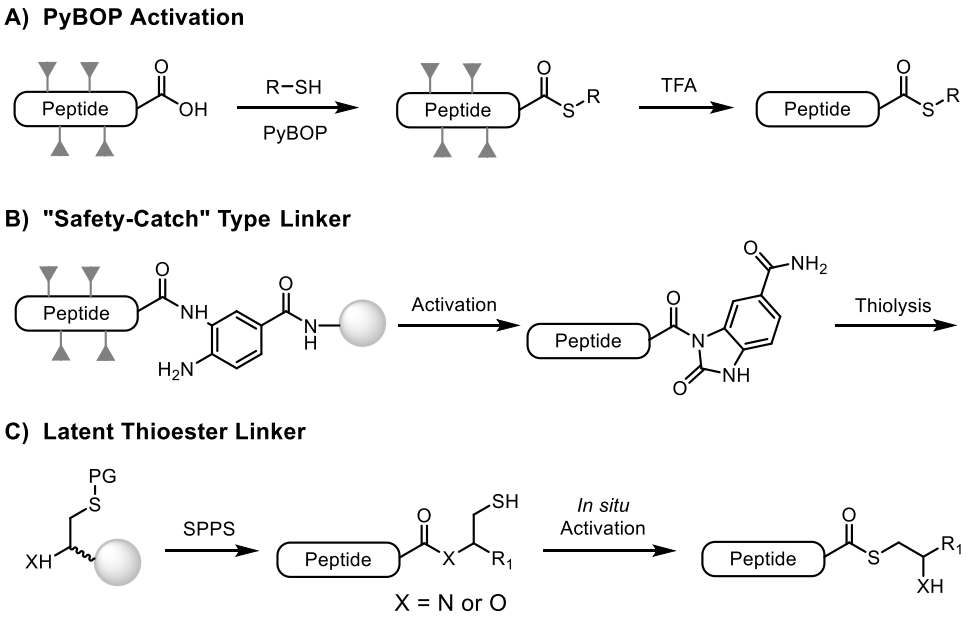


Figure 2-4. Thioesterification Strategies for Fmoc-based SPPS.

Due to the diversity of "safety-catch" type linkers, Nbz linker is shown here as an example. Triangles indicate side chain protecting groups on peptides. TFA = trifluoroacetic acid; PG = protecting group SPPS = solid phase peptide synthesis; Nbz = N-acyl-benzimidazolinone.

The two main sub-types of latent thioester linkers are O-to-S and N-to-S linkers. As the names suggest, O-to-S linkers are joined to the peptide through an ester bond and are subsequently activated by O-to-S acyl shift, while N-to-S linkers are the amide bond counterparts. Rather than choosing one type over the other, we decided to explore both strategies because they have complementary advantages: 1) due to the higher stability of amide bond as compared to ester bond, N-to-S linkers are more stable than O-to-S linkers in synthesis and handling, but would also be slower in *in situ* thioester formation; 2) for the same reason, O-to-S acyl shifts are typically conducted in mildly acidic conditions (pH 6.8 ~ 7.0) to avoid ester hydrolysis, whereas N-to-S acyl shifts can be conducted at much higher pH (7.4 ~ 8.4) to accelerate the reaction. Exploring both types gave us

access to a wide range of reactivity and reactions conditions, which would be beneficial in fine-tuning synthetic procedures for different target proteins.

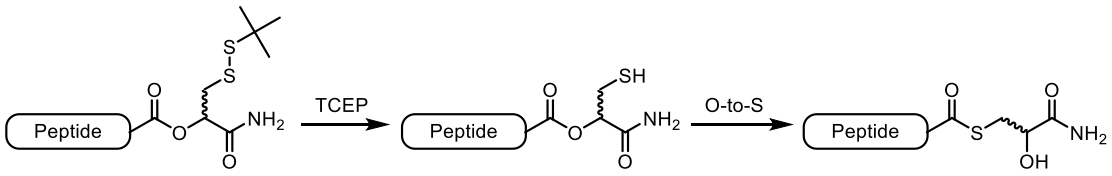


Figure 2-5. ChB Latent Thioester Linker utilizing O-to-S Acyl Shift.

The O-to-S linker we chose to adapt for the thioamide is ChB, an α-hydroxyl analog of Cys, as developed by Botti *et al.*[278] (Figure 2-5). It was initially synthesized by on-resin hydroxylation of Cys, and subsequently refined by Muir *et al.* into a monomer building block that can be directly incorporated by SPPS[186]. The N-to-S linker we chose was CPG_o, a Cys-Pro-Gla motif, as developed by Aimoto *et al.*[171] (Figure 2-6). Pro is a well-known substrate for diketopiperazine (DKP) formation; glycolic acid promotes this cyclization by serving as a good leaving group, which in turn shifts the N-to-S acyl shift equilibrium towards the thioester. In our design of C^bPG_o, we would added a *t*BuS− protecting group on Cys to avoid complications from dimerization and aggregation.

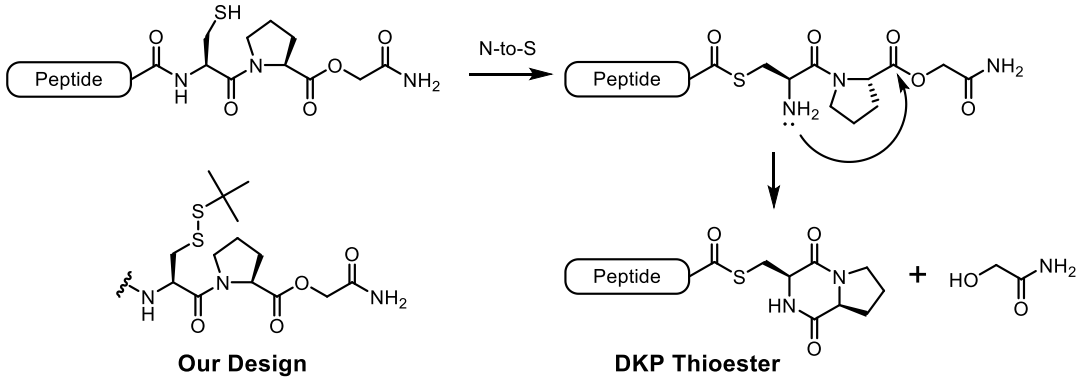


Figure 2-6. CPG_o Latent Thioester Linker utilizing N-to-S Acyl Shift.

2.2 Results and Discussion

We began by exploring the O-to-S latent thioester linker ChB. It can be synthesized in its protected form TBS-ChB-OH (**6)** in five steps from 1-thioglycerol (Figure 2-7) based on a route developed by Muir *et al.*[186], where successive oxidations were used to obtain the carboxylic acid functionality. (We note that all synthetic routes from Cys or cystine proved unfruitful as summarized in Figure 2-15.) The precursor can then be used as a building block for incorporation by SPPS – **6** is first coupled onto Rink amide resin, and deprotected to reveal the hydroxyl group; the next amino acid is installed using *in situ* carbodiimide activation, after which the rest of the sequence can be assembled using standard SPPS (Figure 2-8).

To test the compatibility of the ChB thioester linker with a thioamide, we synthesized model peptide **7** with thioleucine (Leu^S) in the sequence. As expected, the latent ester strategy was a sufficiently small deviation from standard SPPS that **7** could be readily synthesized and purified without complications. When subjected to a test ligation with N-terminal Cys peptide **8**, the latent thioester was rapidly converted into the desired product **9c** within 10 min (Figure 2-9). Intermediates **9a** and **9b** could be observed early in the reaction, which corresponded to the O-linked and S-linked isomers as latent or active thioesters. The UV-Vis absorption profile of **9c** clearly demonstrated the presence of two 7-methoxycoumarin chromophores and the integrity of thioamide, giving us confidence that the ChB linker is compatible with the synthesis of thioamide-containing thioesters.

We next investigated the N-to-S latent thioester linker C^bPG_o. The synthesis was much more straight-forward, where the only precursor necessary was an Fmoc-Xaa-Csb-OH dipeptide (**10)** to pre-load the C-terminal residue in order to avoid spontaneous diketo-piperazine (DKP) ester formation[171] (Figure 2-18). **10** can be readily derivatized from Fmoc-protected amino acids in one step by isobutylchloroformate (IBCF)

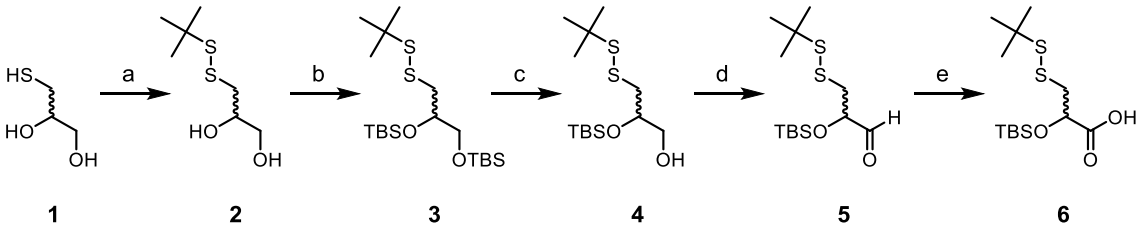


Figure 2-7. Synthesis of TBS-ChB-OH Precursor.

Conditions: a) *t*BuSH, I_2, EtOH, 89%; b) TBS-Cl, imidazole, DMAP, DMF, 99%; c) 1:1:5 TFA/H_2O/THF, 0 ˚C, 61%; d) Dess-Martin periodinane, CH_2Cl_2, 65%; e) $NaClO_2$, NaH_2PO_4, 2-methyl-2-butene, 4:9 H_2O/tBuOH, 31%. TBS = *t*-butyldimethylsilyl; DMAP = 4-dimethylamino-pyridine; DMF = *N*,*N*-dimethylformamide; TFA = trifluoroacetic acid; THF = tetrahydrofuran.

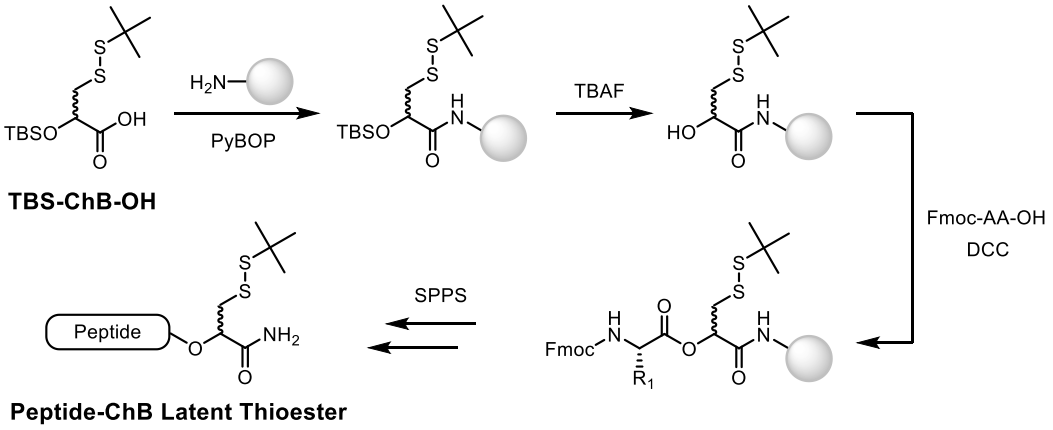


Figure 2-8. Synthesis Scheme for Thioamide-Containing Peptide-ChB Thioester.

We note that the symmetric anhydride intermediate from carbodiimide acitivation method was necessary for successful ester bond; uronium reagents such as HBTU were much less effective. TBAF = tetra-*n*-butylammonium fluoride; Fmoc-AA-OH = Fmoc protected amino acid building block; DCC = dicyclohexylcarbodiimide.

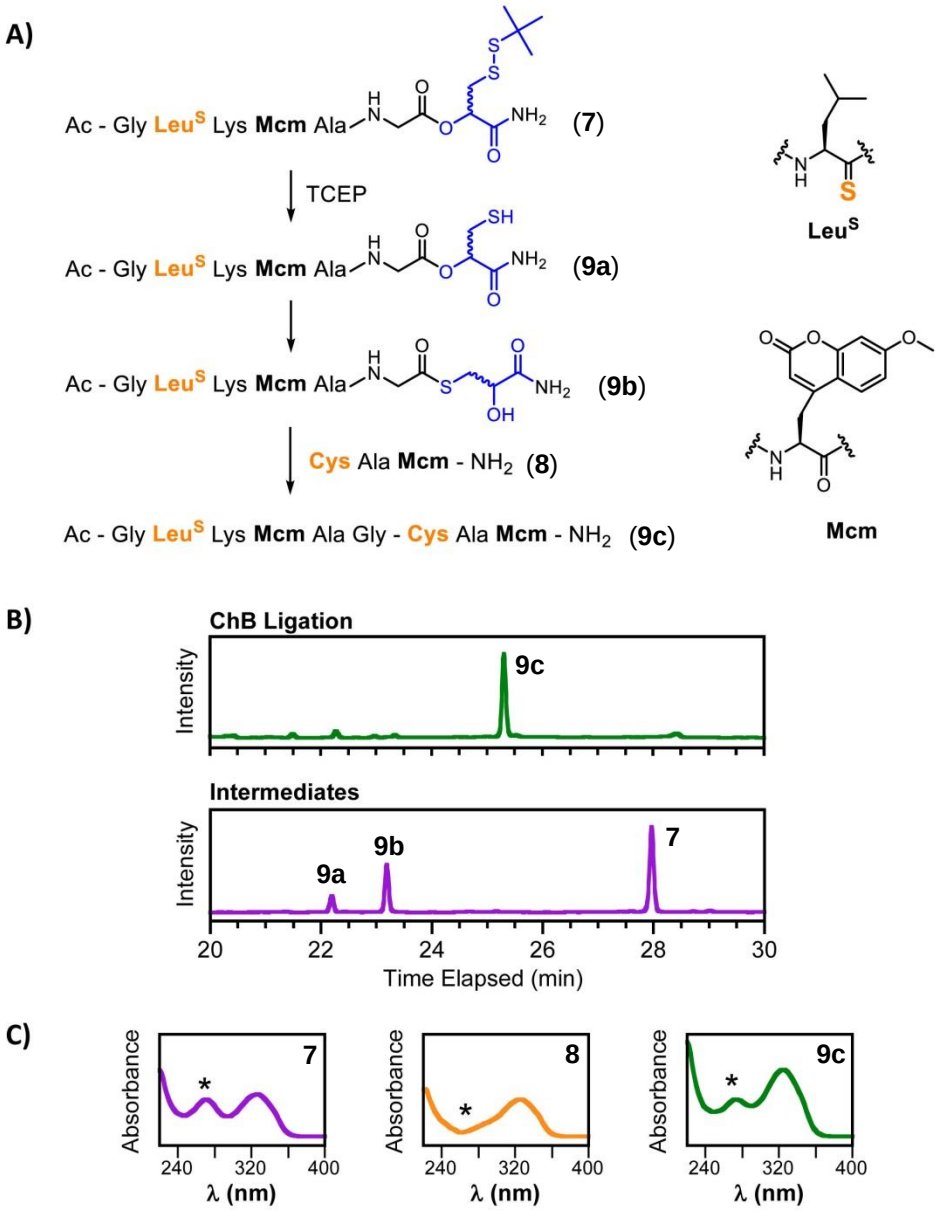


Figure 2-9. Ligation of Thioamide-Containing Peptide-ChB Thioester with CA-Mcm-NH_2.

A) Schematic representation of the test reaction. B) HPLC chromatograms monitored at 325 nm. Ligation was conducted using 1:10 Peptide-ChB **7**/Cys-Peptide **8** for a 10 min reaction; intermediates were isolated using 1:1 Peptide-ChB **7**/Cys-Peptide **8** for a 30 sec reaction. Excess **8** eluted at 18.5 min; it is omitted in these graphics for clarity. MALDI MS: $[\mathbf{7}+H]^+$, expected 939.37, found 939.23; $[\mathbf{9a}+H]^+$, expected 851.34, found 851.22; $[\mathbf{9b}+H]^+$, expected 851.34, found 851.21, $[\mathbf{9c}+H]^+$, expected 1166.46, found 1166.30. **9a** was positive towards Ellman's reagent, while **9b** was negative. C) UV-Vis absorption profiles of reactants and product, showing the integrity of thioamide and two Mcm chromophores in the product. Conditions: 0.5 mM **7**, 0.5 or 5 mM **8**, 20 mM TCEP, 200 mM Na_2HPO_4, pH 6.8.

activation, a common amide bond formation method that is compatible with all 20 natural amino acids[124]. To construct the C^bPG_o linker, we devised an efficient synthesis from bromoacetic acid (Figure 2-10) – after its coupling onto the resin, an S_N2 reaction was used to form the Pro-Gla ester bond, and then **10** was installed by PyBOP; standard SPPS was then used to complete subsequent synthesis and purification.

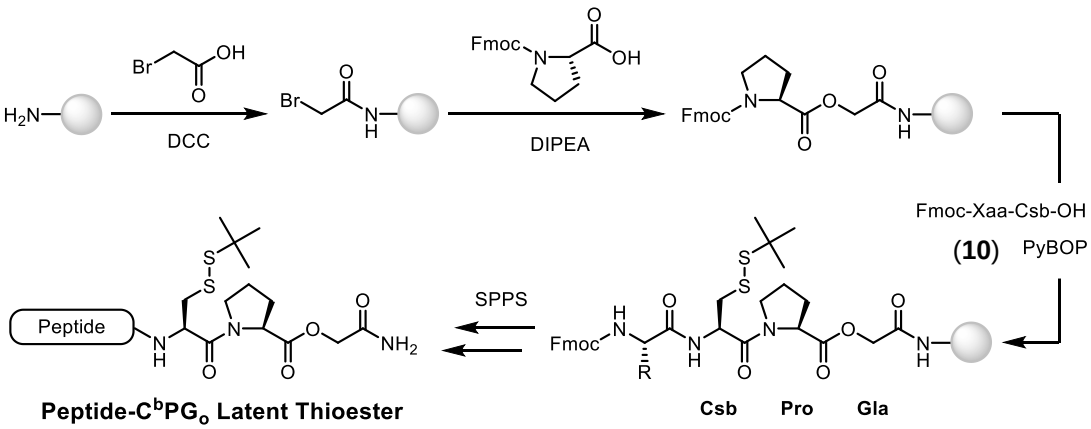


Figure 2-10. Synthesis Scheme for Thioamide-Containing Peptide-C^bPG_o Thioester.

We note that $BrCH_2COOH$ loading must be conducted at neutral to acidic conditions without catalytic DMAP to avoid glycolic acid polymer formation. Synthesis of Fmoc-Xaa-Csb-OH **10** is detailed in *Materials and Methods*. DCC = dicyclohexylcarbodiimide; DMAP = 4-dimethyl-aminopyridine; DIPEA = *N*,*N*-diisopropylethylamine.

Model peptide **11**, analogous to peptide **7**, was synthesized and evaluated in a test ligation with Cys (Figure 2-11). Upon treatment with reducing agent tris(2-carboxyethyl) phosphine (TCEP), **11** was rapidly reduced to the Cys form **11a** in 10 min; the subsequent N-to-S acyl shift took place gradually and approached completion after 6 h. The reaction rate was much slower than that of the latent ChB thioester, presumably due to the higher stability of the amide bond in the N-to-S acyl shift than the ester bond in the

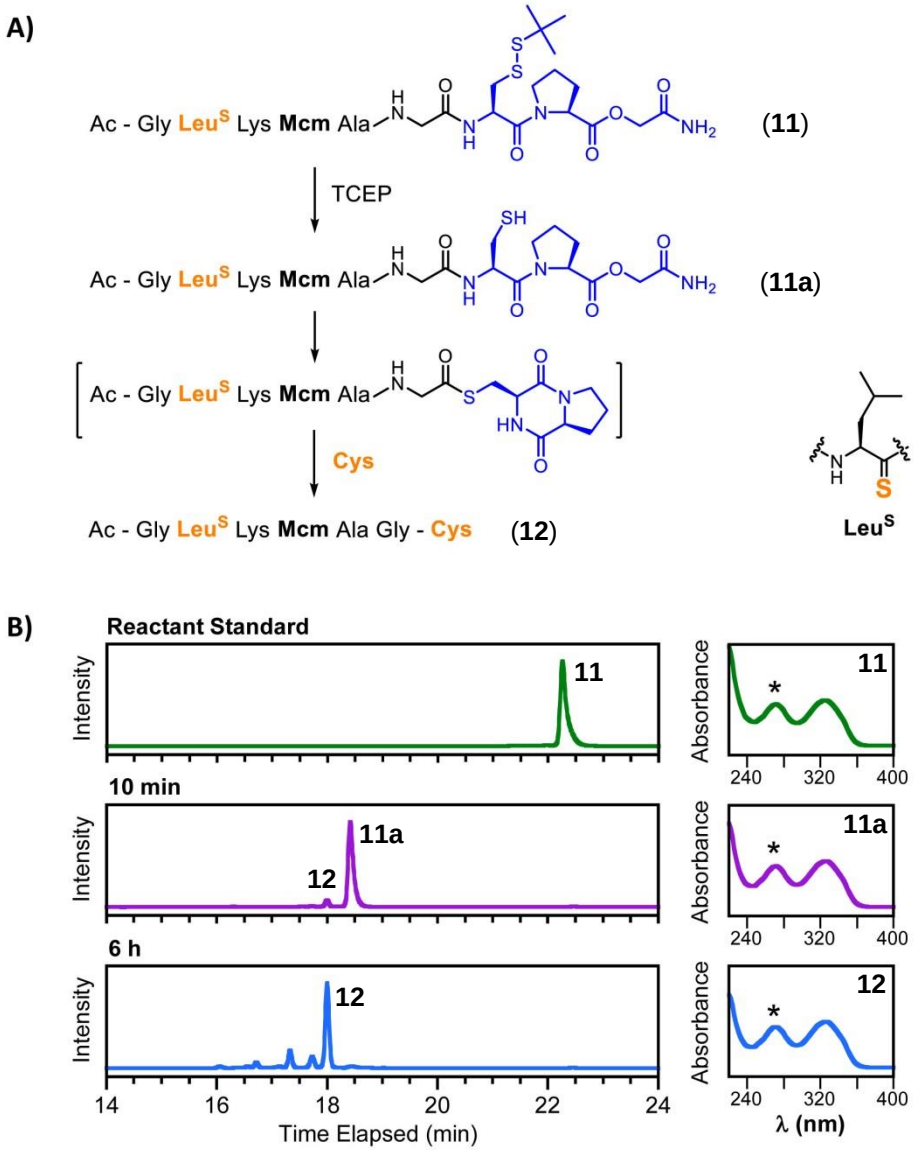


Figure 2-11. Ligation Thioamide-Containing Peptide-C^{b}PG$_{o}$ Thioester with Cys.

A) Schematic representation of the test reaction. B) HPLC chromatogram monitored at 325 nm, and UV-Vis absorption spectra of isolated peptides. MALDI MS: [**11** + H]$^{+}$, expected 1093.34, found 1093.46; [**11a** + H]$^{+}$, expected 1005.41, found 1005.37; [**12** + H]$^{+}$, expected 851.34, found 851.16. Star sign indicates thioamide absorption at 272 nm. See *Materials and Methods* for DKP intermediate characterization. Conditions: 1 mM peptide-C^{b}PG$_{o}$ **11**, 10 mM Cys, 50 mM TCEP, 200 mM Na_2HPO_4, pH 8.4.

O-to-S acyl shift. Using a large excess of Cys, we were able to achieve a pseudo first-order reaction with respect to **11**, from which a second-order reaction constant (1.53 ± 0.06) × 10^{-2} $M^{-1}{\cdot}s^{-1}$ was obtained,with an empirical half-life of 75 ± 1 min (Figure 2-12). As a further validation, we showed that the thioamide was fully intact in the final product, and that DKP thioester intermediate could be observed at a very low substrate concentration in the absence of Cys (Figure 2-19). Lastly, we showed that both ChB and C^bPG_o latent thioesters were tolerant of denaturants – ligation rates were slower, but otherwise no additional complications were observed (Figure 2-17, Figure 2-20).

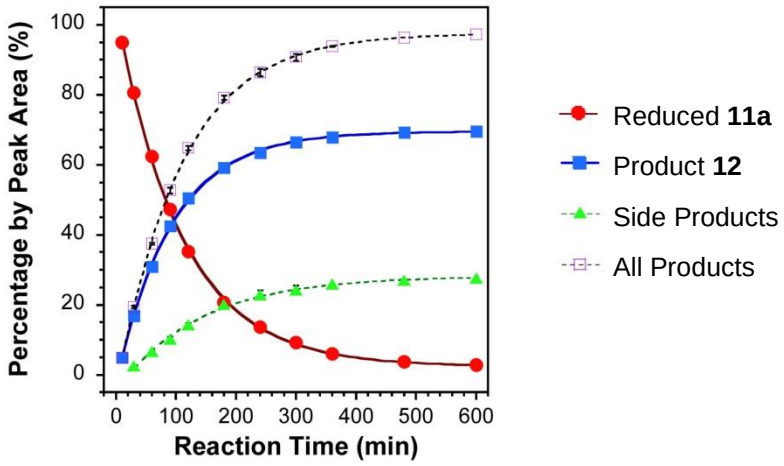


Figure 2-12. Triplicated Ligation Kinetics between C^bPG_o Thioester and Cys.

Side products were defined as all products other than the desired product **12**. See *Material and Methods* for detailed descriptions of quantification and data fitting methods.

Having validated both ChB and C^bPG_o latent thioester strategies in the context of thioamide incorporation, we were eager to find out if these methods would be applicable to a full-length protein. Our chosen target was α-synuclein (αS), a 140 residue protein that is implicated in Parkinson's disease (PD) [279]. αS is a small pre-synaptic protein that

is ubiquitously expressed in the brain; upon conformational changes, it aggregates into oligomers and fibrils, leading to neuronal death and thus the degenerative symptoms in PD. By incorporating a Trp/thioamide fluorophore-quencher pair into αS, we would be able to track the relative movements of the two labels during aggregation using fluorescence spectroscopy, and then reconstruct the conformational changes involved.

We chose the C^bPG_o linker for this particular application because of its ease of synthesis. Using an improved procedure, we synthesized Ac-$\alpha S_{1\text{-}8}V^S{}_3$-$C^bPG_o$ **13** and ligated it to the other fragment $\alpha S_{9\text{-}140}C_9$ that was expressed in *E. coli*. The ligation proceeded smoothly, yielding full-length Ac-$\alpha SV^S{}_3C_9$ as the desired product (Figure 2-13). The integrity of thioamide can be confirmed by trypsin digest, as well as UV-Vis absorption (Figure 2-23). To generate the Trp/thioamide dually labeled protein, αSC_9W_{94} was use as the C-terminal fragment instead, which also proceeded without complications to yield Ac-$\alpha SV^S{}_3C_9W_{94}$ (Figure 2-22).

As a proof-of-concept experiment to demonstrate the utility of thioamide in protein misfolding studies, we subjected our semi-synthetic doubly labeled Ac-$\alpha SV^S{}_3C_9W_{94}$ to *in vitro* aggregation and monitored the Trp fluorescence change over time (Figure 2-14). There is no native Trp in the αS sequence, therefore selective excitation of Trp can be achieved with 295 nm incident light. To account for the environmental effects on Trp emission, all fluorescence data were compared to an oxoamide control, where αSW_{94} with no thioamide in its sequence was aggregated side by side with the thioamide-labeled

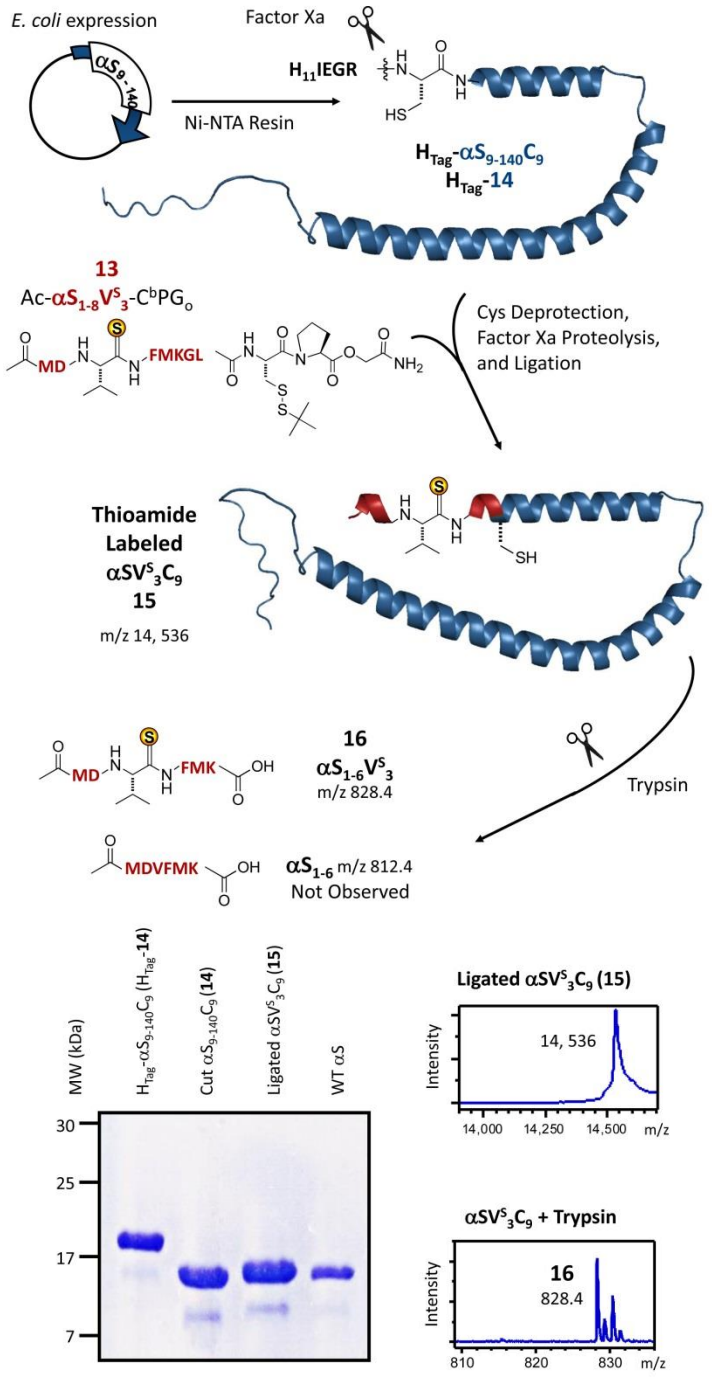


Figure 2-13. Synthesis of Thioamide-Containing Full-length αS using C^bPG_o Linker.

C-Terminal fragment of $\alpha S_{9\text{-}140}$ with an N-terminal His tag, an S9C mutation, and IEGR cleavage sequence was expressed and purified on Ni-NTA resin. After Factor Xa cleavage to generate an N-terminal Cys ($\alpha S_{9\text{-}140}C_9$), the protein fragment is ligated Ac-$\alpha S_{1\text{-}8}V^S_3$-$C^bPG_o$ latent thioester. Top Right: PAGE gel analysis of ligation. Far Right: MALDI-TOF MS of full length (calcd m/z 14,536) or trypsinized Ac-$\alpha SV^S_3C_9$ (Ac-$\alpha S_{1\text{-}5}V^S_3$, calcd m/z 828.4). No desulfurized oxopeptide (calcd m/z 812.4) was observed. With Solongo Batjargal.

samples. To isolate intramolecular misfolding events, we carried out our aggregation using 1:30 doubly labeled Ac-αSV$^{S}_{3}$C$_{9}$W$_{94}$ in WT αS, so that statistically there should not be intermolecular Trp/thioamide interactions within 30 Å (the contact quenching distance for Trp/thioamide pair[280]). Finally, with a residual Cys in our semi-synthetic construct, we conducted our aggregation in the presence of β-mercaptoethanol (BME) to prevent disulfide bond formation that may interfere with the aggregation process.

We were excited to find that as aggregation proceeded, Trp fluorescence quenching was clearly observed, which corresponded to compaction between the Trp and thioamide labels as one would expect in oligomer and fibril formation. Comparing to aggregation progress as monitored by thioflavin T (ThT) fluorescence, Trp fluorescence quenching took place before the rise of ThT fluorescence (ThT preferentially binds to fibrils[54]), indicating that our Trp/thioamide dual labeling method revealed conformational changes at the early oligomer formation stage (< 24 h), which had been a difficult target in prior efforts[281]. While an extensive data set from various labeling positions is certainly necessary to draw definitive conclusions about the exact conformations involved, these studies clearly demonstrated the feasibility of thioamide incorporation into full-length proteins *via* NCL, as well as their great potential as minimalist probes in studying the misfolding of pathogenic proteins.

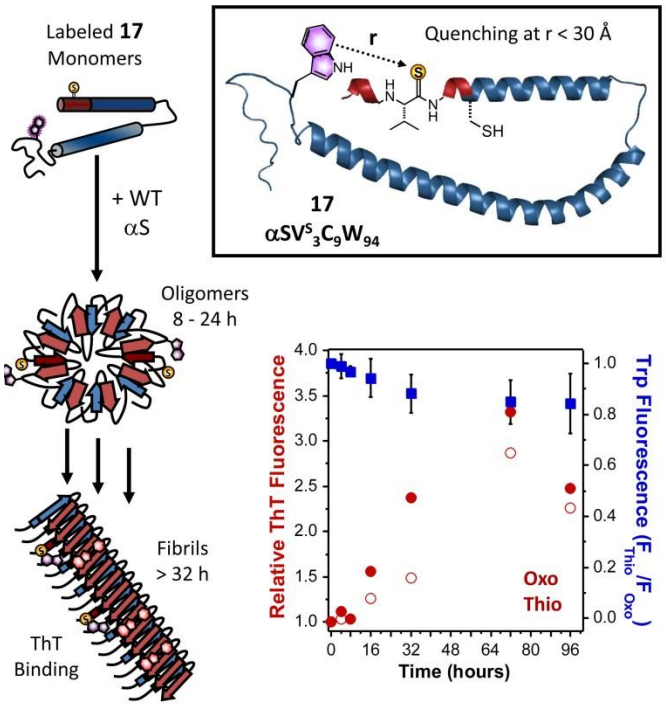


Figure 2-14. Monitoring αS Misfolding Using Thioamide Fluorescence Quenching.

Monomeric Ac-αS$_{1\text{-}140}$V$^{S}_{3}$C$_{9}$W$_{94}$ was mixed in a 1:30 ratio with WT αS, and shaken at 37 °C for several days. Samples were taken at defined intervals and Trp fluorescence were measured at 295 nm. Thioflavin T (ThT) was used as a secondary assay to confirm fibril formation. See *Materials and Methods* for detailed description. With Rebecca F. Wissner.

2.3 Conclusion

We have established two efficient thioesterification methods for the incorporation of thioamide into full-length proteins by NCL. ChB, an O-to-S latent thioester linker, functions at mildly acidic pH and exhibits rapid reaction kinetics, while C^bPG_o, an N-to-S latent thioester, utilizes mildly basic reaction conditions and is more stable toward purification and storage. Both methods are fully compatible with thioamide, giving a wide range of conditions for protein-specific optimization.

As a proof-of-concept example, we successfully extended the thioesterification and NCL method to synthesizing a thioamide-labeled variant of αS, a protein implicated in Parkinson's disease. We demonstrated the utility of the thioamide as a minimalist fluorescence quencher probe in an aggregation study with a Trp/thioamide dually labeled αS construct; conformational compaction was observed at the oligomer stage determined by Trp fluorescence quenching by the thioamide. Our semi-synthesis method has enabled the site-specific incorporation of thioamide into full-length proteins, greatly expanding its potential applications as a minimalist probe in protein misfolding studies.

2.4 Materials and Methods

General Information Fmoc protected amino acids and peptide synthesis reagents were purchased from EMD Millipore (Billerica, MA). All other reagents and solvents were purchased from Fisher Scientific (Waltham, MA) or Sigma-Aldrich (St. Louis, MO) unless otherwise specified. High resolution electrospray ionization mass spectra (ESI-HRMS) were collected with a Waters LCT Premier XE liquid chromatograph/mass spectrometer (Milford, MA). Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer. UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (Agilent Technologies, Santa Clara, CA). Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA). Matrix assisted laser desorption/ionization with time-of-flight detector (MALDI-TOF) mass spectra were acquired on a Bruker Ultraflex III instrument (Billerica, MA). Analytical HPLC was performed on an Agilent 1100 Series HPLC system equipped with a photodiode array detector (Santa Clara, CA). Preparative HPLC was performed on a Varian Prostar HPLC (Agilent Technologies, Santa Clara, CA). HPLC columns were purchased from W. R. Grace & Compnay (Columbia, MD). Fluorescence spectra were collected with a Varian Cary Eclipse fluorescence spectrophotometer with a Peltier multicell holder (Agilent Technologies, Santa Clara, CA). Protein purification was conducted on a BioCad Sprint fast protein liquid chromatography (FPLC) system (Agilent Technologies, Santa Clara, CA).

Synthesis of 3-(*t*-Butyldisulfanyl)propane-1,2-diol (**2**) 3-Mercaptopropane-1,2-diol (**1**, 960 µL, 90% aq., 10 mmol, 1 equiv) was mixed with 2-methylpropane-2-thiol (tBuSH, 3.4 mL, 30 mmol, 3 equiv), and chilled to 0 °C on ice. A solution of iodine (2.7918 g, 11 mmol, 1.1 equiv) in ethanol (50 mL) was added dropwise over 2 h, until the reaction changed from colorless to red. The reaction was allowed to proceed for another 2 h on ice, and then quenched with saturated NaS_2O_3 (1 mL). The crude mixture was concentrated by rotary evaporation, and then purified by flash chromatography in 2:3 ethyl acetate/petroleum ether. The product was isolated as a yellow oil (1.7391 g, 8.9 mmol, 89% yield). R_f 0.23 in 2:3 ethyl acetate/petroleum ether; ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.78 (m, 1H), 3.52 (dd, *J* = 11.3, 4.7 Hz, 1H), 3.48 (dd, *J* = 11.3, 5.7 Hz, 1H), 2.85 (dd, *J* = 13.2, 5.4 Hz, 1H), 2.73 (dd, *J* = 13.4, 7.4 Hz, 1 H), 1.29 (s, 9H). ESI$^+$-LRMS calculated for $C_6H_{13}O_2S_2^+$ 181.04, found $[M - CH_3]^+$ 181.10.

Synthesis of 1,2-Bis-(*t*-butyl-dimethyl-silanyloxy)-3-*t*-butyldisulfanyl-propane (**3**) **2** (1.0144 g, 5.2 mmol, 1 equiv), imidazole (2.1105 g, 31 mmol, 6 equiv) and 4-dimethyl-aminopyridine (DMAP, 0.0379 g, 0.3 mmol, 0.06 equiv) were dissolved in anhydrous *N*,*N*-dimethylformamide (DMF, 15 mL) on ice under argon. *tert*-Butyl-dimethylsilyl chloride (TBSCl, 2.3359 g, 16 mmol, 3 equiv) was added, and the reaction was allowed to proceed overnight (≥ 12 hr). Upon completion, the reaction was quenched with 0.5 N NaOH (10 mL), and extracted with ethyl acetate (30 mL × 3). The organic layers were concentrated by rotary evaporation and purified by flash chromatography in 1:19 ethyl acetate/hexanes. The product was isolated as a clear oil (2.1758 g, 5.1 mmol, 99% yield). R_f 0.74 in 1:19 ethyl acetate/hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.87 (m, 1H), 3.58

(dd, *J* = 10.0, 5.1 Hz, 1H), 3.51 (dd, *J* = 10.1, 6.3 Hz, 1H), 2.97 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.71 (dd, *J* = 13.1, 6.7 Hz, 1H), 1.31 (s, 9H), 0.87 (s, 18H), 0.04 (m, 12H); ESI^+-LRMS calculated for $C_{19}H_{44}O_2S_2Si_2Na^+$ 447.22, found $[M + Na]^+$ 447.27.

Synthesis of 2-((*t*-Butyldimethylsilyl)oxy)-3-(*t*-butyldisulfanyl)propan-1-ol (**4**) **3** (0.8676 g, 2 mmol, 1 equiv) was dissolved in tetrahydrofuran (THF, 9 mL) on ice, and then a solution of trifluoroacetic acid (TFA, 1.8 mL, 24 mmol, 12 equiv) in water (1.8 mL) was added dropwise. The reaction was allowed to proceed for 3 h on ice until all reactant was consumed, and then quenched with saturated $NaHCO_3$ (30 mL) and extracted with ethyl acetate (30 mL × 3). The organic layers were concentrated by rotary evaporation and purified by flash chromatography in 1:15 ethyl acetate/hexanes. The product was isolated as a clear oil (0.3889 g, 1.3 mmol, 61% yield). R_f 0.25 in 1:15 ethyl acetate /hexanes; ^1H NMR (500 MHz, $CDCl_3$): δ 3.95 (m, 1H), 3.68 (dd, *J* = 11.3, 3.8 Hz, 1H), 3.60 (dd, *J* = 11.3, 4.2 Hz, 1H), 2.85 (dd, *J* = 13.4, 7.2 Hz, 1H), 2.77 (dd, *J* = 13.4, 5.7 Hz, 1H), 1.85 (s, 1H), 1.31 (s, 9H), 0.88 (s, 9H), 0.11 (s, 3H), 0.09 (s, 3H). ESI^+-LRMS m/z calculated for $C_{13}H_{30}O_2S_2SiNa^+$ $[M + Na]^+$ 333.14, found 333.10.

Synthesis of 2-((*t*-butyldimethylsilyl)oxy)-3-(*t*-butyldisulfanyl)propanal (**5**) Dess-Martin periodane was freshly prepared according to Ireland *et al.*[282] **4** (1.8736 g, 6 mmol, 1 equiv) was dissolved in CH_2Cl_2 (40 mL) under argon. Dess-Martin periodane (3.8379 g, 9 mmol, 1.5 equiv) was slowly added, and then the reaction was allowed to proceed for 3 h at room temperature. Upon completion, the reaction mixture was diluted with CH_2Cl_2 (50 mL) and ethyl acetate (50 mL), filtered, concentrated by rotary evaporation, and then

purified by flash chromatography in 1:19 ethyl acetate/hexanes. The product was isolated as a pale yellow oil (1.2183 g, 4.0 mmol, 65% yield). R_f 0.27 in 1:19 ethyl acetate/ hexanes; ^{1}H NMR (500 mHz, $CDCl_3$): δ 9.65 (d, *J* = 1.1 Hz, 1H), 4.23 (ddd, *J* = 7.9, 4,4, 1,1 Hz, 1H), 2.99 (dd, *J* = 13.3, 4.4 Hz, 1H), 2.81 (dd, *J* = 13.4, 7.9 Hz, 1H), 1.32 (s, 9H), 0.92 (s, 9H), 0.13 (s, 3H), 0.11 (s, 3H). ESI$^-$-LRMS calculated for $C_{13}H_{27}O_2S_2Si^-$ 307.13, found $[M - H]^+$ 307.21.

Synthesis of 2-((*t*-butyldimethylsilyl)oxy)-3-(*t*-butyldisulfanyl)propanoic acid (**6**) **5** (0.2479 g, 0.8 mmol, 1 equiv) was dissolved in 2-methyl-2-propanol (tBuOH, 4 mL) on ice. 2-Methyl-2-butene (4.8 mL, 45 mmol, 58 equiv) was added, followed by a solution of $NaClO_2$ (0.2110 g, 2.3 mmol, 3 equiv) and NaH_2PO_4 (0.2053 g, 1.7 mmol, 2.2 equiv) in water (4 mL). The reaction was vigorously stirred for 2 h under nitrogen, and then diluted with ethyl acetate (10 mL) upon completion. The crude mixture was extracted against 0.1N HCl (10 mL × 3), concentrated by rotatory evaporation, and then purified by flash chromatography in 1:19 methanol/CH_2Cl_2. The product was isolated as a clear oil (0.0782 g, 0.24 mmol, 31% yield). R_f 0.16 in 1:19 methanol /CH_2Cl_2; ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.28 (s, 1H), 4.46 (dd, 1H, *J* = 7.8, 3.9 Hz), 3.11 (dd, 1H, *J* = 13.4, 3.8 Hz), 2.91 (dd, 1H, *J* = 13.4, 7.7 Hz), 1.31 (s, 9H), 0.91 (s, 9H), 0.14 (s, 3H), 0.12 (s, 3H). ESI$^-$-HRMS calculated for $C_{13}H_{27}O_3S_2Si^-$ 323.1171, found $[M - H]^-$ 323.1168.

***L*-Cystine and *L*-Cysteine Routes towards TBS-ChB-OH** Various routes using *L*-cystine or protected Cys as starting material were proven unsuccessful at various stages of synthesis, due to solubility or reactivity limitations.

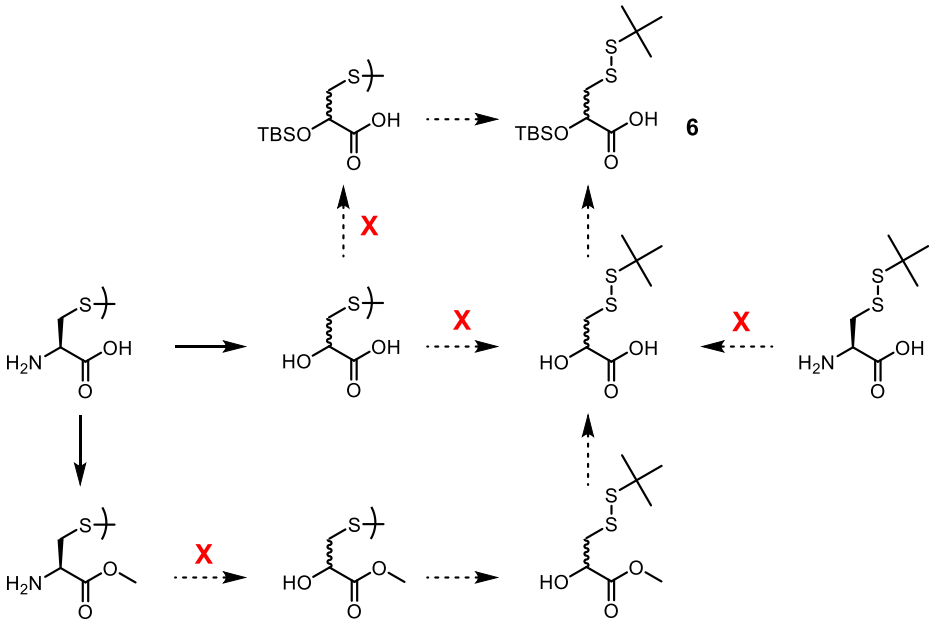


Figure 2-15. Attempted Synthesis towards TBS-ChB-OH from Cys$^{S\text{-}t\text{Bu}}$ or Cystine.

Synthesis of α-*N*-Fmoc-L-thioleucine-benzotriazolide (S1) Thioleucine (Leu^{S} or L^{S}) precursor was synthesized using a similar procedure as previously described in *J. Am. Chem. Soc.* **2012**, *134*, 6088−6091. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.78 (d, *J* = 8.3 Hz, 1H), 8.15 (d, *J* = 8.0 Hz, 1H), 7.77 (d, *J* = 7.3 Hz, 2H), 7.69 (d, *J* = 7.5 Hz, 1H), 7.65 (t, *J* = 9.1 Hz, 2H), 7.55 (t, *J* = 7.6 Hz, 1H), 7.40 (t, *J* = 7.3 Hz, 2H), 7.33 (t, *J* = 7.2 Hz, 2H), 6.39 (t, *J* = 8.8 Hz, 1H), 5.78 (d, *J* = 9.6 Hz, 1H), 4.49 (d, *J* = 7.0 Hz, 1H), 4.41 (t, *J* = 7.2 Hz, 1H), 4.25 (t, *J* = 6.5 Hz, 1H), 1.92 (s, 1H), 1.85 (t, *J* = 10.1 Hz, 1H), 1.15 (d, *J* = 6.2 Hz, 3 H), 0.98 (d, *J* = 6.4 Hz, 3H). ESI^{+}-LRMS calculated for $C_{27}H_{26}N_4O_2SNa^{+}$ 493.17, found $[M + Na]^{+}$ 493.26.

Synthesis of α-*N*-Fmoc-L-thiovaline-benzotriazolide (S2) Thiovaline (Val^S or V^S) precursor was synthesized by Solongo Batjargal using a similar procedure as detailed in *J. Am. Chem. Soc.* **2012**, *134*, 9172−9182.

Peptide Synthesis and Purification Peptides were synthesized using standard Fmoc solid phase peptide synthesis (SPPS) procedure on either Rink Amide or 2-chlorotrityl chloride resin (100 - 200 mesh; 0.6 mmol substitution/g). For ChB and C^bPG_o latent thioesters, the linkers were first installed, and then the remaining sequences were completed using the standard procedure. For coupling, 5 equiv of amino acid and 5 equiv of HBTU were dissolved in DMF, and then added to the resin. 10 equiv of *N*,*N*-diisopropylethylamine (DIPEA) was then added, and the mixture was stirred for 30 min. For deprotection, 20% piperidine in DMF was stirred with the resin for 20 min. Thioleucine (Leu^S or L^S) and thiovaline (Val^S or V^S) were introduced as activated benzotriazole precursors[283,115]; 3 equiv of the precursor was dissolved in DMF, and stirred with the resin for 1 h in the presence of 10 equiv DIPEA. For N-terminally acetylated peptides, an acetylation cocktail of 1:1:8 acetic anhydride/N-methylmorpholine/DMF was stirred with the resin twice, each for 10 min, after the last deprotection. Upon completion of SPPS, resins were rinsed thoroughly with CH_2Cl_2 and dried under vacuum.

For cleavage, resins were treated with a cleavage cocktail (38:1:1 trifluoroacetic acid/triisopropylsilane/H_2O unless otherwise specified) for 30 min. The solution was then collected by filtration, and dried by rotary evaporation. For purification, the crude

residues were brought up in 1:10 CH_3CN/H_2O, and then purified by reverse phase HPLC with acidified (with 0.1% trifluoroacetic acid) CH_3CN/H_2O gradients. Individual fractions were charac-terized by MALDI-TOF MS, and dried by lyophilization. When necessary, the isolated products were subjected to multiple rounds of purification until 99.5% pure by MALDI-TOF MS and analytical HPLC. Solvent radients, retention times and MALDI-TOF MS results are listed in Table 2-1 through Table 2-3.

Table 2-1. Peptide Purification Methods and Retention Time.

Peptide	Gradient	Retention Time	Column
Ac-G**L**^{S}KXAG-ChB-NH_2 (**7**)	**1**	20.9 min	Vydac C18 semi-prep
CAX-NH_2 (**8**)	**2**	18.5 min	Vydac C18 semi-prep
Ac-G**L**^{S}KXAG-C^{b}PG_o (**11**)	**1**	19.2 min	Vydac C18 semi-prep
Ac-MD**V**^{S}FMKGL-C^{b}PG_o (**13**)	**5**	21.0 min	Vydac C18 semi-prep
$\alpha S_{9\text{-}140}C_9$ (**14**)	**6**	24.4 min	Vydac C4 prep

* X = 7-methoxycoumarinylalanine; L^S = thioleucine; V^S = thiovaline.

Table 2-2. HPLC Gradients for Peptide Purification and Characterization.

No.	Time (min)	%B	**No.**	Time (min)	**%B**	**No.**	Time (min)	**%B**
1	0:00	2	**2**	0:00	2	**3**	0:00	2
	5:00	2		5:00	2		6:00	2
	10:00	30		10:00	15		30:00	50
	30:00	50		25:00	30		40:00	100
	35:00	100		30:00	100		45:00	100
	40:00	100		35:00	100		50:00	2
	45:00	2		40:00	2			
4	0:00	2	**5**	0:00	2	**6**	0:00	5
	5:00	2		5:00	2		5:00	5
	10:00	20		10:00	40		15:00	35
	25:00	50		25:00	55		30:00	50
	30:00	100		30:00	100		35:00	100
	35:00	100		35:00	100		40:00	100
	40:00	2		40:00	2		45:00	2

* Solvent A: 0.1% trifluoroacetic acid in water; Solvent B: 0.1% trifluoroacetic acid in acetonitrile

Table 2-3. MALDI-TOF MS Characterization of Purified Peptides.

Peptide	$[M + H]^+$		$[M + Na]^+$	
	Calc'd	**Found**	**Calc'd**	**Found**
Ac-G**L**^{S}KXAG-ChB-NH_2 (**7**)	939.37	939.23	961.36	961.20
CAX-NH_2 (**8**)	437.14	436.99	459.13	458.96
Ac-G**L**^{S}KXAG-$C^{b}PG_o$ (**11**)	1093.44	1093.36	1115.43	1115.34
Ac-MD**V**^{S}FMKGL-$C^{b}PG_o$ (**13**)	1357.58	1357.38	1379.56	1379.50
$\alpha S_{9\text{-}140}C_9$ (**14**)	13552.2	13552.8	-	-

* X = 7-methoxycoumarinylalanine; L^S = thioleucine; V^S = thiovaline.

Synthesis of Peptide-ChB Latent Thioester (**7**) Rink amide resin (25 μmol, 0.6 mmol/g substitution, 1 equiv) was swollen in dimethylformaide (DMF) for 1 h, and then deprotected twice with 2 mL of 20% piperdine in DMF for 20 min each. For ChB installation, TBS-ChB-OH (**6**, 0.0089 g, 27.5 μmol, 1.1 equiv) and PyBOP (0.0143 g, 27.5 μmol, 1.1 equiv) were dissolved in 2 mL of DMF, and then added to the resin. DIPEA (43.5 μL, 250 μmol, 10 equiv) was added, and then the reaction was allowed to proceed for 90 min with stirring. Upon completion, the resin was thoroughly rinsed with DMF and CH_2Cl_2, and then stirred in tetrahydrofuran (THF) for 1 h for equilibration. For TBS deprotection, 2 mL of 1.0 M tetra-*n*-butylammonium (TBAF) in THF was added. The reaction was allowed to proceed for 4 h, and then thoroughly rinsed with THF.

For coupling of the next residue Gly, Fmoc-Gly-OH (0.0516 g, 125 μmol. 5 equiv) and dicyclohexylcarbodiimide (DCC, 0.0129 g, 62.5 μmol, 2.5 equiv) were dissolved in 2 mL of THF, and then pre-activated for 30 min. Dicyclohexylurea (DCU) side product

was removed by filtration, and then the clear solution was added to the resin. Catalytic amount of 4-dimethylaminopyridine (DMAP, 0.0003 g, 2.5 µmol, 0.1 equiv) was added, and then the reaction was allowed to proceed for 45 min. Upon completion, the resin was thoroughly rinsed with THF, and then stirred in DMF for 1 h for equilibration. The rest of the sequence was elongated using standard SPPS protocol, cleaved with 3:7 trifluoro-acetic acid/CH_2Cl_2 (3 mL), and purified by reverse phase HPLC (Table 2-1 to Table 2-3).

Test Ligation of Peptide-ChB Latent Thioester A ligation buffer (20 mM TCEP, 200 mM Na_2HPO_4, pH 6.8) was freshly prepared and degassed by argon purging. **7** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}, 1 equiv) and **8** (100 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}, 10 equiv) was each dissolved in 10 µL of ligation buffer and then combined. The reaction was allowed to proceed at 37 ˚C for various amount of time under argon atmosphere. Upon completion, 5 µL of the reaction mixture was removed, diluted into 795 uL of H_2O, and analyzed by reverse phase HPLC on a Luna C8 analytical column using gradient **3**. Fractions were collected for MALDI-TOF MS analysis, and UV-Vis absorption profiles were extracted from data acquired by photodiode array detector on analytical HPLC.

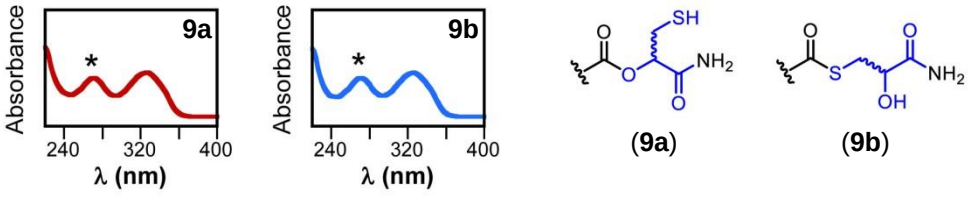


Figure 2-16. UV-Vis Absorption Spectra for Intermediate **9a** and **9b**.

Denaturant Tolerance of ChB Ligation Reactions were conducted similarly to the standard ChB ligation procedure, except that a ligation buffer containing 6 M Gdn·HCl was used. The presence of denaturant resulted in slower reaction kinetics, but otherwise no additional complications were observed.

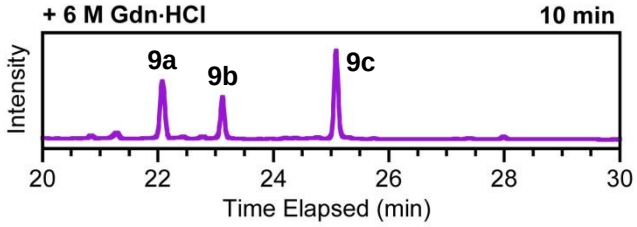


Figure 2-17. Denaturant Tolerance of ChB Ligation between **7** and **8**.

Synthesis of Peptide-C^{b}PG$_{o}$ Latent Thioester (11) Ac-GL^{S}KXAG-C^{b}PG$_{o}$ **11** (X = 7-methoxycoumarinylalanine) was synthesized on Rink amide resin (100 μmol, 0.6 mmol/g). The resin was deprotected with 20% piperidine in dimethylformamide (DMF), thoroughly rinsed, and then dried on vacuum. Bromoacetic acid (0.0695 g, 500 μmol, 5 equiv) was pre-activated with dicyclohexylcarbodiimide (DCC, 0.0516 g, 250 μmol, 2.5 equiv) for 30 min in anhydrous DMF (6 mL); dicyclohexylurea (DCU) side product was removed by filtration, and then the clear solution was added to the resin. The reaction was allowed to proceed with stirring for 30 min, and then thoroughly rinsed with DMF. Subsequently, Fmoc-Pro-OH (0.1687 g, 500 μmol, 5 equiv) was dissolved in anhydrous DMF (6 mL) with diisopropylethylamine (DIPEA, 174.3 μL, 1 mmol, 10 equiv), and coupled to the resin for 30 min.

The next two residues were introduced as a dipeptide, to avoid premature diketo-piperazine (DKP) formation on resin (Figure 2-18). To accomplish this, Fmoc-Gly-Cys(tButhio)-OH was synthesized on 2-chlorotrityl resin at 125 μmol scale using standard SPPS, and cleaved with 1:1:8 acetic acid/trifluoroethanol/CH_2Cl_2. The crude product (0.0537 g, 110 μmol, 1.1 equiv) was dissolved in anhydrous DMF (6 mL) with PyBOP (0.0572 g, 110 μmol, 1.1 equiv) and DIPEA (174.3 μL, 1 mmol, 10 equiv), and then coupled to the resin for 1 h. The rest of the peptide was elongated and acetylated with standard SPPS. For cleavage, the resin was treated with 30:70 trifluoroacetic acid/CH_2Cl_2 (6 mL) for 30 min, dried by rotary evaporation, and then purified by reverse phase HPLC (Table 2-1 through Table 2-3).

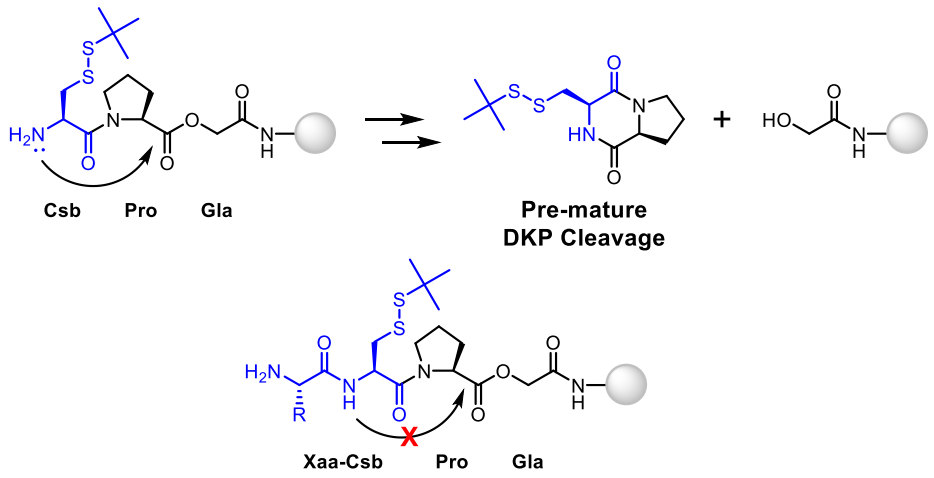


Figure 2-18. Necessity of Fmoc-Xaa-Csb-OH in Preventing DKP Formation in SPPS.

Ligation between Peptide-C^bPG_o and Cys A 2x ligation buffer (100 mM TCEP, 400 mM Na_2HPO_4, pH 8.4) was freshly prepared and degassed by argon purging. Ac-GL^SKXAG-C^bPG_o **11** (100 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in argon-purged

water (30 μL), and mixed with 20 μL of 5x Cys stock (50 mM in argon-purged water). 50 μL of 2x ligation buffer was added, and then the reaction was allowed to proceed at 37 ˚C for 10 h. At appropriate time points, a 4 μL sample was removed from the reaction mixture, quenched with 796 μL of 0.1% trifluoroacetic in water, and analyzed by reverse phase HPLC on a Luna C8 analytical column using gradient **4**. Fractions were collected for MALDI-TOF MS analysis, and UV-Viv absorption profiles were extracted from data acquired by photodiode array detector on analytical HPLC.

Kinetic Characterization of C^bPG_o Ligation To obtain a kinetic profile of C^bPG_o ligation, HPLC chromatograms of samples taken at various time points were integrated for areas under the curve based on absorbance at 325 nm (λ_{max} for Mcm, 7-methoxycoumarinylalanine). Peak identities were assigned based on MALDI MS, and then LC area percentages (LCAP) were calculated using equation **S1**:

LCAP = Area of individual peak at 325 nm / Area of all peaks at 325 nm (**S1**)

For data fitting, LCAP for each peak was averaged from three independent trials, and plotted against the time point at which they were obtained. Since all Mcm-containing peptides originate from the reactant peptide **11**, their total concentration will stay the same as initial concentration of **11** (1 mM) and their individual concentrations (in terms of molarity) will also be proportional to their LCAPs. With Cys in large excess (10 equiv) and the rapid conversion of C^bPG_o to its reduced form, the reaction can be assumed as a

pseudo-first-order reaction of the reduced reactant **11a**. Therefore, LCAPs were fitted to a single exponential decay function **S2** using Microcal Origin 6.0 Pro ((Northampton, MA).

$$LCAP = LCAP_0 + A \cdot \exp [-t / \tau] \qquad (\mathbf{S2})$$

The fitting parameters were then converted into pseudo first order rate constant k_1', second order rate constant k_2, and half-life $t_{1/2}$ using equations **S3** through **S5**.

$$k_1' = 1 / \tau \qquad (\mathbf{S3})$$

$$k_2 = k_1 / [Cys] \qquad (\mathbf{S4})$$

$$t_{1/2} = \tau \cdot \ln 2 \qquad (\mathbf{S5})$$

With an *R*-value of 0.9998, the experimental pseudo first order rate constant k_1' was calculated to be $(1.53 \pm 0.06) \times 10^{-4}$ s^{-1}, corresponding to a second order rate constant k_2 of $(1.53 \pm 0.06) \times 10^{-2}$ s^{-1} and a reaction half-life $t_{1/2}$ of (75 ± 1) min.

Diketopiperazine (DKP) Thioester Intermediate The DKP thioester intermediate **S3** could be observed when 0.1 mM peptide-C^bPG_o **11** was incubated with ligation buffer in the absence of Cys (Figure 2-19). Briefly, **11** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in 50 μL of argon-purged water, and then 50 μL of 2x argon-purged ligation buffer stock (100 mM TCEP, 400 mM Na_2HPO_4, pH 8.4) was added. The reaction was allowed to proceed at 37 ˚C for 45 min, and then quenched with 900 μL of 0.1% trifluoroacetic acid in water and analyzed by reverse phase HPLC on a Vydac C18 analytical column using gradient **4**. Fractions were collected for MALDI-TOF MS.

Denaturant Tolerance of C^bPG_o Ligation Three Cys stocks (10 mM Cys, 20 mM TCEP, 200 mM Na_2HPO4, pH 8.0) were prepared with various additives: A) none; B) 6 M Gdn·HCl; C) 6 M Gdn·HCl and 50 mM 4-mercaptophenylacetic acid (MPAA). For each condition, **11** (20 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in the corresponding Cys stock (20 µL), and allowed to react at 37 ˚C for 2 h. Upon completion, the reaction was diluted into 780 µL of 0.1% trifluoroacetic acid in water, and analyzed by reverse phase HPLC (Figure 2-20). The results showed that the presence of Gdn·HCl led to slower kinetics, whereas the addition of MPAA could partially "rescue" the reactivity.

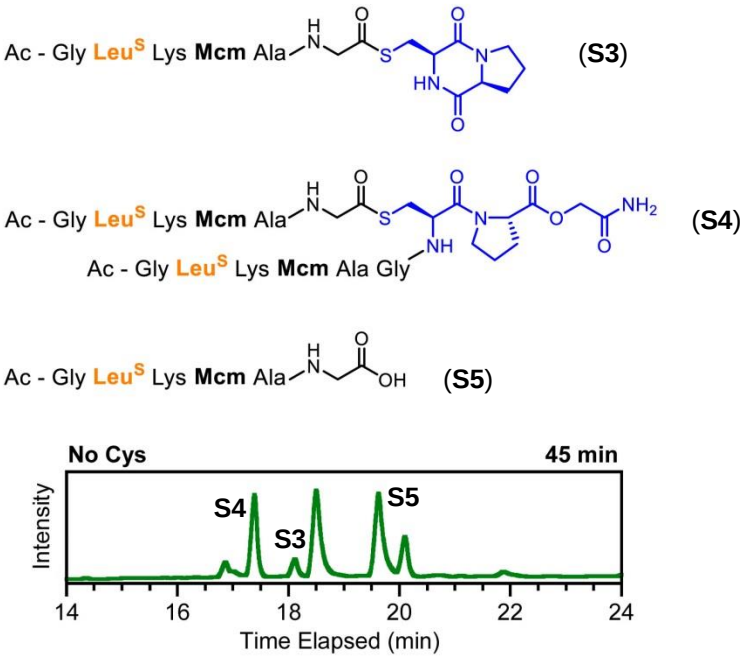


Figure 2-19. Diketopiperazine (DKP) Intermediate in C^bPG_o Ligation.

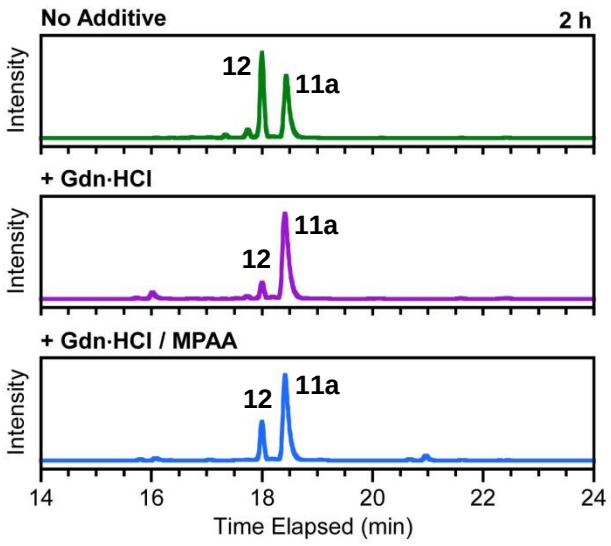


Figure 2-20. Effects of Denaturant and Thiol Additive on C^bPG_o Ligation.

Synthesis of Fmoc-Leu-Cys(S-*t*Bu)-OH (S6) Fmoc-Leu-OH (1.0602 g, 3 mmol, 1.5 equiv) was dissolved in 20 mL of tetrahydrofuran (THF), and chilled to 0 °C on ice. *N*-Methylmorpholine (NMM, 769 µL, 7 mmol, 3.5 equiv) and isobutyl chloroformate (IBCF, 290 µL, 2.2 mmol, 1.1 equiv) were added, and then the reaction was stirred for 15 min on ice. *S*-*t*-butylmercapto-*L*-cysteine (0.4186 g, 2 mmol, 1 equiv) was added, and the reaction was allowed to proceed at room temperature overnight (≥ 14 hr) under an argon atmosphere. Upon completion, solvent was removed by rotary evaporation. The residue was brought up in 40 mL of ethyl acetate, and extracted against 40 mL each of 1 M HCl and brine. The crude product was purified by flash chromatography in 40:60 ethyl acetate/petroleum ether with 0.1% acetic acid. A white foam was yielded as the final product (0.7643 g, 1.29 mmol, 70% yield). R_f 0.5 in 4:6 ethyl acetate/petroleum ether with 0.1% acetic acid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.38 (s, 1H), 7.73 (d, *J* = 7.1 Hz, 2H), 7.59 (t, *J* = 6.8 Hz, 2H), 7.41 (d, *J* = 7.0 Hz, 1H), 7.38 (t, *J* = 7.2 Hz, 2H), 7.29 (t, *J*

= 7.2 Hz, 2H), 5.95 (d, *J* = 8.3 Hz, 1H), 4.86 (m, 1H), 4.45 (m, 1H), 4.38 (m, 2H), 4.21 (t, *J* = 6.5 Hz, 1H), 3.28 (dd, *J* = 13.4, 3.9 Hz, 1H), 3.17 (dd, *J* = 13.3, 6.0 Hz, 1H), 1.69 (m, 2H), 1.59 (m, 2H), 1.26 (s, 9H), 0.94 (m, 6H). ^{13}C NMR (500 MHz, $CDCl_3$): δ 173.10, 156.72, 144.09, 143.86, 141.48, 127.93, 126.33, 125.41, 120.17, 68.90, 67.56, 53.60, 52.71, 48.37, 47.29, 41.89, 41.54, 30.02, 29.96, 28.08, 24.87, 23.15, 22.34, 22.29. ESI$^+$-HRMS: calculated for $C_{28}H_{37}N_2O_5S_2^+$ 545.2144, found $[M + H]^+$ 545.2150.

Synthesis of Ac-α$S_{1\text{-}8}V^S{}_3$-C^bPG_o (13) Rink amide resin (100 μmol, 0.6 mmol/g substitution) was swollen in anhydrous DMF for 1 h, and then deprotected twice with 6 mL of 20% piperidine in anhydrous DMF, each for 20 min. $BrCH_2COOH$ (0.0695 g, 500 μmol, 5 equiv) was dissolved in 6 mL of anhydrous DMF, and pre-activated with diisopropylcarbodiimide (DIC, 78 μL, 500 μmol, 5 equiv) for 20 min. The mixture was then transferred to the resin, and the coupling reaction was allowed to proceed for 45 min. For Pro installation, Fmoc-Pro-OH (0.1687 g, 500 μmol, 5 equiv) was dissolved in 6 mL of anhydrous DMF and transferred to the resin. DIPEA (174 μL, 1 mmol, 10 equiv) was added, and then the reaction was allowed to proceed for 4 h. A second coupling was repeated with fresh Fmoc-Pro-OH/DIEPA/DMF overnight (≥ 8 h). The resin was then deprotected with 6 mL of 20% piperidine in anhydrous DMF for 20 min.

To install the next two amino acids as dipeptide, Fmoc-Leu-Cys(S-*t*Bu)-OH (0.1634 g, 300 μmol, 3 equiv) was dissolved in 6 mL of anhydrous DMF, and pre-activated with PyBOP (0.1561 g, 300 μmol, 3 equiv) and DIPEA (174 μL, 1 mmol, 10 equiv) for 3 min.

The mixture was then transferred to the resin, and the coupling reaction was allowed to proceed for 4 h. The rest of the sequence was assembled using standard SPPS procedure.

For cleavage, the resin was treated with 6 mL of 1:1:2:20 1,2-ethanedithiol/triiso-propylsilane/thioanisole/trifluoracetic acid for 30 min. The solution was then collected by filtration, concentrated by rotary evaporation, and then precipitated with 10 mL of cold diethyl ether. The pellets were collected by centrifugation at 4 krpm for 20 min, and then brought up in 50:50 H_2O/CH_3CN for reverse phase HPLC purification.

Over-expression and Purification of $\alpha S_{9\text{-}140}C_9$ or $\alpha S_{9\text{-}140}C_9W_{94}$ Expression and purification were performed by Solongo Batjargal as detailed in *J. Am. Chem. Soc.* **2012**, *134*, 9172−9182. Briefly, a plasmid encoding for H_{Tag}-$\alpha S_{9-140}C_9$ or H_{Tag}-$\alpha S_{9-140}C_9W_{94}$ were transformed into *E. coli* BL21 DE3 cells, and over-expressed with isopropyl-*β*-D-thiogalactopyranoside (IPTG) induction. After lysis of the cells, target protein was pulled down by Ni-NTA resin, and treated with Factor Xa to remove the H_{Tag}. The crude was then dialyzed against water, and further purified by reverse phase HPLC on a Vydac C4 prep column using gradient **6** (Tables Table ***2-1*** through Table ***2-3***).

Ligation between Ac-$\alpha S_{1\text{-}8}V^S_3$-$C^bPG_o$ and $\alpha S_{9\text{-}140}C_9$ or $\alpha S_{9\text{-}140}C_9W_{94}$ Latent thioester Ac-$\alpha S_{1-8}V^S_3$-C^bPG_o **13** (90 nmol, 1.5 equiv) and expressed protein fragment $\alpha S_{9\text{-}140}C_9$ or $\alpha S_{9\text{-}140}C_9W_{94}$ (60 nmol, 1 equiv) were dissolved in 60 μL of freshly degassed ligation buffer (6 M Gdn·HCl, 200 mM Na_2HPO_4, 20 mM TCEP, 1% v/v thiophenol, and 1% v/v benzylmercaptan, pH 8.0). The reaction was allowed to proceed for 24 h at 37 °C,

and then quenched with 940 µL of 0.1% trifluoroacetic acid in water. The crude was dialyzed against water, and purified by reverse phase HPLC. Fractions were collected, characterized by MALDI-TOF MS, and then lyophilized until further use.

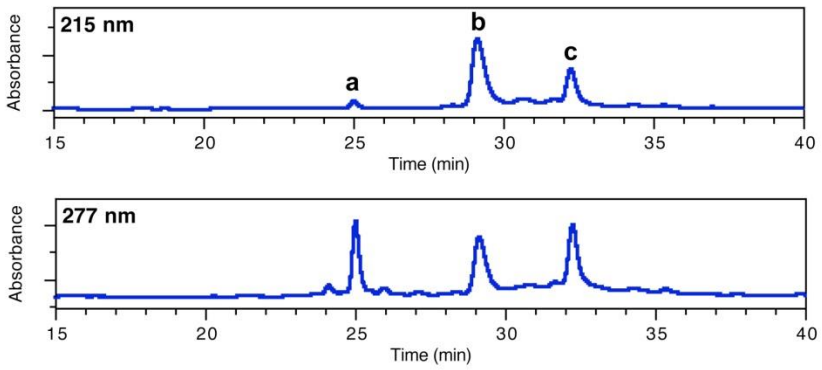


Figure 2-21. HPLC Chromatogram of Ligation between $\alpha S_{1\text{-}8}V^{S}_{3}$-CbPG$_o$ and $\alpha S_{9\text{-}140}C_9$.

Peaks a, b,and c correspond to hydrolyzed Ac-$\alpha S_{1\text{-}8}V^{S}_{3}$-C^{b}PG$_o$, $\alpha S_{9\text{-}140}C_9$, and the ligation product, Ac-$\alpha S_{1\text{-}140}V^{S}_{3}C_9$. UV absorbance monitored at 215 and 277 nm.

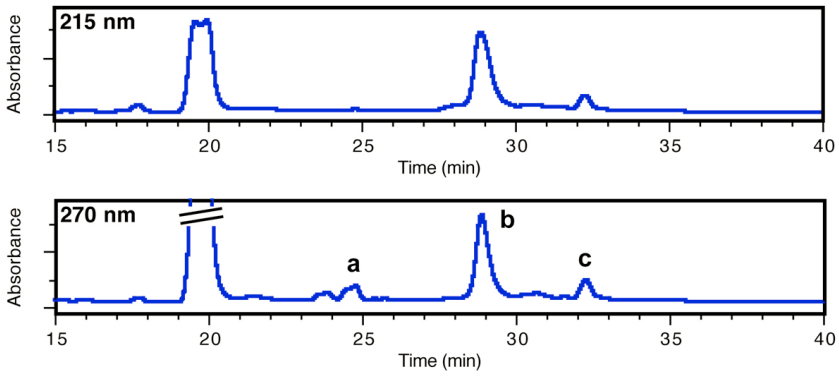


Figure 2-22. HPLC Chromatogram of Ligation between $\alpha S_{1\text{-}8}V^{S}_{3}$-CbPG$_o$ and $\alpha S_{9\text{-}140}C_9W_{94}$.

Peaks a, b, and c correspond to hydrolyzed Ac-$\alpha S_{1\text{-}8}VS_3$-C^{b}PG$_o$, $\alpha S_{9\text{-}140}C_9W_{94}$, and ligation product, Ac-$\alpha S_{1\text{-}140}VS_3C_9W_{94}$. UV absorbance monitored at 215 and 270 nm. The large peak at 20 min is excess thiophenol additive.

Trypsin Digest Analysis of Ac-αSV$^{S}_{3}$C$_9$ Ligated Ac-αSV$^{S}_{3}$C$_9$ **15** (20 μg) was dissolved in 18 μL of digestion buffer (25 mM NH_4HCO_3, pH 7.5), and treated with 2 μL of sequencing-grade modified trypsin (0.1 μg/μL). The digestion was allowed to proceed at 37 °C for 4 h, and then analyzed by MALDI-TOF MS.

UV-Vis Absorption Spectroscopy Analysis αS$_{9\text{-}140}$C$_9$ contains four Tyr with a molar absorption of 5,340 M^{-1} cm^{-1} at 273 nm; Ac-αSV$^{S}_{3}$C$_9$ contains four Tyr and one thioamide with a molar absorption of 15,368 M^{-1} cm^{-1} at 273 nm. Therefore, the molar absorption ratio (Ac-αSV$^{S}_{3}$C$_9$: αS$_{9\text{-}140}$C$_9$) at 273 nm is 2.88. The absorption spectra of the two proteins were normalized by assuming that the molar absorption of Ac-αSV$^{S}_{3}$C$_9$ (140 residues) at 218 nm is 1.07 times that of αS$_{9\text{-}140}$C$_9$ (131 residues). The experimental absorbance ratio (Ac-αSV$^{S}_{3}$C$_9$: αS$_{9\text{-}140}$C$_9$) at 273 nm is 2.84, indicating 2.84/2.88 = 99 % thioamide incorporation.

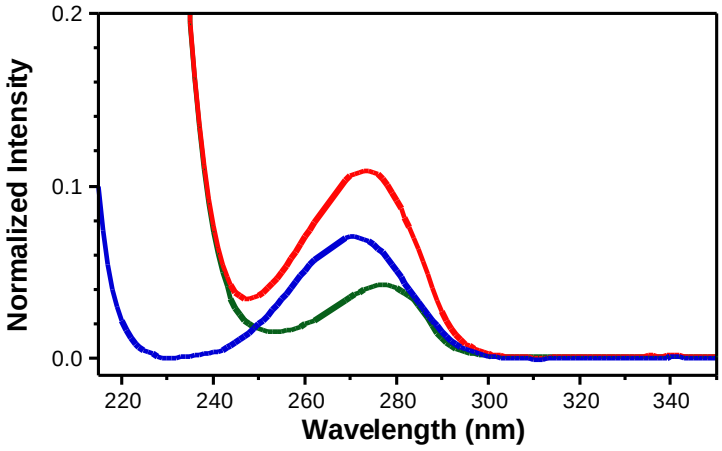


Figure 2-23. Normalized UV-Vis Absorption Spectra of αS Ligation. αS$_{9\text{-}140}$C$_9$ (green line), αSV$^{S}_{3}$C$_9$ red line), and difference spectrum (blue line).

Aggregation of Ac-αS$_{1\text{-}140}$V$^{S}_{5}$C$_{9}$W$_{94}$ Aggregation experiments were performed according to literature precedence.[284,285] First, 350 μL samples for aggregation assays were prepared (97 μM WT αS with 3 μM αSW$_{94}$ or 3 μM Ac-αSV$^{S}_{3}$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 4−6 days with continuous shaking at 1,100 rpm. Periodically, 40 μL aliquots were removed to monitor changes in Trp fluorescence and ThT binding.

For fluorescence spectroscopy, 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 below.

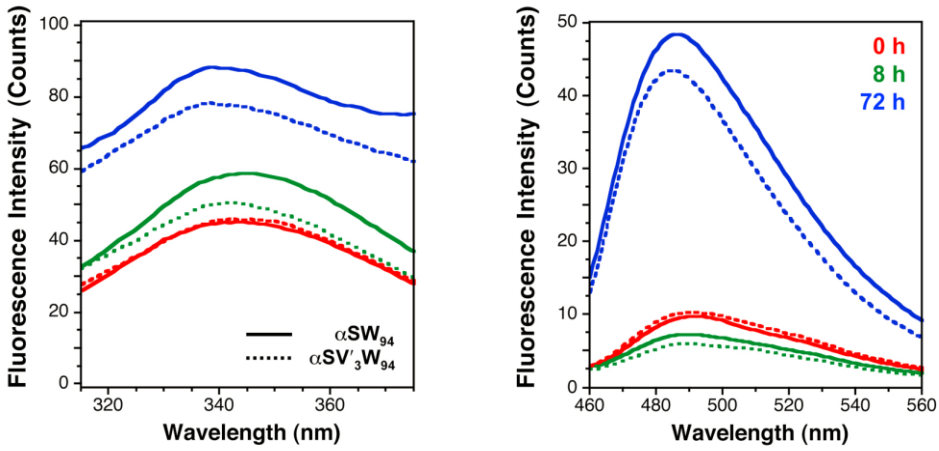


Figure 2-24. Examples Fluorescence Spectra from αS Aggregation Experiments.

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 Ac-$\alpha SV^{S}_{3}C_{9}W_{94}$/WT mix. With Rebecca F. Wissner.

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.

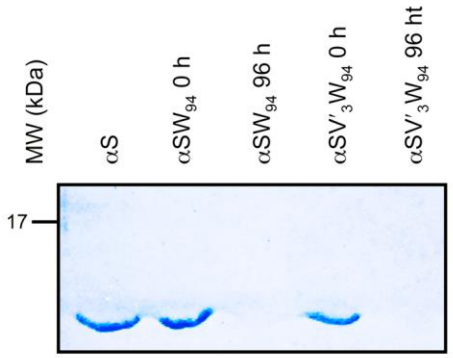


Figure 2-25. PAGE Gel Analysis of Aggregation Experiments.

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. With Rebecca F. Wissner.

2.5 Acknowledgement

This work was supported by funding from the University of Pennsylvania, Searle Scholars Program (10-SSP-214 to EJP), and the National Science Foundation (NSF CHE-1020205 to EJP). Instruments supported by the National Science Foundation and National Institutes of Health include: HRMS (NIH RR-023444) and MALDI-MS (NSF MRI-0820996). We thank Dr. Rakesh Kohli for assistance with HRMS analysis, Prof. Elizabeth Rhoades for the gift of the pT7 plasmid containing α-synuclein.

Chapter 3 . Chemoselective Desulfurization and Deselenization for Traceless Incorporation of Thioamides into Peptides and Proteins

3.1 Introduction

While traditional native chemical ligation (NCL), as developed by Kent[128], was proven successful in incorporating thioamides into peptides and proteins, one inherent limitation of this method is the necessity of a Cys at the ligation site (Figure 3-1A). This greatly restricts the scope of thioamide incorporation – Cys is among the least abundant amino acids (2.4% frequency in the human proteome[286]) and does not exist in the native sequences of many protein targets, including α-synuclein, a Parkinson's disease protein of interest to our laboratory For these systems, thioamide incorporation via NCL would leave an artifact of residual Cys in the ligated product, undermining the utility of the thioamide as a minimalist probe and introducing complications during biophysical studies. In fact, it has been clearly demonstrated that Cys mutations in αS can alter its aggregation and cellular toxicity.[287]

To circumvent the Cys limitation in NCL, various groups have developed traceless methods by either masking or "erasing" the side chain thiol. Masking techniques utilize alkyl halides to convert Cys into Lys/Glu/Gln analogs, or homocysteine (Hcs) into Met after ligation[130] (Figure 3-1C). The former is not truly traceless since it results in a thioether linkage (and an extra methylene group as in the case of Glu/Gln) that is chemically and electrostatically different from the native residue. Hcs methylation does yield a native Met, and has previously been utilized by our group for the synthesis of thioamide-containing proteins[269,288], granting access to another 2.1% of potential ligation sites (Met abundance out of all 20 natural amino acids).[286]

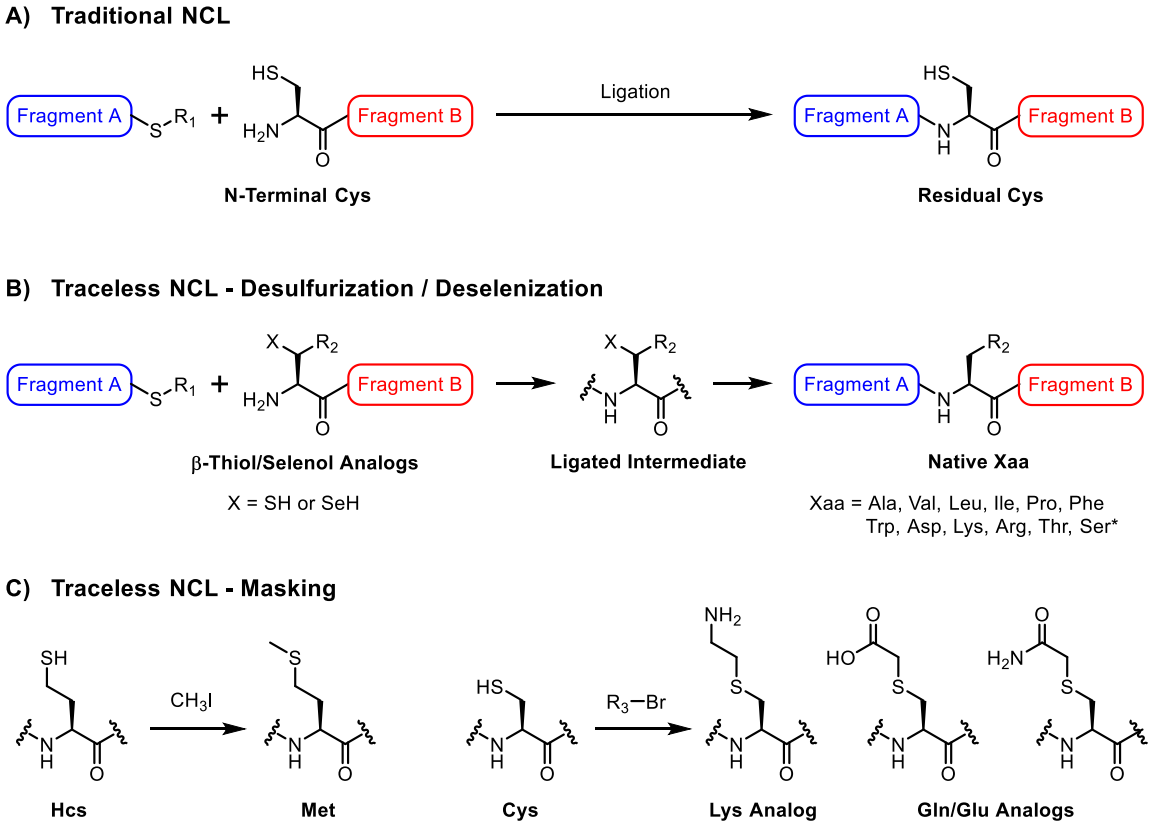


Figure 3-1. Comparison of Traditional and Traceless Native Chemical Ligation (NCL).

A) Traditional NCL utilizes an N-terminal Cys as the nucleophile, and leaves a residual Cys in the ligated product. B) Traceless NCL via desulfurization/deselenization requires a Cys or β-thiol/selenol analog as the nucleophile; the thiol/selenol is chemically removed after NCL. C) Traceless NCL via masking uses either Cys or homocysteine (Hcs) as the nucleophile; the thiol is subsequently masked into a functionalized thioether to mimic the natural side chain. * Conditions are different for Ser as compared to all other amino acids.

The true paradigm shift, though, would come from adopting recently developed desulfurization /deselenization methods, where the N-terminal residue is decorated with a thiol or selenol at its β- or γ-carbon as the NCL handle, and then chemically "erased" after ligation to become the native amino acid[218,202] (Figure 3-1B). These methods are extremely versatile – with nearly identical reaction conditions, it has thus far been concretely demonstrated on 12 amino acids[219] and theoretically possible for 19 out the 20

amino acids (except for Gly which does not have a β-carbon). In addition, the β-selenol analogs offer a second level of selectivity against native Cys residues, where the mild deselenization conditions using tris(2-carboxyethyl)phosphine (TCEP) are fully compatible with side chain thiols.[216] However, it is unclear whether such methods would be amenable to the incorporation of thioamides into peptides and proteins, particularly in the case of desulfurization. Since both the thioamide and Cys contain a sulfur atom, characterizing and fine-tuning the chemoselectivity would seem quite challenging.

Upon a closer examination at the reactions utilized – namely desulfurization by Raney nickel, desulfurization by VA-044, and deselenization by TCEP – we hypothesize that Raney nickel would be too harsh but the latter two should be feasible for thioamide-containing peptides and proteins. Raney nickel is well-established for its broad substrate scope where thiols, thioethers and thionoesters (thiocarbonyl esters) are all effectively desulfurized, largely because of the high nickel-sulfur affinity.[289] The only well-characterized precedence of Raney nickel/thioamide reactions was published by Kornfield, where he reported a reductive desulfurization into amine (Figure 3-2).[290,291] Although the exact reaction conditions would be different in our traceless NCL (for example, it will be conducted in aqueous solution instead of organic solvents), one would assume that Raney nickel could lead to complications with thioamides.

On the other hand, radical initiated desulfurization and deselenization do not depend on metal-sulfur affinity, therefore as long as thioamide could tolerates an environment with certain concentrations of radicals, it is possible to conduct traceless ligation while

leaving the thioamide intact. Based on our prior knowledge that thioamide can function as an acceptor for photoinduced electron transfer (PET)[115,114] – where a single electron is transferred from thioamide to a paired fluorophore, generating a transient thioamide radical cation in the process (Figure 3-3) – we believed that the thioamide could be stable towards ambient radicals in desulfurization/deselenization reactions. If our hypothesis was valid, we would greatly expand the scope of potential ligation sites and remove the possibility of complications from thiol side chains in biophysical assays.

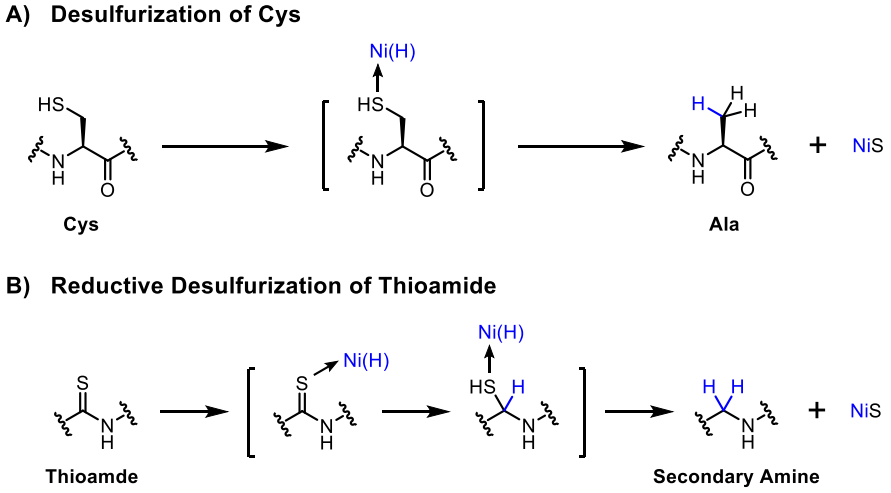


Figure 3-2. Raney Nickel Desulfurization of Cys and Thioamide.

A) Mechanism for Cys desulfurization. B) Proposed mechanism for reductive thioamide desulfurization as reported by Kornfield. * Raney nickel is denoted as Ni(H) to represent that the reagent contains H_2 as part of the formulation; it is generally understood that the hydrogen in the resulting C–H bond comes from the Raney nickel reagent.

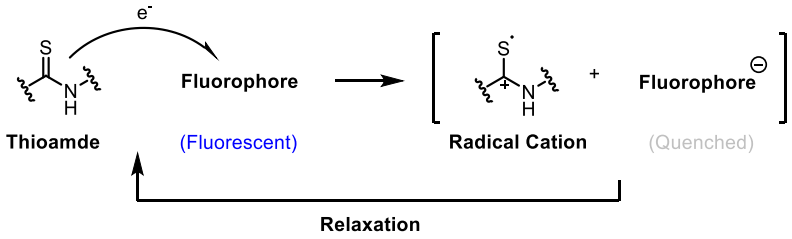


Figure 3-3. Transient Thioamide Radical Cation in Photo-induced Electron Transfer (PET).

3.2 Results and Discussion

We began by establishing the compatibility of the thioamide with various desulfurization/deselenization methods. First, we tested Raney nickel desulfurization on model peptide **1**. Upon treatment with Raney nickel, peptide **1** is fully consumed; surprisingly, instead of a secondary amine from reductive desulfurization[290], cleavage at the thioamide bond was observed, yielding peptide **1a** with an N-terminal amine at the cleavage site (Figure 3-4). We reasoned that since the reaction was conducted in an aqueous solution (instead of organic solvents as in the previous study), water could serve as an excellent nucleophile to attack the thiocarbonyl group that is activated by nickel-sulfur binding (Figure 3-5). While Raney nickel has been a useful reagent in other contexts, it is evident that such method would not be a viable option for the incorporation of thioamides.

We next sought to explore the compatibility of the thioamide with radical initiated desulfurization/deselenization. We chose to start with TCEP induced deselenization, to take advantage of our prior knowledge that the thioamide *per se* is stable towards TCEP.[283] Without attempting to synthesize a complex Sec-containing peptide in this initial trial, we designed a one-pot ligation-deselenization reaction, where a thioamide-containing CPG_o thioester **2a** (CPG_o = Cys-Pro-Gla latent thioester; see Chapter 2) was ligated to selenocystine, and then treated with TCEP for deselenization. We were delighted to find that both the ligation and deselenization proceeded smoothly and chemoselectively, where peptide **2d** was generated as the final product (Figure 3-6). **2d** exhibited the characteristic thioamide π-π^* absorption at 272 nm, giving us confidence

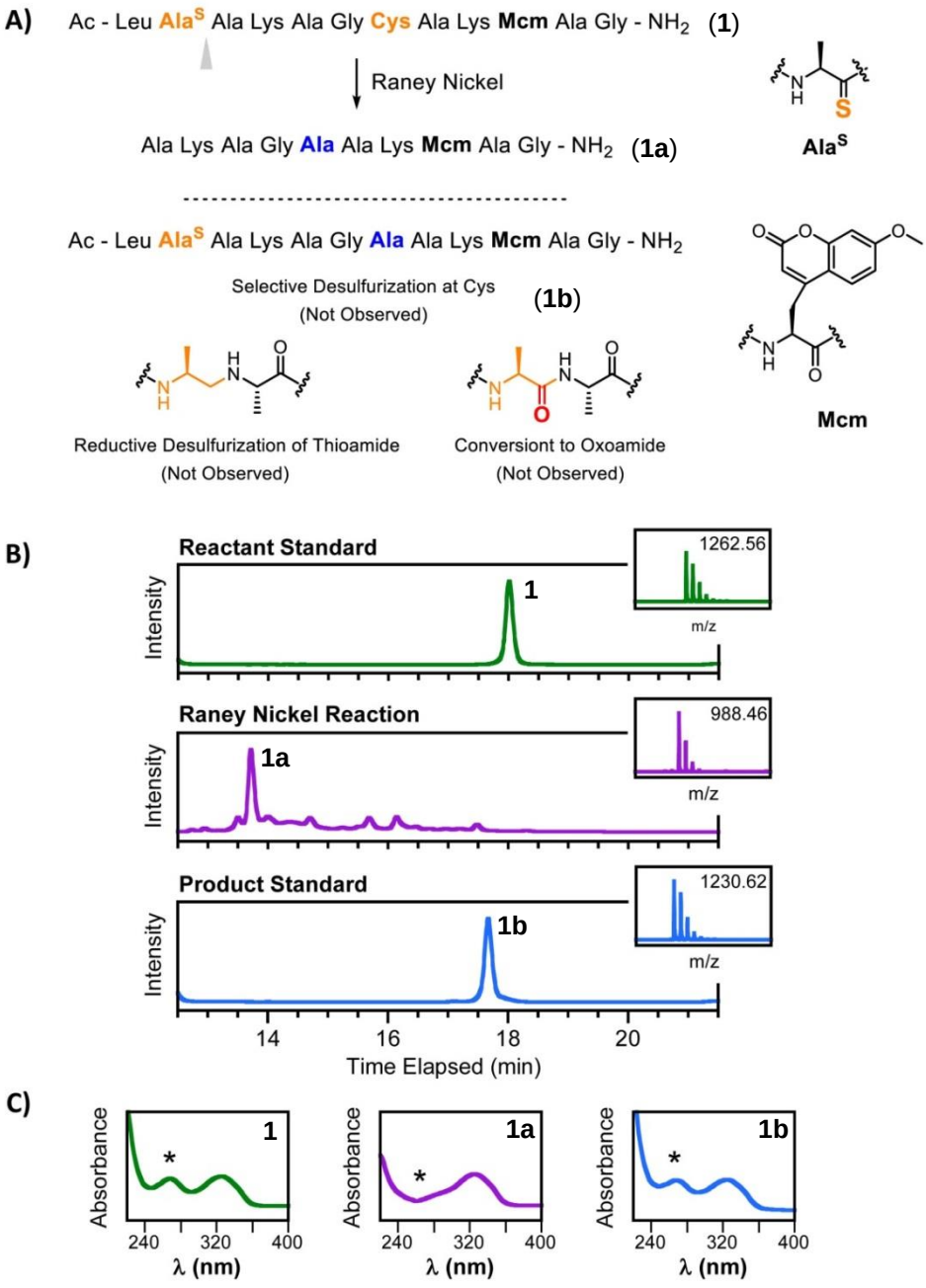


Figure 3-4. Non-Selective Desulfurization of Cys and Thioamide by Raney Nickel.

A) Schematic representation of the test reaction. B) HPLC chromatogram monitored at 325 nm, and MALDI-TOF mass spectra for major peak in each chromatogram; [**1** + H]$^+$: expected 1262.96, found 1262.56; [**1a** + H]$^+$: expected 988.53, found 988.46; [**1b** + H]$^+$: expected 1230.63, found 1230.62. C) UV-Vis absorption spectra of isolated peptides. The star sign indicates thioamide absorption at 272 nm, which was absent in **1a**. Conditions: 0.1 mM peptide **1**, 0.1% w/v Raney nickel, 100 mM Na_2HPO_4, 10 mM TCEP, pH 5.8, 12 h. TCEP = tris(2-carboxyethyl)phosphine; Mcm = 7-methoxycoumarinylalanine.

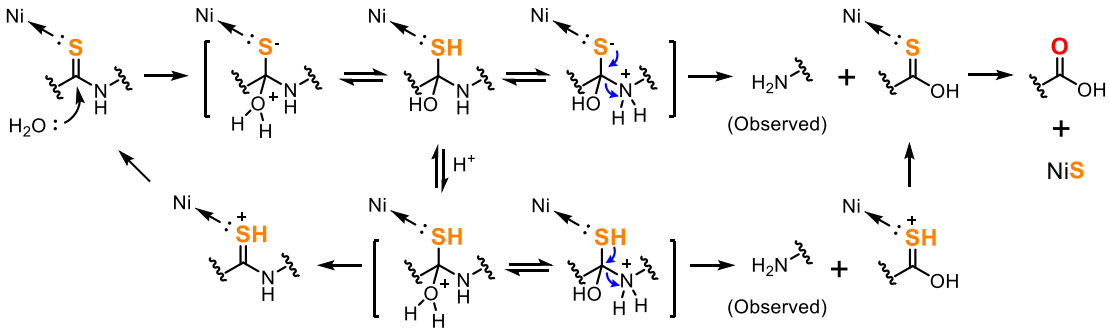


Figure 3-5. Proposed Mechanism for Raney Nickel Induced Thioamide Bond Cleavage.

that the thioamide remained intact throughout the process. It is interesting to note that due to the unique activity of selenol[292,293], two ligation intermediates were isolated − namely homodimer **2b** *via* a Sec diselenide bond and monomer **2c** that was sequestered at Sec through a hemiselenide bond to a −SPh group from ambient thiophenol used in the ligation – instead of free selenol. In subsequent deselenization, both species were reduced into selenol upon treatment with dithiothreitol (DTT) and converted into Ala.

The positive results from selective deselenization prompted us to further investigate the possibility of selective desulfurization with VA-044 as radical initiator This is a more challenging method for two reasons: 1) there is a smaller energy difference between a C=S double bond (standard bond enthalpy 137 kcal·mol^{-1}) and C−S single bond (65 kcal·mol^{-1}), than with a C−Se single bond (56 kcal·mol^{-1})[77], requiring the method to be more stringently discriminating against the thioamide; 2) mechanistically, deselenization of Sec does not require ambient radicals and is simply initiated by spontaneous cleavage of the Se−X (X = Se, S, H) bond, whereas desulfurization of Cys requires high

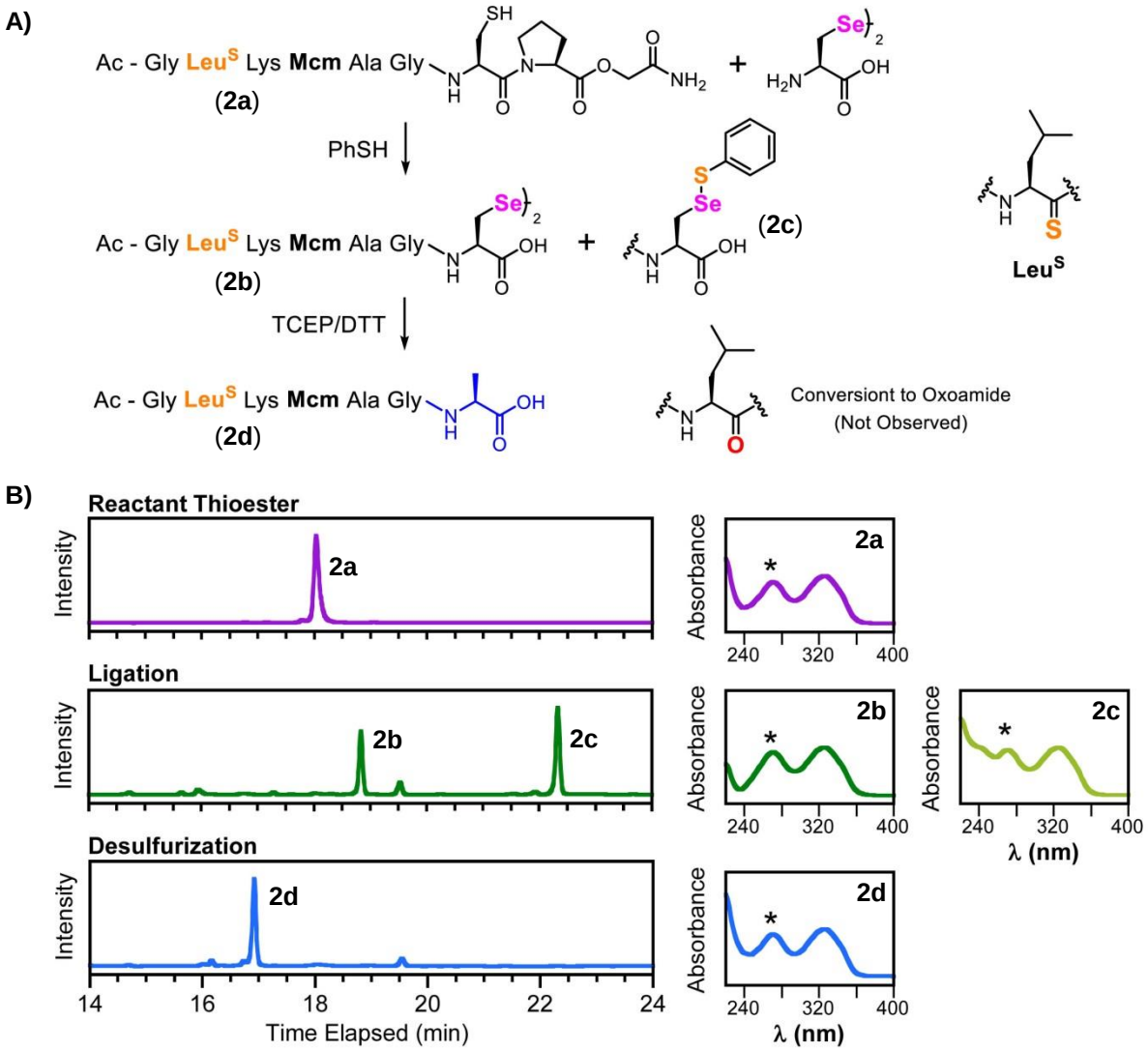


Figure 3-6. Selective Deselenization of Sec in the Presence of Thioamide.

A) Schematic representation of the test reaction; a reduced version of CPG_o instead of the tBuS-protected C^bPG_o was used to eliminate the need for TCEP as a reducing reagent in ligation. B) HPLC chromatogram monitored at 325 nm, and UV-Vis absorption spectra of isolated peptides. The star sign indicates thioamide absorption at 272 nm; **2c** also contained a peak around 260 nm that is characteristic of −SPh group. Ligation Conditions: 0.2 mM peptide **2a**, 2 mM selenocystine, sat. PhSH, 200 mM Na_2HPO_4, pH 8.0, 12 h. Deselenization Conditions: 0.2 mM ligated peptide (without purification), 20 mM TCEP, 20 mM DTT, 200 mM Na_2HPO_4, pH 8.0, overnight. MALDI MS: [**2a** + H]$^+$, expected 1005.42, found 1005.35; [**2b** + H]$^+$, expected 1795.55, found 1795.41; [**2c** + H]$^+$, expected 1007.29, found 1007.25; [**2d** + H]$^+$, expected 819.37, found 819.19. TCEP = tris(2-carboxyethyl) phosphine; DTT = dithiothreitol; Mcm = 7-methoxycoumarinylalanine.

concentrations of ambient radicals that are initiated by VA-044 and propagated by thiol additives to sever the C−S bond (Figure 3-7).

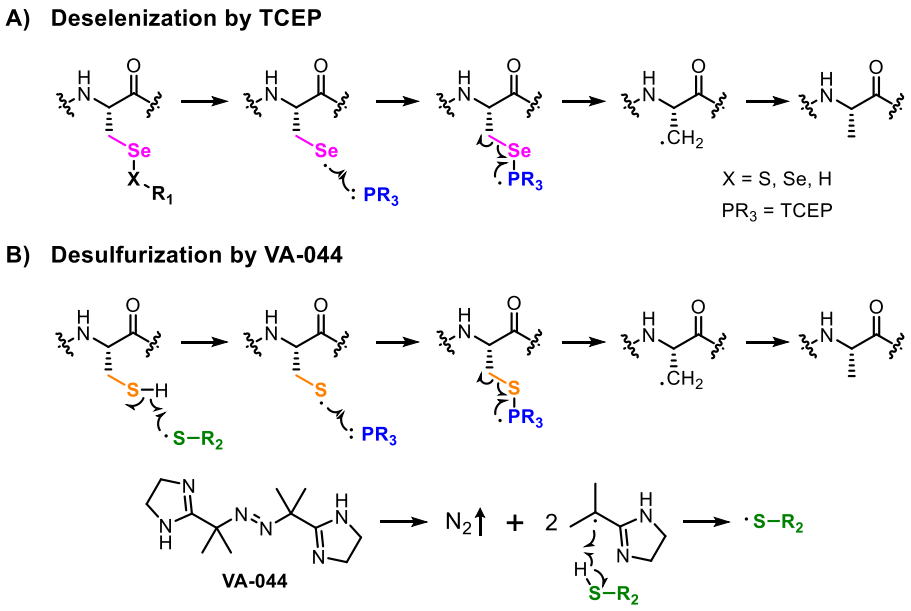


Figure 3-7. Proposed Mechanisms for TCEP Deselenization and VA-044 Desulfurization.

We carried out our initial test on peptide **3**, using an intermediate VA-044 concentration and short reaction time to minimize potential thioamide complications. We were excited to find that Cys can be selectively desulfurized in the presence of a thioamide, generating **3a** as the major product in 80% yield in 10 min (Figure 3-8). The nearly identical UV-Vis absorption profiles of isolated product as compared to the reactant demonstrated that the thioamide remained intact in Cys desulfurization (normalized $A_{272\ nm}$ = 0.48 for both product and reactant). These results gave us confidence that the thioamide was indeed compatible with the radical initiated desulfurization/deselenization methods. Given that β-thiol analogs are much easier to

handle than β-selenol compounds in terms of synthetic efforts and air-sensitivity, we decided to pursue VA-044 desulfurization as the primary method for traceless NCL incorporation of thioamides, and reserve TCEP deselenization for situations where additional selectivity against Cys is needed.

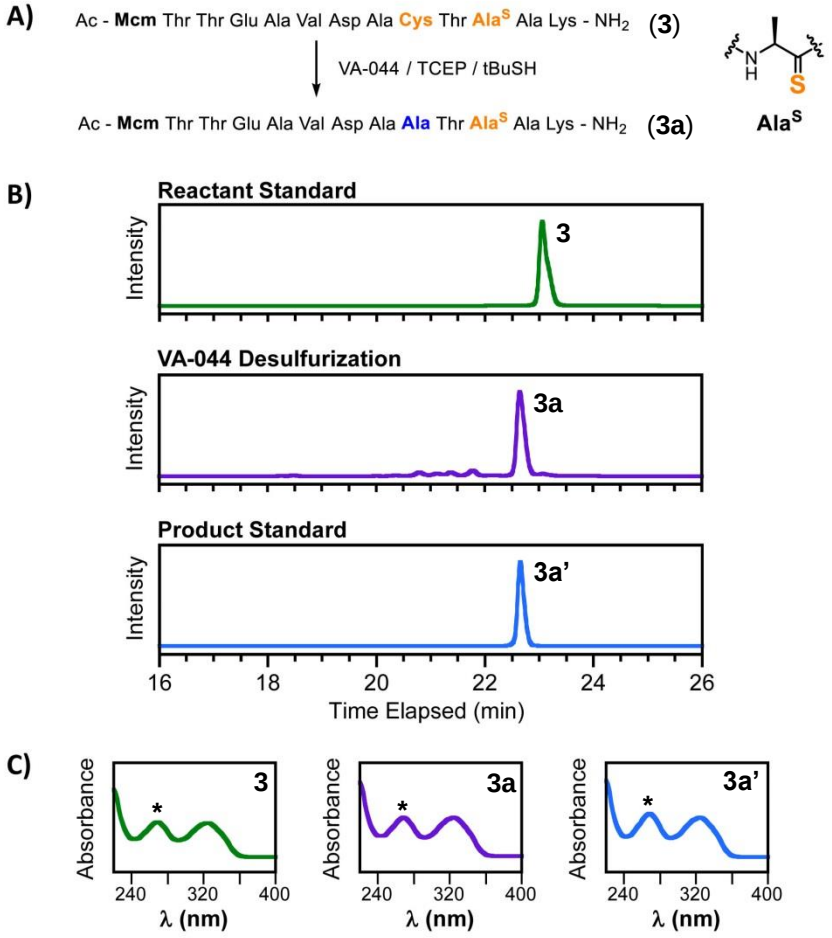


Figure 3-8. Selective Desulfurization of Cys in the Presence of Thioamide.

A) Schematic representation of the test reaction. B) HPLC chromatogram monitored at 325 nm. **3a'** denotes a genuine product standard synthesized *via* solid phase peptide synthesis (SPPS) for comparison. MALDI-TOF MS: [**3** + H]$^+$, expected 1482.62, found 1482.66; [**3a** + H]$^+$, expected 1450.64, found 1450.70; [**3a'** + H]$^+$, expected 1450.64, found 1450.74. C) UV-Vis absorption of isolated peptides. Star indicates thioamide absorption. Conditions: 0.1 mM peptide **3**, 10 mM VA-044, 40 mM TCEP, 200 mM Na_2HPO_4, 10% tBuSH (v/v) pH 7.0.

To optimize the VA-044 desulfurization reaction, we systematically varied individual components of the reaction to identify their optimal concentration. We found that high TCEP concentration was necessary to avoid disulfide bond formation (Figure 3-22), which was likely an off-pathway product from collisional quenching of thiol radicals (Figure 3-9). A high concentration of TCEP serves the dual function of ensuring the prompt scavenging of thiol radical for on-pathway desulfurization, and reducing the disulfide bonded side product to reverse the off-pathway reaction. We also found that the reaction was tolerant to a wide range of VA-044 concentrations, and no complications were observed with as high as 1000 equivalents of VA-044 (Figure 3-23). We reasoned that since the majority of reactive radicals were propagated from the thiol additive present at 10% (v/v), the reaction was relatively insensitive to VA-044 excess as long as it was above a critical concentration (empirically 10 equivalents for this test reaction). In addition, we showed that dissolved oxygen – which may function as collisional radical quencher – was well-tolerated (Figure 3-24). Only minor alterations in reaction kinetics were observed with no additional complications, eliminating the need to keep the reaction buffer degassed and making it easier for other research groups to adopt our methodology.

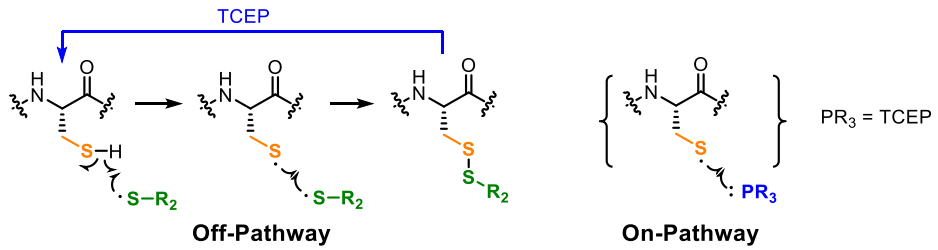


Figure 3-9. Proposed Mechanism for Off-Pathway Disulfide Bond Formation.

Upon a closer examination of the robustness of chemoselectivity against the thioamide, we observed that after prolonged exposure to the radical environment, a thioamide-to-oxoamide conversion side product **3b** gradually accumulated over time (Figure 3-26). Although **3b** only represented a 6% side reaction after 2 h, it became significant after 18 h (17% side product formation). We hypothesize that this is due to the amide-to-imine equilibrium, where the stable C=S bond is converted to a labile C−S bond as part of the transformation (Figure 3-10). Since this equilibrium is intrinsic to the thioamide bond, we turned to "suicide scavengers" in order to suppress the side reaction.

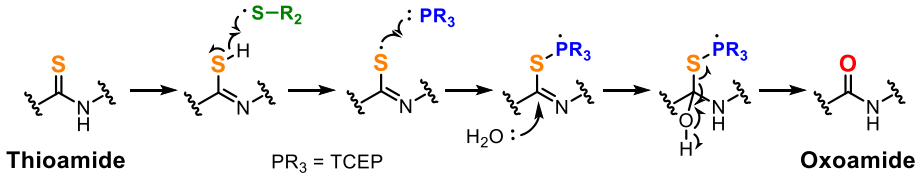


Figure 3-10. Proposed Mechanism for Radical Initiated Thioamide Desulfurization.

After screening a few thiocarbonyl-containing small molecules as candidates, we identified thioacetamide as the most effective "suicide scavenger". Its utility as a desulfurization chaperon can be clearly demonstrated on model peptide **3a'** (we use a prime symbol to denote the genuine peptide standards in order to distinguish them from species identified in desulfurization reactions) – when we treated **3a'** with VA-044 in the absence of thioamide, side product **3b** was generated at 16% by 18 h; in the presence of thioacetamide, **3a'** was fully stable for the same time frame (Figure 3-11B). When we applied this additive to the Cys/thioamide peptide **3**, the protective effect was equally profound – by suppressing thioamide side reaction, the yield of desired product **3a** was

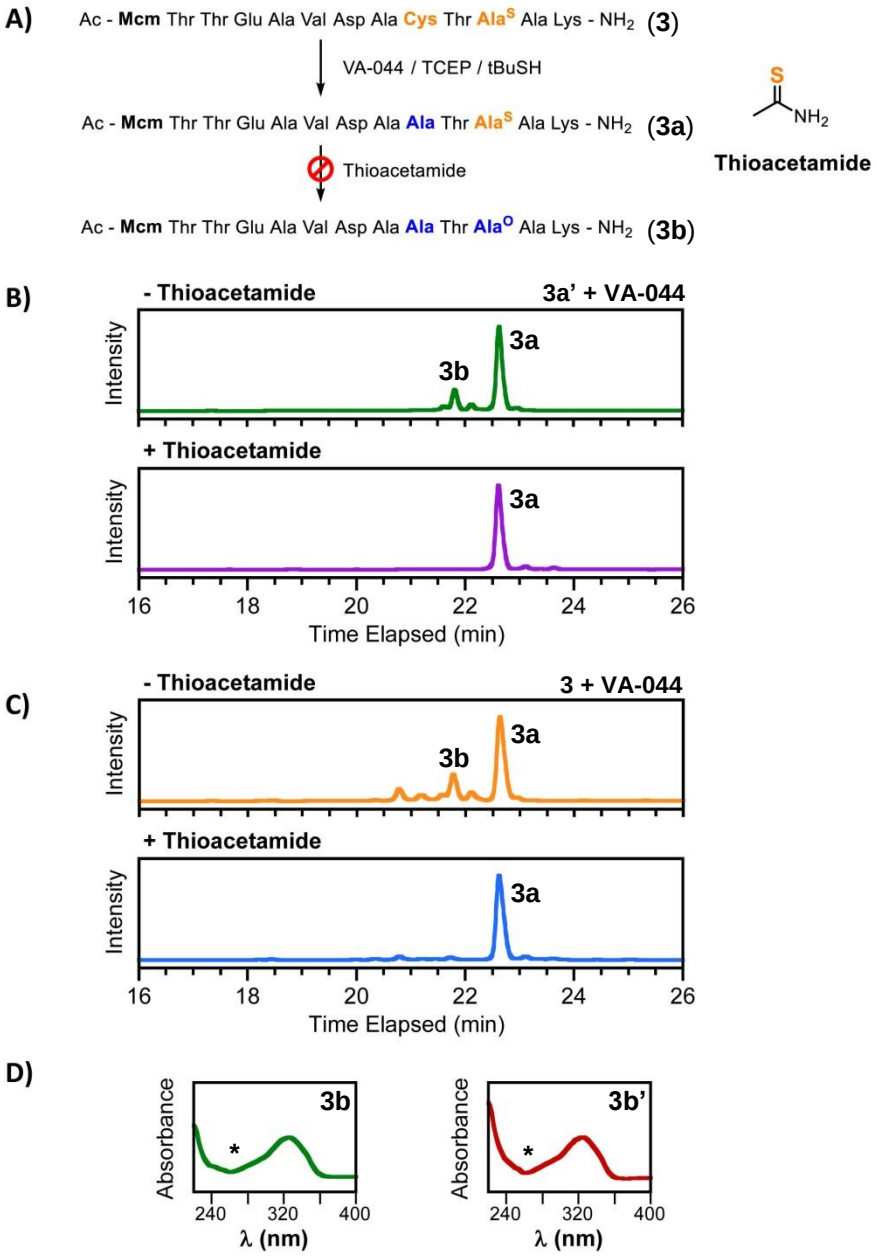


Figure 3-11. Thioacetamide as Chaperon to Suppress Thioamide Side Reaction.

A) Schematic representation of the test reaction. B) Treatment of control peptide **3a'** with VA-044 in the absence and presence of thioacetamide. HPLC chromatograms monitored at 325 nm. MALDI-TOF MS: [**3a** + H]$^+$, expected 1450.64, found 1450.74; [**3b** + H]$^+$, expected 1434.67, found 1434.87. C) Desulfurization of peptide **3** in absence and presence of thioacetamide. MALDI-TOF MS: [**3a** + H]$^+$, found 1450.46; [**3b** + H]$^+$, found 1434.44. D) UV-Vis absorption spectra of isolated **3b** compared to a genuine standard **3b'**. **3a'**, **3b'** denotes the genuine standards synthesized by solid phase peptide synthesis. Star sign indicates the absence of thioamide absorption at 272 nm. Conditions: 0.1 mM peptide **3a'** or **3**, 10 mM VA-044, 40 mM TCEP, 200 mM Na_2HPO_4, 10% tBuSH (v/v) pH 7.0, 18 h.

improved from 58% to 88% after an 18 h reaction (Figure 3-11C). It is worth noting that the residual peak at 21.8 min represented a 2% Cys-to-Ser conversion side reaction rather than thioamide-to-oxoamide conversion, which is rigo-rously characterized in *Materials and Methods* (Figure 3-30). As a further validation, we demonstrated that the supper-ssion was effective over a wide range of thioacetamide concentrations (Figure 3-27), and that the oxygen and denaturant tolerance of VA-044 desulfurization was unaltered by the addition of thioacetamide (Figure 3-28).

Having established a robust method for selective Cys desulfurization on thioamide-containing peptides, we next sought to investigate its application in traceless native chemical ligation (NCL) – joining two peptide fragments using Cys as a ligation handle, and subsequently "erasing" it into an Ala after ligation. There are two general ways to achieve ligation-desulfurization – with or without an intermediate purification step on the ligated Cys peptide. The latter is obviously advantageous in terms of purification efforts, but chemically more challenging as it requires all reagents used in ligation to be compa-tible with desulfurization. A particular reagent of concern is the aromatic thiol, typically thiophenol (PhSH) or 4-mercaptophenylacetic acid (MPAA), that is a common additive to accelerate ligation by converting alkyl thioesters into reactive aryl thioesters[136], but also a radical quencher that would sequester the reactivity of tBuS· radicals[294] (Figure 3-12).

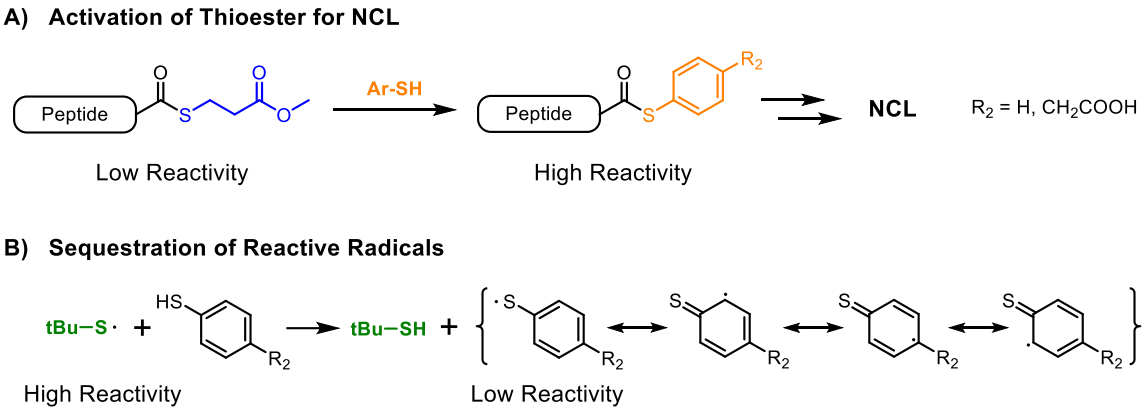


Figure 3-12. Effects of Aromatic Thiols on One-Pot Ligation Desulfurization.

To solve this problem, we attempted using CF_3CH_2SH as a non-aromatic substitute for aromatic thiols as advocated in a prior study from Payne's group;[294] unfortunately, this method did not yield successful ligation in our test reactions, presumably due to the high volatility of CF_3CH_2SH (boiling point 34 °C). Upon a closer examination of properties of the various thiols[295], we hypothesized – contrary to previous claims by other groups[294] – that PhSH can be effectively removed by lyophilization if the pH was kept sufficiently low to maintain it in the thiol-free form (Figure 3-13). The thiol-free residue can then be brought up in fresh buffer and desulfurized, without the need of additional purification steps for PhSH removal.

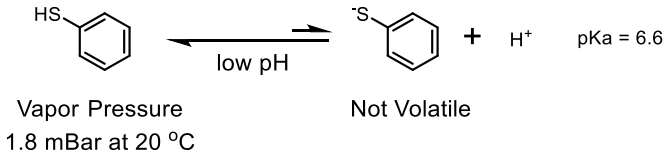


Figure 3-13. Schematic Representation of Thiol/Thiolate Equilibrium of PhSH.

See *Materials and Methods* for theoretical calculations of residual PhSH concentrations from a saturated PhSH solution after lyophilization at different pH.

Using thioester **4a** and Cys-containing peptide **4b**, we demonstrated that one-pot ligation-desulfurization can be successfully performed in the presence of aromatic thiol PhSH, by adopting the simple procedure of acidification and lyophilization (Figure 3-14). Both ligation product **4c** and desulfurization product **4d** perfectly matched the genuine peptide standards − **3** and **3a'**, respectively − in MALDI-TOF MS, UV-Vis absorption and HPLC retention time. Notably, the presence of 7-methoxycoumarin and thioamide chromophores deriving from the two fragments was clearly exemplified. With thio-acetamide present as a desulfurization chaperon, no thioamide-to-oxoamide conversion was observed.

With an effective strategy for one-pot ligation-desulfurization, one constraint still remains − VA-044 initiated desulfurization would indiscriminately desulfurize all Cys in the sequence, limiting its application to proteins without endogenous Cys. To remove this constraint, we turned to selective deselenization, where Sec can be selectively converted into Ala by TCEP while leaving Cys intact[216]. Having previously demonstrated the compatibility of the thioamide with deselenization (see Figure 3-6), we synthesized a Sec/Cys/thioamide-containing peptide **5**, and subjected it to the standard deselenization procedure. Much to our surprise, while the thioamide was unaffected throughout the reaction, both selective deselenization on Sec and non-selective desulfurization on Cys took place, giving a mixture of 55% Ala/Cys product **5a** and 12% Ala/Ala product **5b** (Figure 3-15).

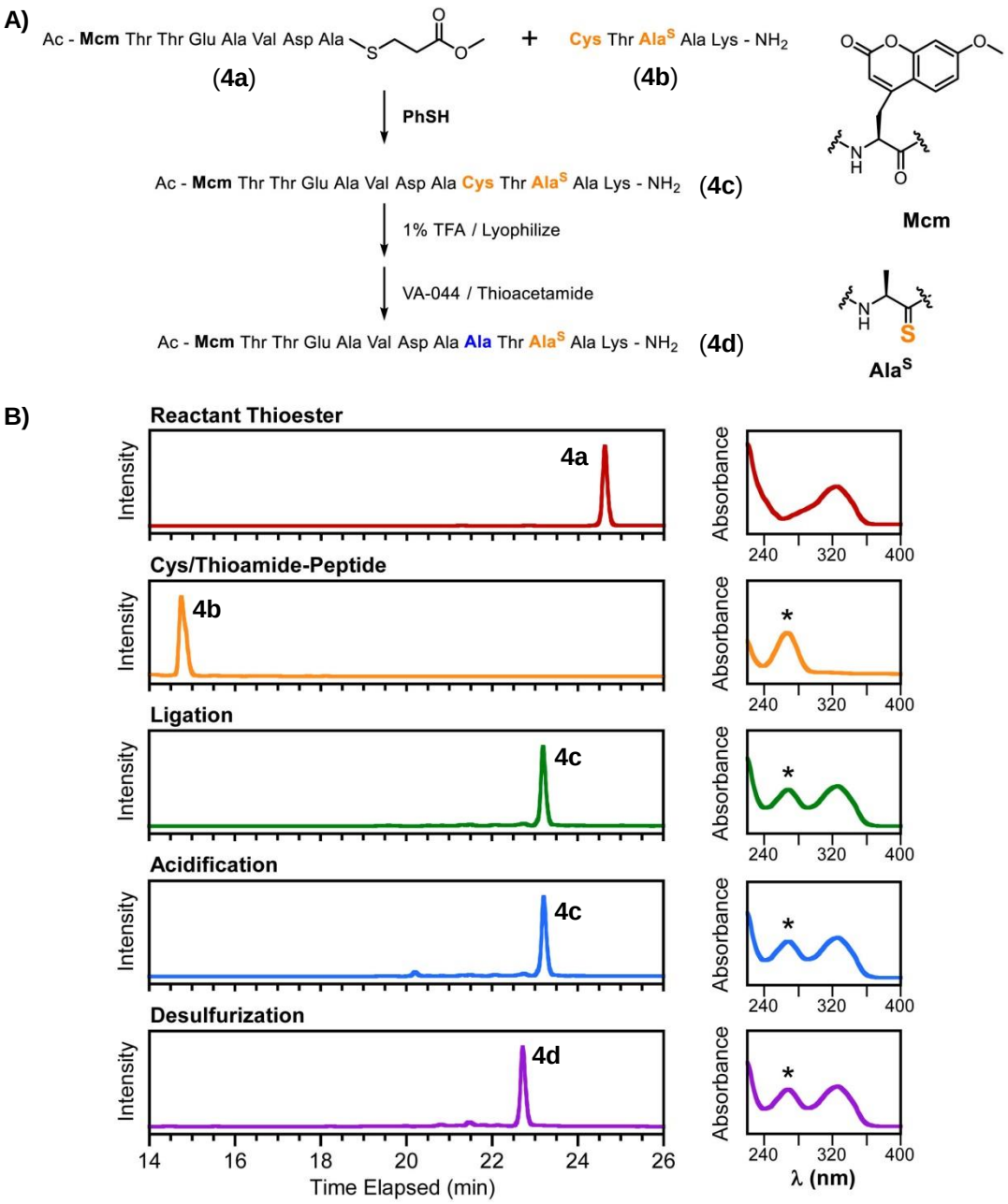


Figure 3-14. One-Pot Ligation-Desulfurization of Thioamide-Containing Peptides.

A) Schematic representation of the test reaction. B) HPLC chromatograms monitored at 325 nm (**4b** at 272 nm) and UV-Vis absorption spectra of major peak in each chromatogram. MALDI-TOF MS: [**4c** + H]$^+$, expected 1482.62, found 1482.47; [**4d** + H]$^+$, expected 1450.64, found 1450.82. Star sign indicates thioamide absorption at 272 nm. Ligation: 1 mM thioester **4a**, 1 mM peptide **4b**, 40 mM TCEP, 200 mM Na_2HPO_4, pH 7.0, overnight. Desulfurization: 0.1 mM reaction crude **4c**, 10 mM VA-044, 40 mM TCEP, 200 mM Na_2HPO_4, 10% tBuSH (v/v) pH 7.0, 2 h. See Figure 3-8 for genuine standards **3** and **3a'**. See *Materials and Methods* for residual PhSH characterization.

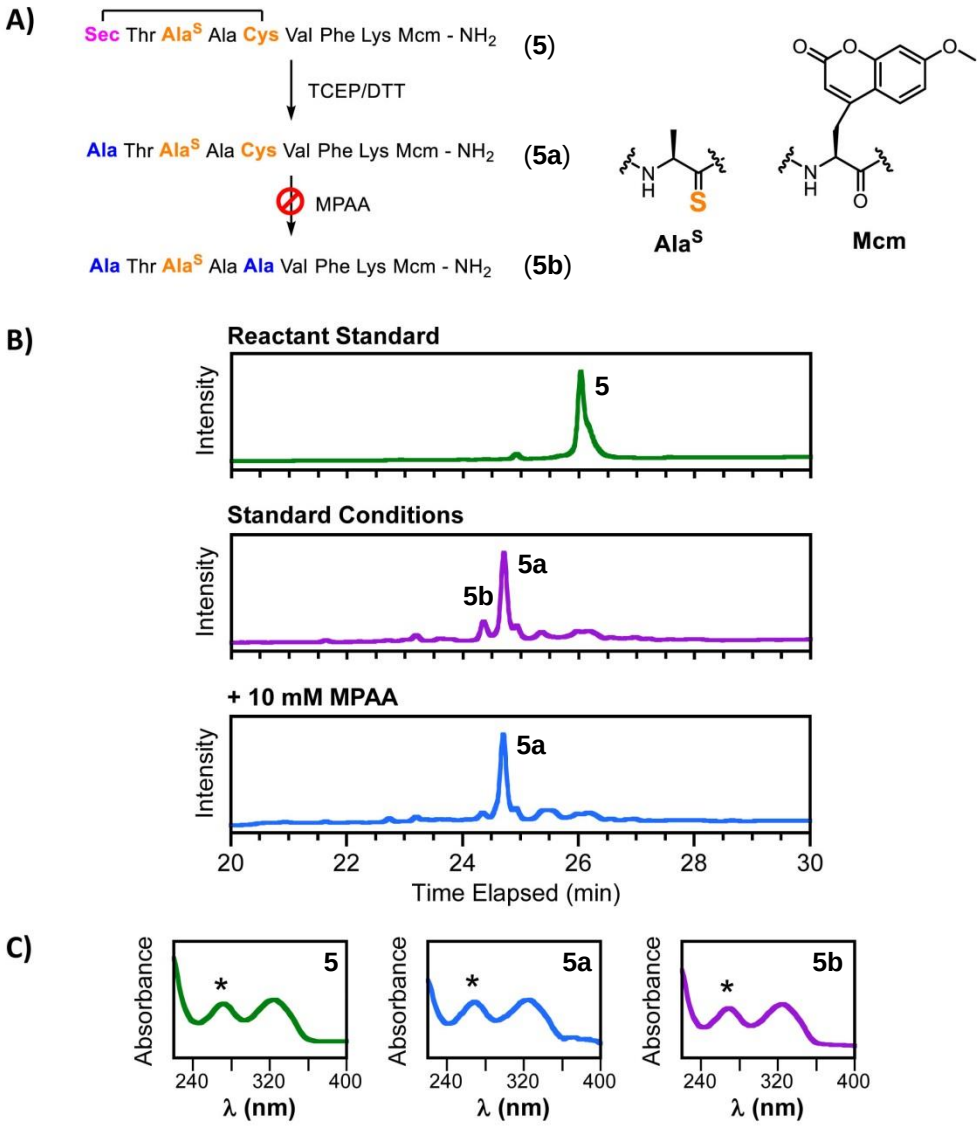


Figure 3-15. Selective Deselenization of Sec in the Presence of Cys and Thioamide.

A) Schematic representation of the test reaction. B) HPLC chromatograms monitored at 325 nm. MALDI-TOF MS: [**5** + H]$^+$, expected 1148.37, found 1148.56; [**5a** + H]$^+$, expected 1070.48, found 1070.74; [**5b** + H]$^+$, expected 1038.50, found 1038.73. C) UV-Vis absorption spectra of reactant and isolated products. Star sign indicates thioamide absorption at 272 nm. Conditions: 0.1 mM peptide **5**, 40 mM TCEP, 40 mM DTT, with or without 10 mM MPAA, 200 mM Na_2HPO_4, pH 7.0, 18 h. TCEP = tris(2-carboxyethyl) phosphine; DTT = dithiothreitol; MPAA = 4-mercaptophenylacetic acid; Mcm = 7-methoxycoumarinylalanine.

After examining the reaction mechanism, we realized that the side reaction originates from homolytic cleavage of Se−S bond as the first step of the deselenization reaction, which also generates a sulfur radical that can proceed with desulfurization (Figure 3-16). Since these intramolecular Se−S bonds formed spontaneously during peptide synthesis and purification (in fact, we used hemiselenide protected Boc-Sec(S-iPr)-OH as the building block, but could only isolate the intramolelar hemiselenide **5** as product), we would need to develop a strategy to suppress the sulfur radical. Knowing that an aromatic thiol is tolerated in one-pot ligation-deselenization (see Figure 3-6) but not in desulfurization[294,296], we hypothesized that aromatic thiol may be able to selectively quench R−S· radicals but not R−Se· radicals, which is theoretically favorable based on a comparison of the bond dissociation energy of Se−H bond (66 kcal mol^{-1}) and S−H bond (87 kcal mol^{-1})[297] (Figure 3-17). If this is true, we should be able to selectively suppress desulfurization side reaction using aromatic thiols, while allowing desired deselenization to proceed without complications. Indeed, using MPAA as an aromatic thiol additive, we were able to effectively suppress desulfurization, giving the desired product **5a** in 62% yield (Figure 3-15). While more in-depth mechanistic studies are certainly necessary to fully characterize the reactivity of the various radicals involved, we have provided an effective empirical method for the synthesis of thioamide-containing proteins that also contain native Cys residues.

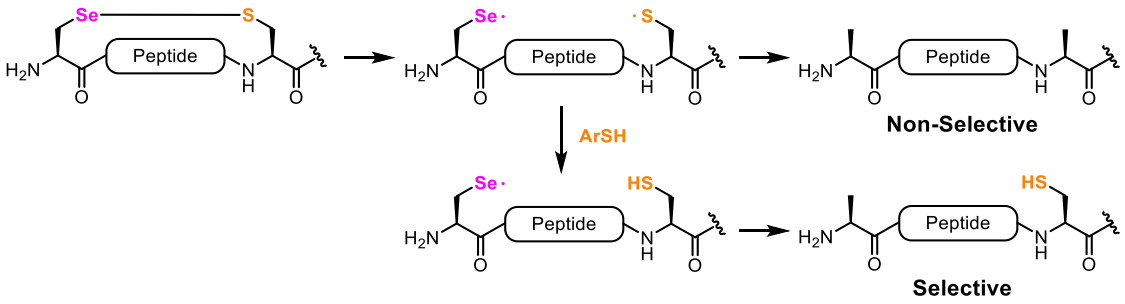


Figure 3-16. Proposed Mechanism for Non-Selective Desulfurization.

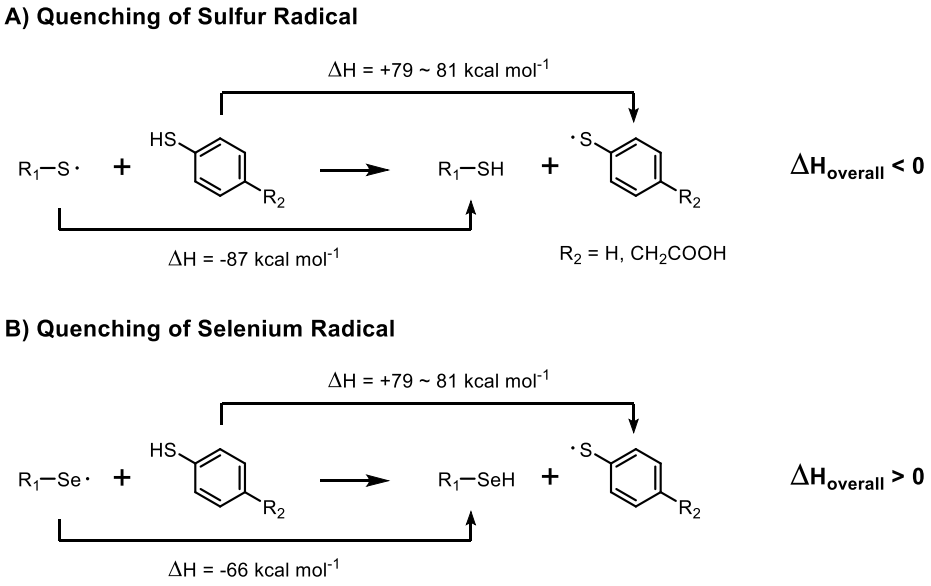


Figure 3-17. Thermodynamics Prediction of Sulfur and Selenium Radical Quenching.

Bond dissociation energy for aromatic thiols differs by substitution[296], so a range is given instead of a single value; this difference does not alter the overall predicted trends.

Finally, to demonstrate the possibility of expanding traceless NCL into non-Cys/Sec residues, we performed a proof-of-concept desulfurization reaction with a Pen-containing peptide **6** (Pen = penicillamine). Pen is the β-thiol analog of Val; similar to Cys, it can be

used as a ligation handle and subsequently "erased" into the native Val[202]. We chose Pen because it is a most sterically demanding member among all β-thiol analogs; the ambient radicals will need to access a quantanery carbon center than is one bond away from the peptide backbone in order to intiate desulfurization. Upon treatment with VA-044 in the presence of thioacetamide, **6** was converted to desired **6a** in 73% yield, with its thioamide intact (Figure 3-18). This clearly showed the feasibility of β-thiol and β-selenol analogs for thioamide incorporation *via* traceless NCL, lifting the last restriction on semi-synthesis of thioamide-containing peptides and proteins.

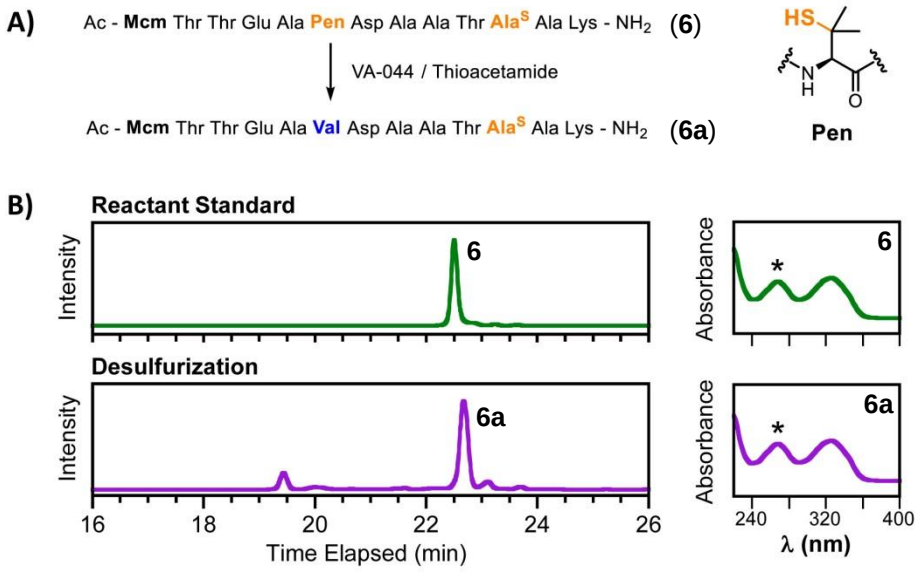


Figure 3-18. Selective Deselenization of Sec in the Presence of Cys and Thioamide.

A) Schematic representation of the test reaction. B) HPLC chromatograms monitored at 325 nm, and UV-Vis absorption of isolated peptides. MALDI-TOF MS: [**6** + H]$^+$, expected 1482.62, found 1482.56; [**6a** + H]$^+$, expected 1450.64, found 1450.82. Retention time: 22.4 min for **6**, and 22.7 min for **6a**; the difference is small but reproducible; identities of the peptides were confirmed by mass spectrometry. Star sign indicates thioamide absorption at 272 nm. Conditions: 0.1 mM peptide **6**, 10 mM VA-044, 40 mM TCEP, 200 mM Na_2HPO_4, 10% tBuSH (v/v), pH 7.0, 18 h. TCEP = tris(2-carboxyethyl) phosphine; Mcm = 7-methoxycoumarinyl-alanine. With D. Miklos Szantai-Kis.

3.3 Conclusion

We have devised a method for the traceless incorporation of thioamide into peptides and proteins using chemoselective desulfurization/deselenization in combination with native chemical ligation (NCL). We demonstrated that Cys can be selectively converted into Ala using VA-044 as a radical initiator in the presence of a backbone thioamide, and discovered that a small molecule, thioacetamide, can be used as an additive to improve the robustness of this chemoselectivity. We also showed that by acidifying the ligation mixture, PhSH can be effectively removed by lyophilization, enabling one-pot ligation-desulfurization that had not previously been achieved. For sequences with additional native Cys, we showed the utility of selective deselenization, where Sec was converted into Ala in the presence of Cys and thioamides, and identified aromatic thiols as additives to enhance the chemo-selectivity. To expand into ligation sites other than Cys/Sec, we demonstrated selective desulfurization of β-thiol analogs, using conversion of penicillamine into Val as a proof-of-concept example. In conclusion, we have greatly expanded the scope of thioamide incorporation − it is now possible to achieve NCL at nearly any amino acid site in a target protein, which should greatly advance the adoption of the thioamide as a minimalist probe.

3.4 Future Directions

With the methodology for thioamide incorporation in place, we are now capable of performing “positional scanning” experiments, where a combinatorial library of target

proteins can be generated, with thioamide and fluorophores at various positions. Specifically, there are 15 good ligation sites in α-synuclein, where the junction is between a small amino acid Gly or Ala and an "erasable" Ala or Val handle (Figure 3-19). This would not only enable us to incorporate a thioamide at the N- or C-terminal region, but also allow us to perform a tandem ligation to incorporate thioamide directly into the fibril core (residues 65 to 95) without worrying about the perturbation of residual Cys.

```
1      6      11     16      21     26     31
MDVFM  KGLSK  AKEG|V VA|A|AE KTKQG |VAEA|A GKTKE

35     41     46     51      56     61     66
G|VLYV GSKTK  EG|VVH G|VATV  AEKTK  EQVTN  VGG|AV

71     76     81     86      91     96     101
VTG|VT A|VAQK TVEG|A GSIA|A |ATGFV KKDQL  GKNEE

105    111    116    121     126    131    136
G|APQE GILED  MPVDP  DNEAY   EMPSE  EGYQD  YEPEA
```

Figure 3-19. α-Synuclein Sequence and Potential Traceless NCL Sites.

Of particular interest would be tandem ligation between $\alpha S_{1\text{-}66}$, $\alpha S_{67\text{-}89}$ and $\alpha S_{90\text{-}140}$. We can synthesize the middle fragment $\alpha S_{67\text{-}89}$ using standard SPPS, with thioamides at any Val, Ala, Glu, Gln or Lys sites. The C-terminal fragment $\alpha S_{90\text{-}140}$ would contain a fluorophore at F94 position, which has been previously shown to be at a rigid position just outside the fibril core[47]. By generating fibrils using these doubly labeled proteins, we can extrapolate thioamide-fluorophore distances from fluorescence spectroscopy both in the mature fibril and during the aggregation process, and start to decode the biophysical process that is the very basis of Parkinson's disease.

3.5 Materials and Methods

General Information VA-044 was purchased from Wako Pure Chemical Industries (Osaka, Japan). Fmoc protected amino acids and peptide synthesis reagents were purchased from EMD Millipore (Billerica, MA). All other reagents and solvents were purchased from Fisher Scientific (Waltham, MA) or Sigma-Aldrich (St. Louis, MO) unless otherwise specified. High resolution electrospray ionization mass spectra (ESI-HRMS) were collected with a Waters LCT Premier XE liquid chromatograph/mass spectrometer (Milford, MA). Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer. UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (Agilent Technologies, Santa Clara, CA). Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA). Matrix assisted laser desorption/ionization with time-of-flight detector (MALDI-TOF) mass spectra were acquired on a Bruker Ultraflex III instrument (Billerica, MA). Analytical HPLC was performed on an Agilent 1100 Series HPLC system (Santa Clara, CA). Preparative HPLC was performed on a Varian Prostar HPLC system (Agilent Technologies, Santa Clara, CA). HPLC columns were purchased from W. R. Grace & Compnay (Columbia, MD).

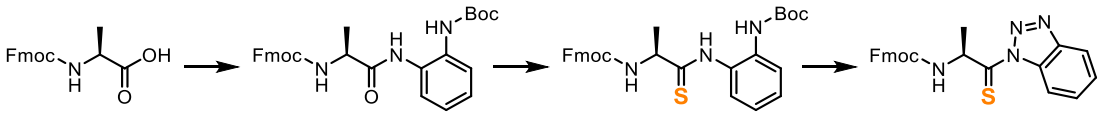


Figure 3-20. Synthesis Scheme of Thioalanine Precursors for Thioamide Incorporation.

Synthesis of *N*-Boc-1,2-phenylenediamine (S1) 1,2-Phenylenediamine (5.9478 g, 55 mmol, 1.1 equiv) was dissolved in anhydrous tetrahydrofuran (THF, 50 mL). A solution of of di-*tert*-butyldicarbonate (Boc_2O, 10.9125 g, 50 mmol, 1 equiv) in anhydrous THF (50 mL) was added dropwise over 4 h on ice. Upon completion, the reaction was quenched by slow addition of H_2O (100 mL), and extracted with ethyl acetate (100 mL×3). The organic layers were combined, dried by rotary evaporation, and then purified by flash chromatography with 2:8 ethyl acetate/petroleum ether. A white flake was isolated as the final product (6.9823 g, 33.5 mmol, 67% yield). R_f 0.18 in 2:8 ethyl acetate/petroleum ether. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 7.27 (d, J = 7.6 Hz, 1H), 6.99 (t, J = 8.3 Hz, 1H), 6.77 (t, J = 7.6 Hz, 1H), 6.71 (d, J = 8.1 Hz, 1H), 6.60 (s, 1H), 3.76 (s, 2H), 1.54 (s, 9H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 154.16, 140.23, 126.17, 124.82, 119.44, 117.53, 80.48, 28.48. ESI$^+$-HRMS: calculated for $C_{11}H_{17}N_2O_2^+$: 209.12; found $[M + H]^+$: 209.14.

Synthesis of α-*N*-Fmoc-L-alanine-(*N*-Boc)-2-aminoanilide (S2). In an argon atmosphere, Fmoc-Ala-OH (1.5567 g, 5mmol, 1 equiv) was dissolved in 50 mL of dry tetrahydrofuran (THF) and chilled in a 1:3 NaCl : ice bath (-10 °C). With stirring, N-methylmorpholine (NMM, 1.10 mL, 10 mmol, 2 equiv) and isobutylchloroformate (IBCF, 0.65 mL, 5 mmol, 1 equiv) were slowly added. The reaction was stirred at -10°C for 15 min, and *N*-Boc-1,2-phenylenediamine (**S1**, 1.0413g, 5 mmol, 1 equiv) was added. The reaction was allowed to proceed at -10°C for 2 h, and then at room temperature overnight (≥ 14 h). Upon completion, the solvent was removed by rotary evaporation.

The residue was brought up in 40 mL of ethyl acetate, and extracted against 40 mL each of 1 M Na_2HPO_4, brine, 5% $NaHCO_3$, and brine. The aqueous layers were combined and extracted against 100 mL of ethyl acetate. The organic layers were combined, dried by rotary evaporation, and purified by flash chromatography in ethyl acetate/petroleum ether (35:65 v/v). A pale yellow solid was yielded as the final product (1.7944 g, 3.6 mmol, 77% yield). R_f 0.23 in 35:65 ethyl acetate/petroleum ether. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 8.73 (s, 1H), 7.70 (d, *J* = 7.6 Hz, 2H), 7.52 (t, *J* = 8.0 Hz, 2H), 7.48 (d, *J* = 7.4 Hz, 1H), 7.33 (t, *J* = 8.0 Hz, 2H), 7.31 (m, 1H), 7.22 (m, 2H), 7.12 (s, 1H), 7.10 (t, *J* = 7.5 Hz, 1H), 7.01 (t, *J* = 7.6 Hz, 1H), 5.87 (d, *J* = 7.1 Hz, 1H), 4.38 (m, 1H), 4.35 (t, *J* = 6.4 Hz, 2H), 4.14 (t, *J* = 6.7 Hz, 1H), 1.44 (s, 9H), 1.41 (d, *J* = 7.1 Hz, 3H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 171.83, 156.25, 154.11, 143.72, 141.28, 131.39, 128.81, 127.79, 127.11, 126.57, 125.40, 125.08, 124.20, 80.75, 68.62, 51.08, 47.09, 28.32, 22.11. ESI^+-HRMS: calculated for $C_{29}H_{31}N_3O_5Na^+$: 524.2161; found $[M + Na]^+$: 524.2148.

Synthesis of *α*-*N*-Fmoc-L-thioalanine-(*N*-Boc)-2-aminoanilide (**S3**). In an argon atmosphere, **S2** (1.6593 g, 3.3 mmol, 1 equiv) and Lawesson's reagent (1.0036 g, 2.5 mmol, 0.75 equiv) were refluxed in 30 mL of CH_2Cl_2 for 18 h. Upon completion, the solvent was removed by rotary evaporation. The crude product was purified by flash chromatography in ethyl acetate/petroleum ether (35:65 v/v) to yield a pale yellow foam (1.5214 g, 2.9 mmol, 89% yield). R_f 0.43 in 35:65 ethyl acetate/petroleum ether. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 10.09 (s, 1H), 7.69 (d, *J* = 7.3 Hz, 2H), 7.61 (d, *J* = 7.4 Hz, 1H), 7.51 (d, *J* = 7.6 Hz,1H), 7.49 (d, *J* = 7.1 Hz,1H), 7.33 (t, *J* = 7.4 Hz, 2H), 7.30 (m, 1H), 7.23 (m, 2H), 7.22 (m, 1H), 7.05 (t, *J* = 7.6 Hz, 1H), 6.91 (s, 1H), 6.04 (d, *J* = 6.4

Hz, 1H), 4.67 (m, 1H), 4.31 (d, *J* = 6.1 Hz, 2H), 4.14 (t, *J* = 7.0 Hz, 1H), 1.50 (d, *J* = 6.6 Hz, 3H), 1.42 (s, 9H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 205.90, 155.94, 153.83, 143.61, 143.56, 141,22, 141.20, 132.95, 129.64, 128.49, 127.78, 127.14, 125.11, 125.07, 124.65, 123.59, 120.01, 81.07, 67.27, 55.58, 46.99, 28.26, 22.42. ESI^+-HRMS: calculated for $C_{29}H_{31}N_3O_4SNa^+$: 540.1933; found $[M + Na]^+$: 540.1914.

Synthesis of α-*N*-Fmoc-L-thioalanine-benzotriazolide (**S4**) In argon atmosphere, **S3** (1.0861 g, 2.1 mmol, 1 equiv) was dissolved in 10 mL of CH_2Cl_2 and chilled to 0 ˚C on ice. 10 mL of an ice cold TFA/CH_2Cl_2 (50:50 v/v) solution was slowly added, and the reaction was allowed to proceed at 0 ˚C for 2 h. Upon completion, the solvent was removed by rotary evaporation. The residue was dissolved in glacial AcOH (9.5 mL) and water (0.5 mL), and then chilled to 0 ˚C on ice. With stirring, $NaNO_2$ (0.2202 g, 3.2 mmol, 1.5 equiv) was added in small portions over 5 min. The reaction was allowed to proceed at 0 ˚C for 30 min, and then quenched with 30 mL of ice cold water. The precipitate was collected by filtration, dried on vacuum for 30 min, and then purified by flash chromatography in ethyl acetate/petroleum ether (35:65 v/v). A yellow foam was yielded as the final product (0.6286 g, 1.5 mmol, 69% yield). R_f 0.54 in 35:65 ethyl acetate/petroleum ether. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 8.75 (d, *J* = 8.3 Hz, 1H), 8.12 (d, *J* = 8.0 Hz, 1H), 7.74 (d, *J* = 7.1 Hz, 2H), 7.64 (m, 2H), 7.61 (m, 1H), 7.52 (t, *J* = 7.4 Hz, 1H), 7.38 (t, *J* = 6.7 Hz, 1H), 7.30 (t, *J* = 6.8 Hz, 1H), 6.31 (m, 1H), 5.91 (d, *J* = 8.8 Hz, 1H), 4.47 (dd, *J* = 10.3, 7.0 Hz, 1H), 4.34 (dd, *J* = 10.3, 7.0 Hz, 1H), 4.22 (t, *J* = 6.6 Hz, 1H), 1.66 (d, *J* = 6.6 Hz, 3H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 210.01, 155.61, 147.18, 144.03, 143.88, 141.41, 132.46, 131.68, 127.83, 127.32, 127.20, 125.32, 125.22,

120.92, 120.11, 116.31, 67.16, 57.27, 47.33, 23.10. ESI$^+$-HRMS: calculated for $C_{24}H_{20}N_4O_2SNa^+$: 451.1205; found $[M + Na]^+$: 451.1195.

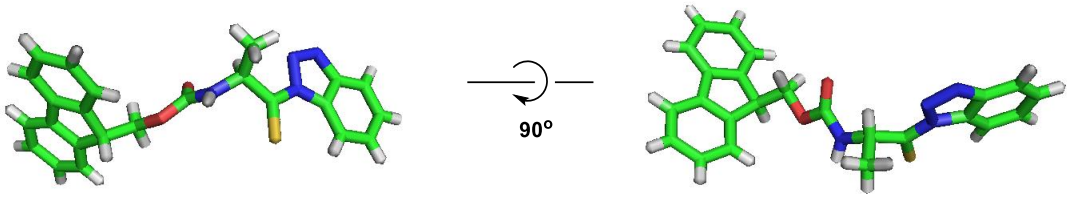


Figure 3-21. Crystal Structure of Thioalanine Precursor.

Crystal was grown in 35:65 ethyl acetate/petroleum ether; data were acquired on a Bruker Kappa APEX II Duo CCD diffractometer (Billerica, MA). Graphics were generated using Schrödinger PyMOL (Cambridge, MA). C=S bond length observed: 1.6 Å.

Peptide Synthesis and Purification Peptides were synthesized using standard Fmoc solid phase peptide synthesis (SPPS) procedure on either Rink Amide or 2-chlorotrityl chloride resin (100 - 200 mesh; 0.6 mmol substitution/g). For coupling, 5 equiv of amino acid and 5 equiv of HBTU were dissolved in dimethylforamide (DMF), pre-activated for 1 min in the presence of 10 equiv of *N*,*N*-diisopropylethylamine (DIPEA), and then stirred with the resin for 30 min at room temperature. For deprotection, 20% piperidine in DMF was stirred with the resin for 20 min. Thioalanine (denoted Ala^S or A^S) was introduced through activated benzotriazole precursors (**S4**); 2 equiv of the precursor was dissolved in dry CH_2Cl_2, and stirred with the resin for 45 min in the presence of 2 equiv DIPEA. For N-terminally acetylated peptides, an acetylation cocktail of 1:1:8 acetic anhydride/N-methylmorpholine/DMF was stirred with the resin twice, each for 10 min, after the last deprotection. Upon completion of SPPS, resins were rinsed thoroughly with CH_2Cl_2 and dried under vacuum.

For cleavage, resins were treated with a cleavage cocktail (12:1:1:26 trifluoroacetic acid/triisopropylsilane/H_2O/CH_2Cl_2) for 30 min. The solution was then collected by filtration, and dried by rotary evaporation. For purification, the crude residues were brought up in 1:10 CH_3CN/H_2O, and then purified by reverse phase HPLC with acidified (with 0.1% trifluoroacetic acid) CH_3CN/H_2O gradients. Individual fractions were characterized by MALDI-TOF MS, and dried by lyophilization. When necessary, the isolated products were subjected to multiple rounds of purification until 99.5% pure by MALDI-TOF MS and analytical HPLC. Solvent radients, retention times and MALDI-TOF MS results are listed in Table 3-1 through Table 3-3.

Table 3-1. Peptide Purification Methods and Retention Time.

Peptide	Gradient	Retention Time	Column
Ac-L**A**^{S}AKAG**C**AKXAG-NH_2 (**1**)	**1**	23.1 min	Vydac C18 Semi-prep
Ac-L**A**^{S}AKAG**A**AKXAG-NH_2 (**1b**)	**1**	22.3 min	Vydac C18 Semi-prep
Ac-G**L**^{S}KXAG-CPG_o (**2a**)	**4**	22.0 min	YMC-Pack Pro C8 Semi-prep
Ac-XTTEAVDA**C**T**A**^{S}AK-NH_2 (**3**)	**5**	16.0 min	Vydac C18 Semi-prep
Ac-XTTEAVDA**A**T**A**^{S}AK-NH_2 (**3a'**)	**6**	21.6 min	Vydac C18 Semi-prep
Ac-XTTEAVDA**A**T**A**AK-NH_2 (**3b'**)	**6**	19.9 min	Vydac C18 Semi-prep
Ac-XTTEAVDA**S**T**A**^{S}AK-NH_2 (**S7'**)	**6**	19.7 min	Vydac C18 Semi-prep
Ac-XTTEAVDA**C**T**A**AK-NH_2 (**S8'**)	**6**	21.2 min	Vydac C18 Semi-prep
Ac-XTTEAVDA-$SC_2H_4COOCH_3$ (**4a**)	**7**	17.6 min	YMC-Pack Pro C8 Semi-prep
CT**A**^{S}AK-NH_2 (**4b**)	**8**	12.8 min	Vydac C18 Semi-prep
UT**A**^{S}A**C**VFKX-NH_2^{Se-S} (**5**)**	**9**	24.5 min	YMC-Pack Pro C8 Semi-prep
Ac-XTTEA**Y**DAATA^{S}AK-NH_2 (**6**)	**10**	34.6 min	Vydac C18 Prep

* X = 7-methoxycoumarinylalanine; Y = penicillamine; U = selenocysteine; G_o = glycolic acid.

** The "Se-S" superscript denotes that **5** was isolated as an intramolecular hemiselenide species.

Table 3-2. HPLC Gradients for Peptide Purification and Characterization.

No.	Time (min)	%B	**No.**	Time (min)	**%B**	**No.**	Time (min)	**%B**
1	0:00	2	**2**	0:00	2	**3**	0:00	2
	5:00	2		5:00	2		8:00	2
	10:00	15		10:00	20		10:00	10
	25:00	30		25:00	50		30:00	50
	30:00	100		30:00	100		35:00	100
	35:00	100		35:00	100		40:00	100
	40:00	2		40:00	2		45:00	2
4	0:00	2	**5**	0:00	2	**6**	0:00	2
	10:00	2		5:00	2		5:00	2
	15:00	20		10:00	25		10:00	20
	30:00	50		25:00	40		25:00	35
	35:00	100		30:00	100		30:00	100
	40:00	100		35:00	100		35:00	100
	45:00	2		40:00	2		40:00	2
7	0:00	2	**8**	0:00	2	**9**	0:00	2
	5:00	2		5:00	2		5:00	2
	10:00	25		10:00	10		10:00	20
	30:00	45		20:00	20		40:00	50
	35:00	100		25:00	100		45:00	100
	40:00	100		30:00	100		50:00	100
	45:00	2		35:00	2		55:00	2
10	0:00	5						
	5:00	5						
	15:00	22						
	45:00	28						
	50:00	100						
	55:00	100						
	60:00	5						
	60:00	5						

* Solvent A: 0.1% trifluoroacetic acid in water; Solvent B: 0.1% trifluoroacetic acid in acetonitrile

Table 3-3. MALDI-TOF MS Characterization of Purified Peptides.

Peptide	**$[M + H]^+$**		**$[M + Na]^+$**	
	Calc'd	**Found**	**Calc'd**	**Found**
Ac-L**A**^{S}AKAG**C**AKXAG-NH_2 (**1**)	1262.59	1262.63	1282.58	1284.62
Ac-L**A**^{S}AKAG**A**AKXAG-NH_2 (**1b**)	1230.62	1230.78	1252.61	1252.77
Ac-G**L**^{S}KXAG-CPG_o (**2a**)	1005.41	1005.35	1027.40	1027.33
Ac-XTTEAVDA**C**T**A**^{S}AK-NH_2 (**3**)	1482.62	1482.42	1504.61	1504.41

Table 3-3 (cont'd). MALDI-TOF MS Characterization of Purified Peptides.

Peptide	$[M + H]^+$		$[M + Na]^+$	
	Calc'd	**Found**	**Calc'd**	**Found**
Ac-L**A**^{S}AKAG**C**AKXAG-NH_2 (**1**)	1262.59	1262.63	1282.58	1284.62
Ac-L**A**^{S}AKAG**A**AKXAG-NH_2 (**1b**)	1230.62	1230.78	1252.61	1252.77
Ac-G**L**^{S}KXAG-CPG_o (**2a**)	1005.41	1005.35	1027.40	1027.33
Ac-XTTEAVDA**C**T**A**^{S}AK-NH_2 (**3**)	1482.62	1482.42	1504.61	1504.41
Ac-XTTEAVDA**A**T**A**^{S}AK-NH_2 (**3a'**)	1450.64	1450.80	1472.63	1472.79
Ac-XTTEAVDA**A**T**A**AK-NH_2 (**3b'**)	1434.67	1434.81	1456.66	1456.80
Ac-XTTEAVDA**S**T**A**^{S}AK-NH_2 (**S7'**)	1466.64	1466.74	1488.63	1466.72
Ac-XTTEAVDA**C**T**A**AK-NH_2 (**S8'**)	1466.64	1466.80	1488.63	1466.81
Ac-XTTEAVDA-$SC_2H_4COOCH_3$ (**4a**)	1095.41	-	1117.40	1117.45
CT**A**^{S}AK-NH_2 (**4b**)	508.23	508.37	530.22	530.34
UT**A**^{S}A**C**VFKX-NH_2^{Se-S} (**5**)**	1148.37	1148.46	1170.36	1170.44
Ac-XTTEA**Y**DAAT**A**^{S}AK-NH_2 (**6**)	1482.62	1482.56	1504.61	1504.55

* X = 7-methoxycoumarinylalanine; Y = penicillamine; U = selenocysteine; G_o = glycolic acid.
** **5** was isolated as an intramolecular hemiselenide species.

Raney Nickel Desulfurization Peptide **1** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in desulfurization buffer (100 µL, 100 mM Na_2HPO_4, 10 mM TCEP, pH 5.8), and then Raney nickel (0.1 mg) was added. The reaction was allowed to proceed at room temperature for 12 h, and then supernatant was recovered by centrifugation at 13.2 krpm for 20 min. The crude was analyzed by analytical RP-HPLC on a Luna C8 analytical column using gradient **2**. Individual fractions were collected manually, and then analyzed by MALDI-TOF and UV-Vis absorption spectroscopy.

Synthesis of Reduced Ac-GL^{S}KXAG-CPG$_o$ (2a) Ac-GL^{S}KXAG-C^{b}PG$_o$ (**S5**) with a side chain tBuS− protecting group was synthesized as we described previously[283]. To

generate the reduced **2a**, **S5** (200 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was mixed with 100 µL of reduction buffer (40 mM TCEP, 200 mM Na_2HPO_4, pH 6.8) and 20 µL of CH_3CN. The reaction was allowed to proceed at 37 ˚C for 1 h. Upon completion, CH_3CN (80 µL) was added, and the supernatant was recovered by centrifugation at 13.2 krpm for 20 min. The crude was diluted into H_2O (800 µL), and then purified by reverse phase HPLC (Table 3-1). Isolated product was characterized by MALDI MS (Table 3-3), quantified by UV-Vis absorption, split into 20 nmol aliquots, and then lyophilized until further use. X = 7-methoxycoumarinylalanine; G_o = glycolic acid.

One-Pot Ligation Deselenization with Selenocystine A phosphate buffer stock (200 mM Na_2HPO_4, pH 8.0) and a 5 M NaOH solution were freshly degassed using the freeze-pump-thaw method. 2% (v/v) PhSH was added to the buffer stock, and pH was quickly adjusted back to 8.0 using the degassed 5 M NaOH solution under an argon atmosphere. A 1.1x *L*-selenocystine stock (2.2 mM) was prepared by dissolving *L*-selenocystine (0.0007 g, 2.2 µmol) in the above buffer (1 mL). To thioester **2a** (20 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}), 10 µL of CH_3CN and 90 uL of the 1.1x *L*-selenocytine stock (final concentration 2 mM, 10 equiv) was added. The reaction was allowed to proceed at 37 °C for 12 h under an argon atmosphere.

For deselenization, a 50 µL aliquot of the ligation reaction was removed, and directly added to 50 µL of 2x deselenization buffer stock (40 mM TCEP, 40 mM DTT, 200 mM Na_2HPO_4, pH 8.0) that had been freshly degassed by freeze-pump-thaw method. The reaction was allowed to proceed at 37 °C overnight (≥ 14 h) under an argon atmosphere.

Upon completion, both ligation and deselenization crudes were diluted with argon-purged H_2O, and then analyzed by reverse phase HPLC on a Luna C8 column using gradient **2**. Individual fractions were recovered, and analyzed by MALDI-TOF MS and UV-Vis absorption spectroscopy.

VA-044 Desulfurization of Cys/Thioamide-Containing Peptide Peptide **3** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in 80 µL of argon-purged 1.25x buffer stock (50 mM TCEP, 125 mM Na_2HPO_4, pH 7.0). 10 µL of tBuSH was added, followed by 10 µL of freshly prepared 10x VA-044 stock (100 mM in argon-purged water). The reaction was allowed to proceed at 37 °C for 10 min under argon atmosphere, and then quickly quenched by chilling to 0 °C on ice. Excess tBuSH was removed by a stream of argon, and then the crude mixture was diluted into H_2O for analysis by reverse phase HPLC on a Luna C8 analytical column using gradient **3**. Fractions were collected, and then analyzed by MALDI-TOF MS and UV-Vis absorption spectroscopy. Yield was quantified by integrating peak areas in HPLC chromatogram.

TCEP Dosage Dependence The reactions were conducted similarly to the standard procedure, except that 1.25x buffer stocks of various TCEP concentration were used, and that a 2 h reaction time was adopted. At low TCEP concentrations, side product **S6** was observed, with a disulfide bond between Cys and ambient tBuSH (Figure 3-22). The side product was not observed at high TCEP concentrations.

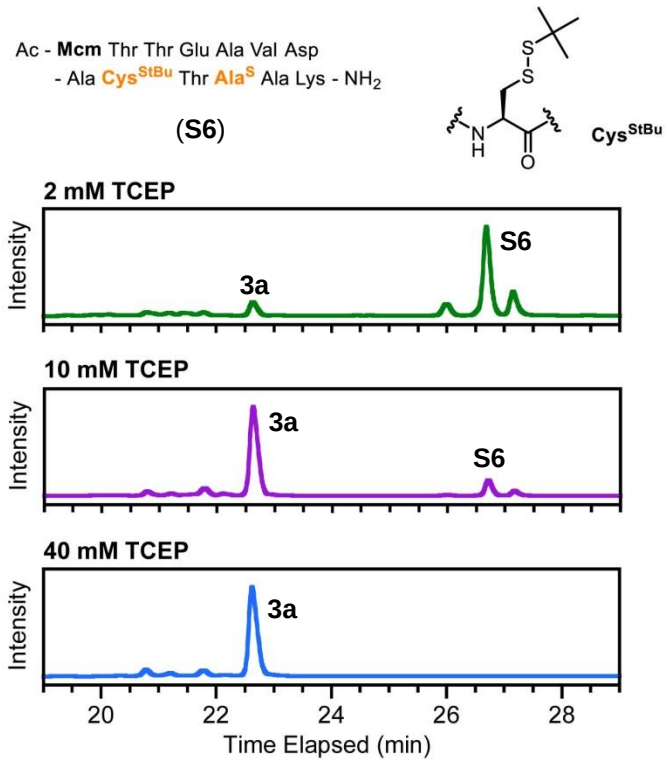


Figure 3-22. TCEP Dosage Dependence and Generation of Disulfide Bonded Side Product. MALDI-TOF MS: [**S6** + H]$^+$, expected 1570.65, found 1570.74.

VA-044 Dosage Dependence Reactions were conducted similarly to the standard procedure, except that 10x VA-044 stocks at various concentrations were used for 2 h or 18 h (Figure 3-23). Above 0.1 mM, no distinction was observed at the range of VA-044 concentrations tested. At very low VA-044 concentration, however, the reaction was significantly slower and messier − at 2 h, while all other conditions showed complete reaction, minimal product formation was observed at 0.01 mM VA-044. At 18 h, the reaction was still not complete, with additional side peaks in the chromagram.

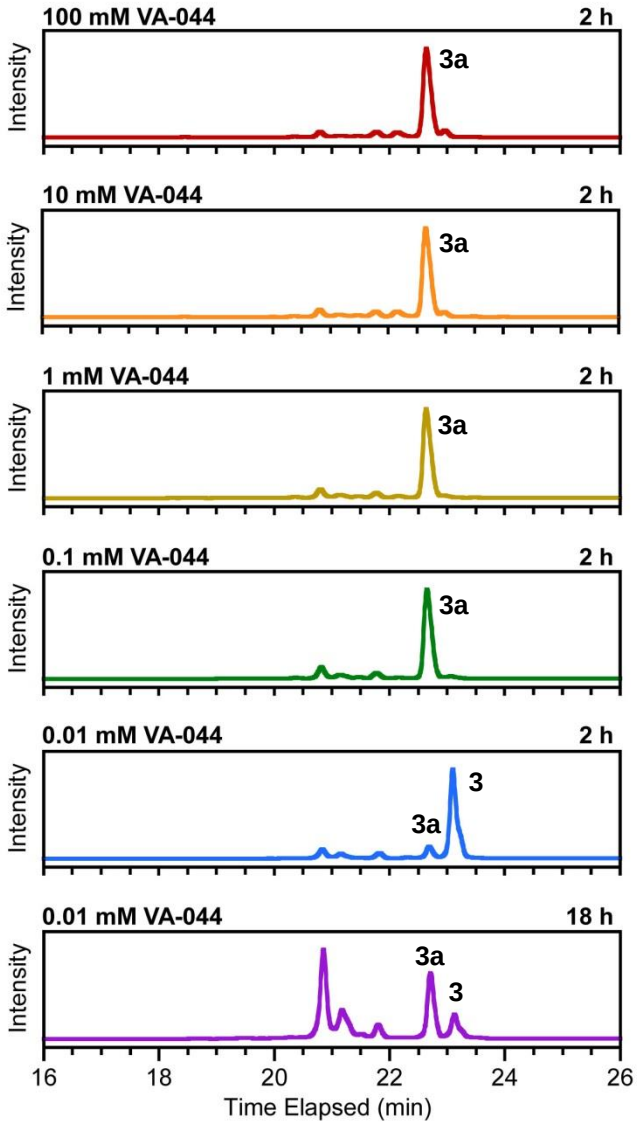


Figure 3-23. VA-044 Dosage Dependence on Chemoselective Cys Desulfurization.

Oxygen Tolerance Reactions were conducted similarly to the standard procedure, except that an undegassed buffer stock was used with 5 min or 10 min reaction time. While both reactions completed within 10 min, the undegassed condition resulted in a slightly slower kinetics – 80% complete (as measured by reactant consumed) by 5 min as compared to 89% complete for the degassed condition (Figure 3-24).

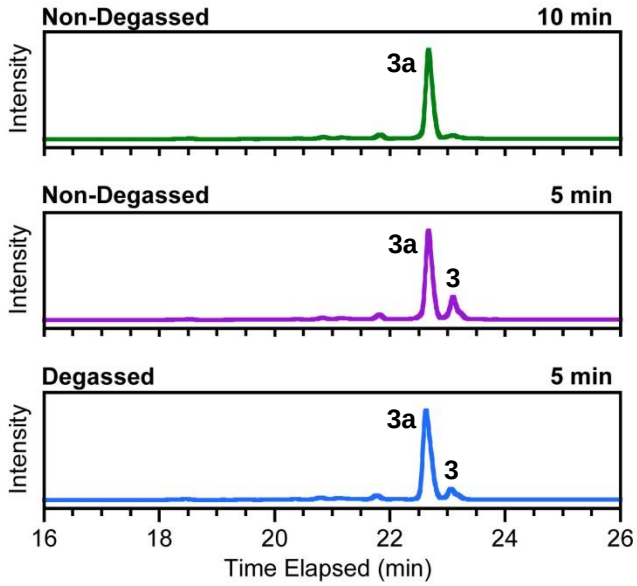


Figure 3-24. Oxygen Tolerance on Chemoselective Cys Desulfurization.

Denaturant Tolerance Reactions were conducted similarly to the standard procedure, except that a 1.25x buffer stock with denaturant was used with 2 h reaction time. No difference was found as compared to denaturant free conditions (Figure 3-25).

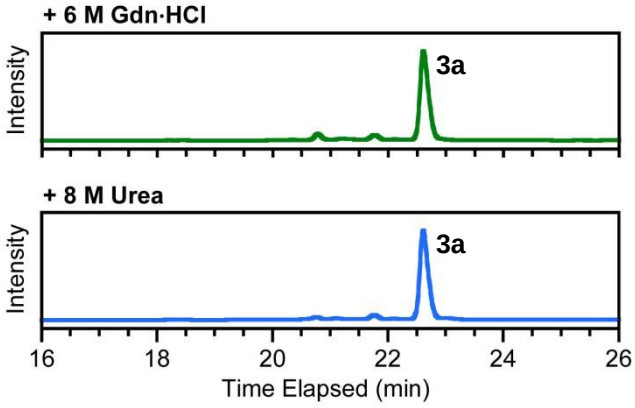


Figure 3-25. Denaturant Tolerance on Chemoselective Cys Desulfurization.

Prolonged Reaction with Cys/Thioamide-Containing Peptide Reactions were conducted similarly to the standard procedure, except that longer reaction times were adopted. Side products, particularly **3b**, accumulated over time (Figure 3-26).

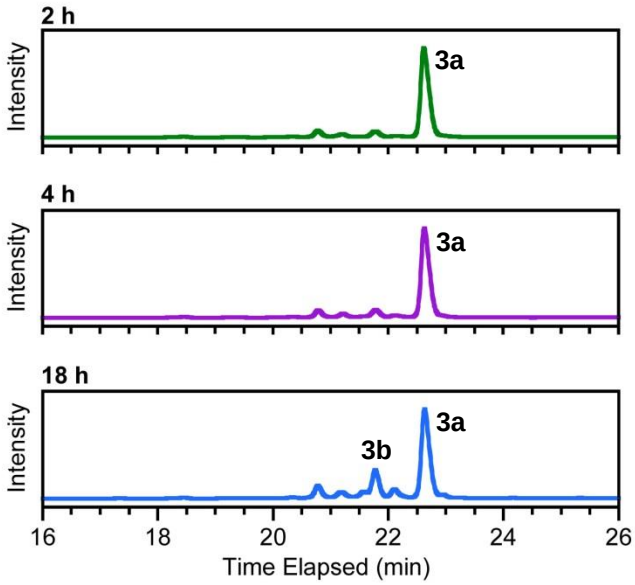


Figure 3-26. Accumulation of Non-Selectively Desulfurized Side Product over Time.

Thioacetamide Dosage Dependence Reactions were conducted similarly to the standard procedure, except that buffers with various concentrations of thioacetamide were used with 18 h reaction time. Suppression of thioamide-to-oxoamide conversion was effective over the wide range of concentrations tested; at very high concentration, thioacetamide may quench ambient radicals, leading to slower kinetics (Figure 3-27).

Oxygen and Denaturant Tolerance in the Presence of Thioacetamide Reactions were conducted similarly to the control reactions without thioacetamide for 18 h. No distinction was found with the addition of thioacetamide (Figure 3-28).

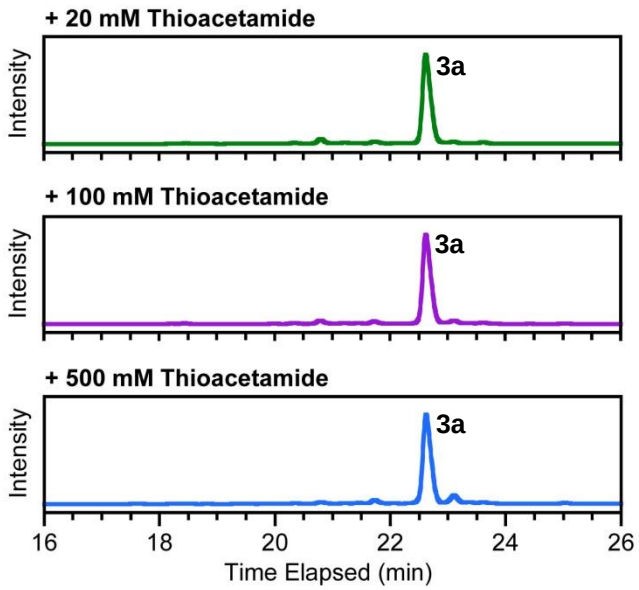


Figure 3-27. Thioacetamide Dosage Dependence on Thioamide Protection.

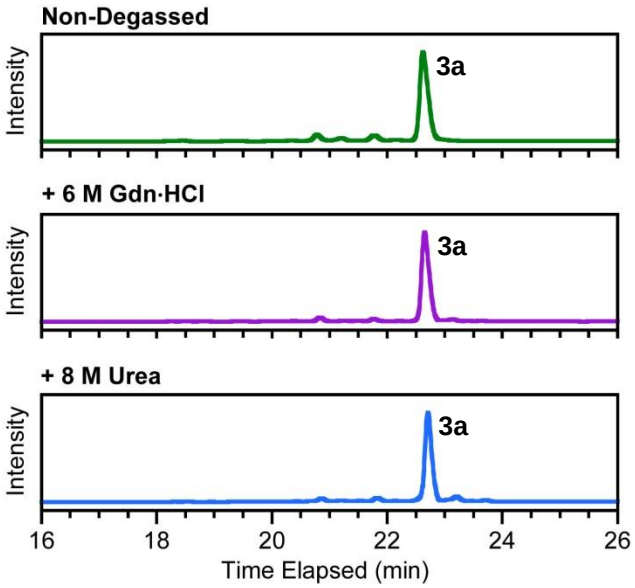


Figure 3-28. Oxygen and Denaturant Tolerance in the Presence of Thioacetamide.

Characterization of Cys-to-Ser Conversion Side Products We were surprised to find a residual peak at 21.8 min even in the presence of thioacetamide, which should

suppress thioamide-to-oxoamide conversion and abolish side product **3b**. Upon a closer examination of the MALDI-TOF MS and UV-Vis absorption spectra (Figure 3-30), we realized that the residual peak represents a separate side product **S7** where a Cys-to-Ser conversion took place with thioamide intact. The conversion could result from quenching of alkyl radical by water or dissolved oxygen in the reaction solution (Figure 3-29); in fact, a similar conversion had been reported for oxygen-related complications in Sec deselenization.[221,222] We believe both water and oxygen contribute to the Cys side reaction observed here – a 2% conversion still took place in argon purged buffer, as compared to 3% in non-degassed buffer.

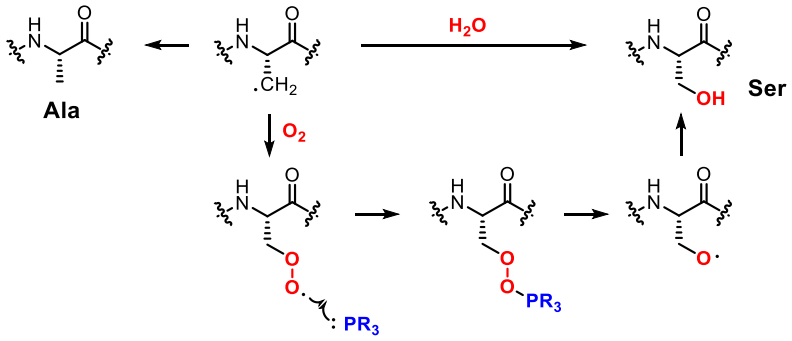


Figure 3-29. Proposed Mechanism for the Cys-to-Ser Conversion Side Reaction.

To prove the identity of the residual peak, we synthesized three control peptides − Ala/oxoamide peptide **3b'**, Ser/thioamide peptide **S7'**, and Cys/ oxoamide peptide **S8'** (prime symbol denotes the genuine peptide standards to distinguish them from species identified in desulfurization reactions). We showed that **3b'** and **S7'** both eluted at 21.8 min under the gradient used, but had vastly different UV-Vis absorption profiles. While **S8'** shares the same expected mass with **S7'**, it eluted at a different retention time, and of

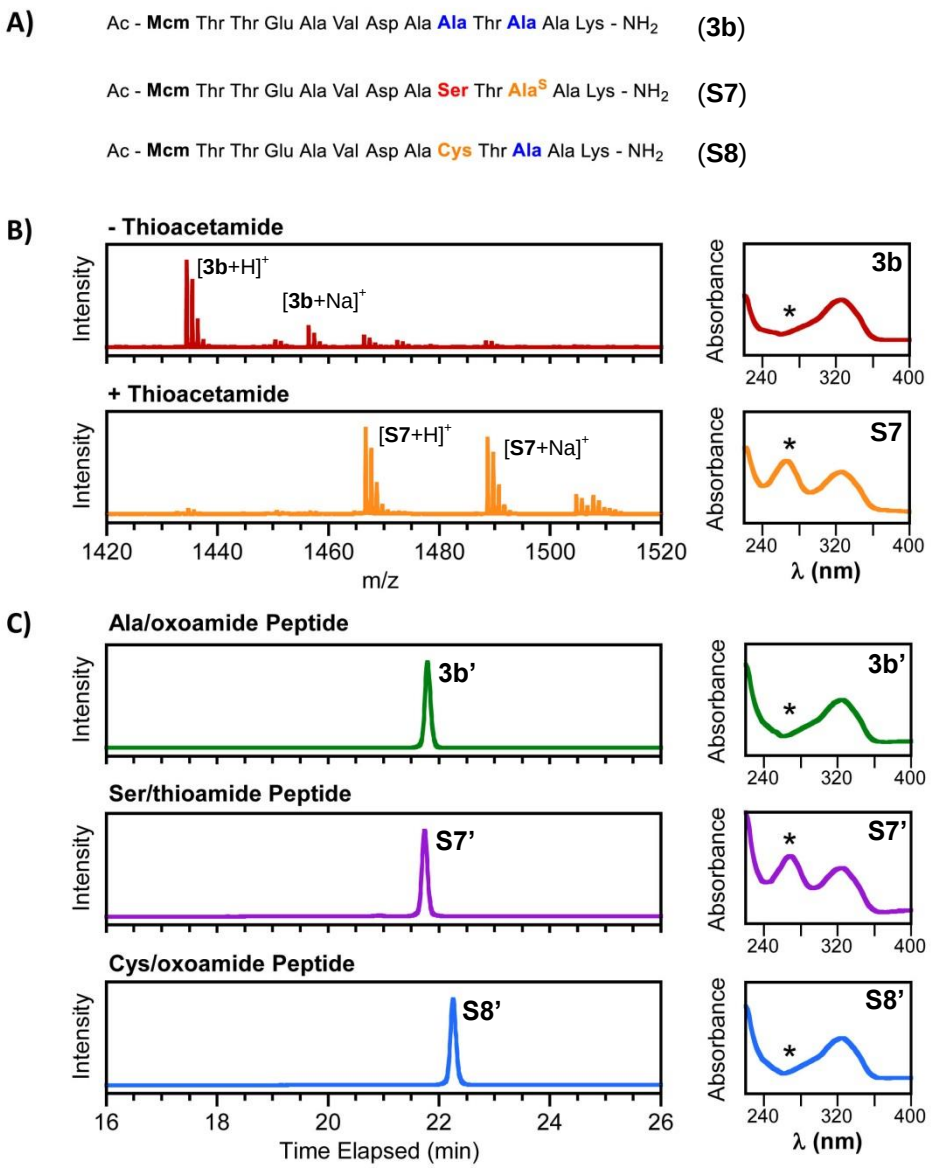


Figure 3-30. Characterization of Cys-to-Ser Conversion Side Reaction.

A) Sequence of all side products involved. B) MALDI-TOF mass spectra and UV-Vis absorption spectra for the 21.8 min peak isolated from reactions with or without thioacetamide. The retention times may be similar, but peak identities were vastly different. C) HPLC chromatrogram monitored at 325 nm and UV-Vis absorption spectra for genuine peptide standards synthesized by solid phase peptide synthesis (SPPS). Only **S7'** matches the profile observed in reaction with thioacetamide as additive. Mcm = 7-methoxycoumarinylalanine.

course, did not exhibit thioamide absorption at 272 nm (Figure 3-30). Therefore, we conclude that the 2% residual peak corresponded to Cys-to-Ser conversion side product **S7**, and that thioamide-to-oxoamide conversion was fully suppressed by thioacetamide.

Synthesis of Ac-Mcm-TTEAVDA-$SC_2H_4COOCH_3$ Thioester (**4a**) Protected precursor Ac-Mcm-$T^{tBu}T^{tBu}E^{OtBu}AVD^{OtBu}$A-OH (**S9)** was synthesized on 2-chlorotrityl chloride resin using standard solid phase peptide synthesis (SPPS) procedure, and cleaved with 1:1:8 acetic acid/trifluoroethanol/CH_2Cl_2. Lyophilized crude **S9** (20 µmol, 1 equiv) was mixed with dimethylformamide (DMF, 1 mL), tetrahydrofuran (THF, 1 mL), *N*,*N*-diisopropylethylamine (DIPEA, 34.8 µL, 200 µmol, 10 equiv) and methyl 3-mercapto-propionate (43.3 µL, 400 µmol, 20 equiv). PyBOP (0.0520 g, 100 µmol, 5 equiv) was added, and then the reaction was allowed to proceed at room temperature for 4 h. Upon completion, the reaction was quenched with 0.1% trifluoroacetic acid in water (10 mL). The precipitate was collected by centrifugation at 13.2 krpm for 20 min, and then cleaved with 3 mL of 1:1:38 triisopropylsilane/methyl 3-mercaptopropionate/trifluoroacetic acid for 30 min. The crude mixture was dried by rotary evaporation, and then brought up in 50:50 acetonitrile/water for purification by reverse phase HPLC. The purified product was quantified by UV-Vis absorption (ε_{325} = 12,000 M^{-1} cm^{-1}) and lyophilized. Mcm = 7-methoxycoumarinylalanine.

Theoretical Prediction of PhS^- Concentration at Various pH Knowing that the pK_a of thiophenol is 6.6 and that its solubility is 0.08% (7.3 mM), we can theoretically predict the concentration of PhS^- at any given pH by solving the acid-base equilibrium.

Assuming that the protonated form PhSH can be fully removed by lyophilization (typical vacuum: 0.024 mBar << vapor pressure of PhSH: 1.8 mBar), the concentration of PhS^- would provide a good estimation for residual aromatic thiol/thiolate after lyophilization.

$$\mathrm{pKa} = -\log\frac{[\mathrm{H^+}]\cdot[\mathrm{PhS^-}]}{[\mathrm{PhSH}]} = 6.6$$

$$[\mathrm{PhS^-}] + [\mathrm{PhSH}] = 0.726\ \mathrm{M}$$

$$\Rightarrow [\mathrm{PhS^-}] = \frac{0.726 \times 10^{-6.6}}{[\mathrm{H^+}] + 10^{-6.6}}$$

When lyophilization is conducted at pH 7, as much as 5.2 mM of 7.3 mM total PhSH would remain in the PhS^- form, which would overwhelm the 0.1 mM peptide. When lyophilization is carried out at pH 1, however, the residual PhS- concentration is only 18 nM, an insignificant amount with respect to the 0.1 mM peptide.

One-Pot Ligation-Desulfurization A phosphate buffer stock (40 mM TCEP, 100 mM Na_2HPO_4, pH 7.0) was freshly prepared and degassed by argon purging. 2% PhSH (v/v) was added, and then pH was quickly adjusted back to 7.0 under argon atmosphere. Thioester **4a** and peptide **4b** (50 nmol each) were each dissolved in 25 µL of the above buffer, and then combined. The reaction was allowed to proceed at 37 °C overnight.

For desulfurization, 10 µL of the ligation crude was removed, acidified with 100 µL of 1% trifluoroacetic acid in water, and then quickly frozen and lyophilized. The crude residue was brought up in 80 µL of freshly degassed 1.25x desulfurization buffer (50 mM TCEP, 125 mM Na_2HPO_4, pH 7.0), and then tBuSH and VA-044 were added to proceed with the standard desulfurization procedure. Upon completion, all crude reaction samples

were diluted into H_2O, and analyzed by reverse phase HPLC on a Luna C8 analytical column using gradient **3**. Individual fractions were collected, and then analyzed by MALDI-TOF MS and UV-Vis absorption spectroscopy.

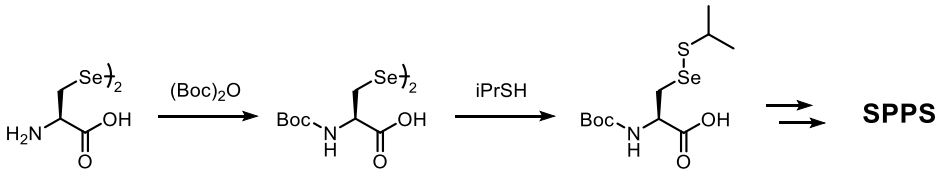


Figure 3-31. Synthetic Scheme for Sec-Containing Peptides.

Synthesis of *N,N'*-di-Boc-*L*-selenocystine (S10) L-Selenocystine (0.1670 g, 0.5 mmol, 1 equiv) was dissolved in argon-purged H_2O (4 mL) and 1,4-dioxane (4 mL), and chilled to 0 °C on ice. Trimethylamine (900 μL, 6 mmol, 12 equiv) was added, followed by Boc anhydride (0.2728 g, 1 mmol, 2 equiv). The reaction was allowed to proceed on ice for 2 h, and then at room temperature overnight (≥ 12 hr) under argon atmosphere. Upon completion, the reaction was diluted with H_2O (18 mL), acidified with 3 M hydrochloric acid (3 mL), and then quickly extracted with ethyl acetate (30 mL × 3). The organic layers were combined, concentrated by rotary evaporation, and then purified by flash chromatography in 0.1% acetic acid in ethyl acetate. A pale yellow foam was isolated as the product (0.2383 g, 0.45 mmol, 89% yield). R_f 0.15 in ethyl acetate with 0.1% acetic acid. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 8.93 (s, 2H), 5.65 (d, J = 6.9 Hz, 1H), 4.57 (m, 2H), 3.48 (m, 2H), 3.40 (dd, J = 12.7, 5.6 Hz, 2 H), 1.44 (s, broad, 18 H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 174.95, 155.84, 81.18, 54.01, 33.34, 28.66. ESI$^-$-HRMS: calculated for $C_{16}H_{27}N_2O_8Se_2^-$, 535.0098; found $[M - H]^-$ 535.0095.

Synthesis of Boc-Sec(S-*i*Pr)-OH (S11) **S10** (0.1681 g, 0.31 mmol, 1 equiv) was dissolved in argon-purged tetrahydrofuran (THF, 5 mL). Triethylamine (439 μL, 3.2 mmol, 10 equiv) was added, followed by 2-propanethiol (599 μL, 6.3 mmol, 20 equiv). The reaction was allowed to proceed at room temperature for 3 h under argon atmosphere. Upon completion, the reaction mixture was dried by rotary evaporation. The residue was brought up in H_2O (10 mL), acidified with 3 M hydrochloric acid (1 mL), and quickly extracted with ethyl acetate (10 mL × 3). Organic layers were combined, concentrated by rotary evaporation, and then purified by flash chromatography in 5:5 ethyl acetate/ petroleum ether with 0.1% acetic acid. A white foam was isolated as the final product (0.2005 g, 0.59 mmol, 93% yield). R_f in 0.27 in 5:5 ethyl acetate/petroleum ether with 0.1% acetic acid. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 5.43 (d, *J* = 6.9 Hz, 1H), 4.63 (dd, *J* = 11.3, 6.9 Hz, 1H), 3.33 (dd, *J* = 12.5, 4.2 Hz, 1H), 3.24 (dd, *J* = 12.1, 6.5 Hz, 1H), 3.03 (septet, *J* = 6.7 Hz, 1H), 1.46 (s, 9H), 1.32 (d, *J* = 6.6 Hz, 6H). ESI$^+$-HRMS: calculated for $C_{11}H_{21}NO_4SSeNa$ 366.0254, found $[M + Na]^+$ 366.0252.

Synthesis of Sec-Containing Peptide Peptide **5** was synthesized using standard solid phase peptide synthesis (SPPS) procedure on Rink amide resin. Sec was introduced as Boc-Sec(S-*i*Pr)-OH (**S11**) using standard HBTU activation method. The peptide was cleaved with 12:1:1:26 trifluoroacetic acid/2-propanethiol/triisopropylsilane/CH_2Cl_2, and then purified by reverse phase HPLC. **5** was isolated as the primary product, with an intramolecular hemiselenide bond between Sec and Cys.

Deselenization of Sec/Cys/Thioamide-Containing Peptide Deselenization buffer (40 mM TCEP, 40 mM DTT, 200 mM phosphate, pH 7.0) was freshly prepared and degassed by the freeze-pump-thaw method. Peptide **6** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in 100 µL of the above buffer, and then the reaction was allowed to proceed at 37 ˚C for 18 h. Upon completion, the reaction was chilled to 0 ˚C on ice, and then 100 µL of argon-purged CH_3CN was added. The supernatant was recovered by centrifugation at 13.2 krpm for 20 min, diluted into argon-purged H_2O (650 µL), and then analyzed by reserve phase HPLC on a Luna C8 analytical column using gradient **3**. Fractions were collected and analyzed for MALDI-TOF MS and UV-Vis absorption.

4-Mercaptophenylacetic Acid (MPAA) as Deselenization Additive Reaction was conducted similarly to the standard procedure, except that a buffer stock containing 10 mM MPAA was used. MPAA was chosen over PhSH for its low volatility.

Synthesis of Pen-Containing Peptide Fmoc-Pen(Trt)-OH was purchased from Sigma-Aldrich (St. Louis, MO), and incorporated into peptide **6** sequence using standard solid phase peptide synthesis (SPPS) procedure.

Desulfurization of Pen/Thioamide-Containing Peptide Reaction was conducted using standard desulfurization procedure with thioacetamide. Briefly, Peptide **6** (10 nmol, ε_{325} = 12,000 M^{-1} cm^{-1}) was dissolved in 80 µL of argon-purged 1.25x buffer stock (50 mM TCEP, 125 mM thioacetamide, 125 mM Na_2HPO_4, pH 7.0). tBuSH and VA-044 stock were then added, and the reaction was allowed to proceed at 37 °C for 18 h. The

crude was analyzed by reverse phase HPLC on a Luna C8 analytical column using gradient **3**, and characterized by MALDI-TOF MS and UV-Vis absorption spectroscopy.

3.6 Acknowledgement

This work was supported by funding from the University of Pennsylvania, the National Institute of Health (5R01NS081033-03 to E.J.P.), and the Searle Scholars Program (10-SSP-214 to E.J.P.). Instruments were supported by the National Science Foundation and National Institutes of Health include: HRMS (NIH RR-023444), X-ray diffractometer (NSF CRIF-0091925), and MALDI-TOF MS (NSF MRI-0820996). We thank Patrick J. Carroll for X-ray crystallography data acquisition, and Rakesh Kohli for assistance with HRMS analysis.

Chapter 4 . Chemoenzymatic Incorporation of Selenocysteine onto the Protein N-terminus by Aminoacyl Transferase (AaT)

4.1 Introduction

Having established a robust method for thioamide incorporation through traceless native chemical ligation (NCL), we have one last challenge to address – the installation of "erasable" ligation handles onto expressed protein fragments. For short peptides, ligation handles such as selenocysteine (Sec) and β-thiol/selenol amino acid analogs can be conveniently introduced by solid phase peptide synthesis (SPPS). However, SPPS has an inherent length limit (typically 70 aa)[270], beyond which synthesis and purification would become prohibitively difficult. Therefore, when an expressed protein C-terminal fragment is required for ligation sites that are far from the C-terminus, a new approach for Sec/analog attachment must be developed.

The conventional way to generate N-terminal Xaa protein fragments is through fusion protein expression. As demonstrated in our preparation of $\alpha S_{9\text{-}140}C_9$,[283] a construct with an N-terminal purification tag and a protease recognition sequence can be expressed, after which the N-terminal Cys is revealed by proteolysis (Figure 4-1A). To apply this method to N-terminal Sec/analog fragments, however, we would also need to adopt unnatural amino acid (Uaa) mutagenesis[272] for translational insertion of Sec/analog into the fusion protein. This would be overwhelmingly complex in terms of validation/optimization efforts – at the bare minimum, we will need to evolve an orthogonal synthetase-$tRNA_{CUA}$ pair for the analogs, ensure that all analogs are metabolically stable and nontoxic for cellular expression, and identify a protease that would tolerate a β-branched side chain at the cleavage site. If the unnatural side chains

are not readily tolerated by the ribosome, we would also need to evolve a modified ribosome for ligation handle incorporation.

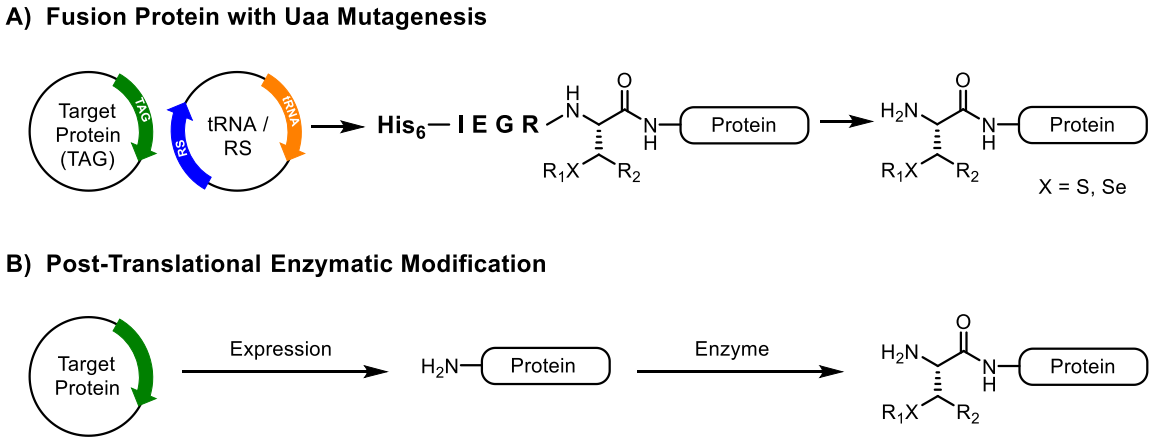


Figure 4-1. Methods for Sec and β-Thiol Analog Installation onto Expressed Proteins.

A) Fusion protein expression with unnatural amino acid (Uaa) mutagenesis. A TAG codon is introduced into the fusion protein construct at the desired site of ligation; a specifically evolved synthetase (RS) charges $tRNA_{CUA}$ with an unnatural amino acid, and "reads through" the TAG codon which would otherwise be interpreted as a "stop" codon. B) Post-translational enzymatic modification. Target protein is produced by regular cellular expression, and then N-terminally modified by an appropriate enzyme to install the Sec/analog.

In contrast, post-translational enzymatic modifications present a much simpler and arguably more versatile option (Figure 4-1B). Rather than evolving an entire cellular expression machinery for these ligation handles, we just need to identify (or evolve) an appropriate enzyme that could transfer synthetic moieties like Sec and β-thiol analogs site-specifically onto the N-termini of expressed proteins. If viable, this strategy would have two major advantages: 1) with minimal deviation from regular recombinant protein production procedure, very little optimization is necessary for target protein expression and handling; 2) conducting the modification *ex vivo* with a single enzyme gives us

nearly full control over key variables, where we can quickly screen substrates, optimize reaction conditions and evolve the enzyme for incorporation of our desired analogs.

A naturally-existing enzyme that performs exactly this type of modifications is *E. coli* aminoacyl transferase (AaT), a component of the prokaryotic N-end degradation pathway (Figure 4-2). AaT recognizes an N-terminal Lys or Arg on an expressed protein, and then attaches a Leu or Phe onto the N-terminus from a tRNA donor[258]. In *E. coli*, this serves as a means to tag proteins for degradation; the L/F-K/R motifs are recognized by an "adapter protein" ClpS, which then transports the tagged proteins to a proteolytic complex for degradation.[256] From a protein engineering perspective, this unique N-terminal amide bond formation capability presents an excellent opportunity for selective

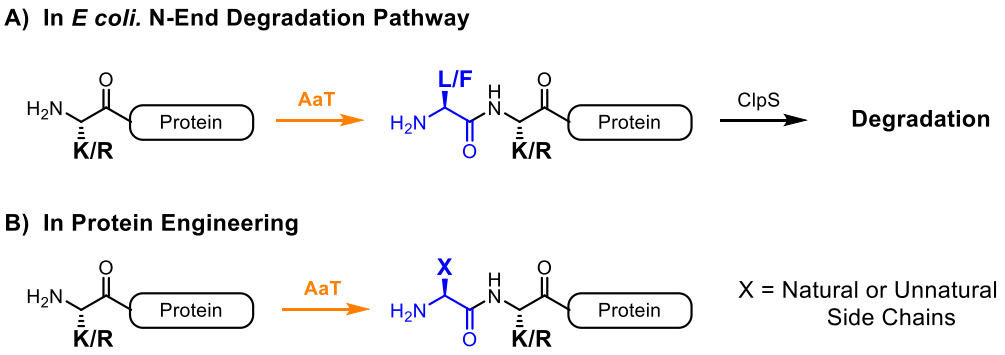


Figure 4-2. Aminoacyl Transferase (AaT) as a Protein Engineering Tool.

A) Biological function of AaT in the *E. coli* N-end degradation pathway. AaT transfers a Leu or Phe onto the exposed Arg or Lys on the protein N-terminus, which results in an N-terminal L/F-R/K motif that is recognized as a degradation signal. B) AaT as a protein engineering tool. When AaT is used in isolate without the other components of the N-end degradation pathway, the enzyme can be repursposed for transfer of natural or unnatural amino acids onto the protein N-terminus.

protein N-terminal modification − when used in isolation (i.e. in the absence of other components of N-end degradation pathway), AaT can be adapted as a tool to incorporate a variety of natural or synthetic amino acids site-specifically onto the protein N-terminus.

AaT was discovered in the 1960s from *E. coli* cell lysate for its ability to transfer Leu or Phe onto proteins[257,258]. It was later established that the AaT modification was specific to the protein N-terminus[259], and that the anticodon of the tRNA was unneccesary for the AaT recognition of aminoacylated tRNA as its substrate[263]. In early protein engineering studies [264,265,267], AaT has been used by Sisido and Tirrell for the proof-of-concept incorporation of unnatural amino acids such as acridone and fluorinated leucine. In our own group, we were able to greatly simplify the substrate synthesis and screening process by adopting a minimal adenosine donor (Figure 4-3)[268] – while full-length aminoacyl-tRNA is the natural substrate, only the last adenosine on the acceptor stem is essential for AaT recognition and activity[261]. This allows us to take an efficient "chemoenzymatic" approach, where adenosine donors of unnatural amino acids are chemically synthesized through simple organic transformations, and then directly screened for activity as AaT substrates without being limited by the RS specificity in the tRNA-charging step. Using this approach, we were able to identify novel AaT substrates such as benzo-phenone or disulfide protected amino acids in our previous work[268,269].

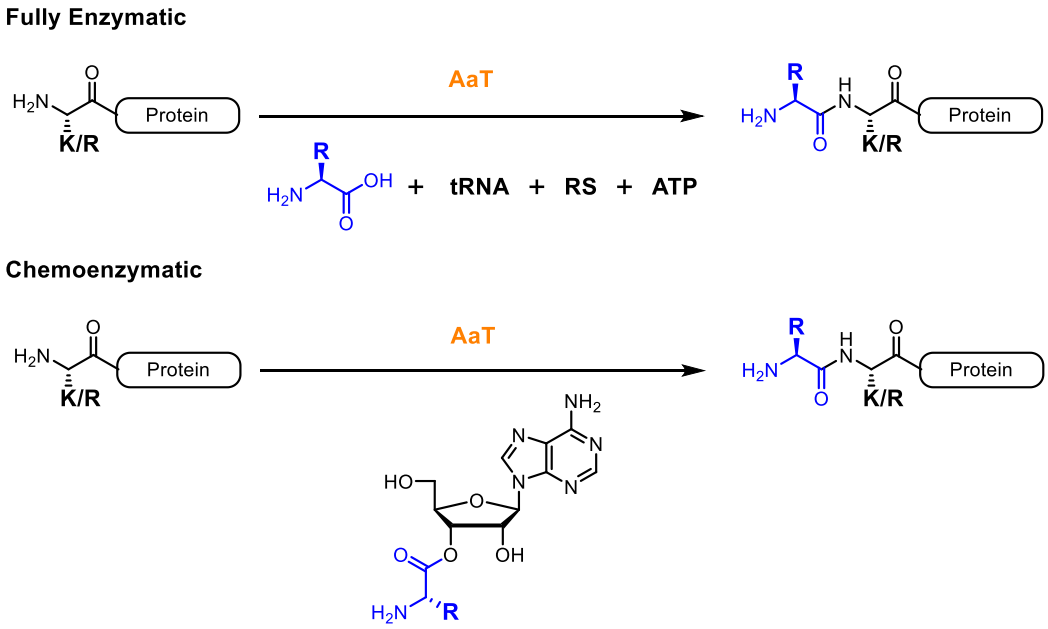


Figure 4-3. Aminoacyl Adenosine Donors as Minimal AaT Substrates.

Top: Fully Enzymatic Approach. Desired amino acid is charged onto a tRNA by its corresponding synthetase (RS), which is then utilized by AaT as a substrate. Bottom: Chemoenzymatic Approach. An amino acid/adenosine conjugate is directly used as donor for AaT mediated protein N-terminal modification. Adapted from Wagner et al.[268]

By examining the known substrate scope and crystal structures of AaT,[261, 264-265, 267-269] we hypothesized that the AaT-mediated N-terminal modification strategy would be applicable to hemiselenide protected selenocysteine (Sec) using wildtype AaT without additional directed evolution. AaT is a somewhat "promiscuous" enzyme; it has equal preference between its natural substrates Leu and Phe, and has been shown to tolerate a variety of unnatural substrates in reconstituted *ex vivo* systems (Figure 4-4). In the crystal structure with a Phe-adenosine substrate bound, there is a small void space next to the ε-carbon of Phe that could tolerate side chain extension (Figure 4-5); with the precedence

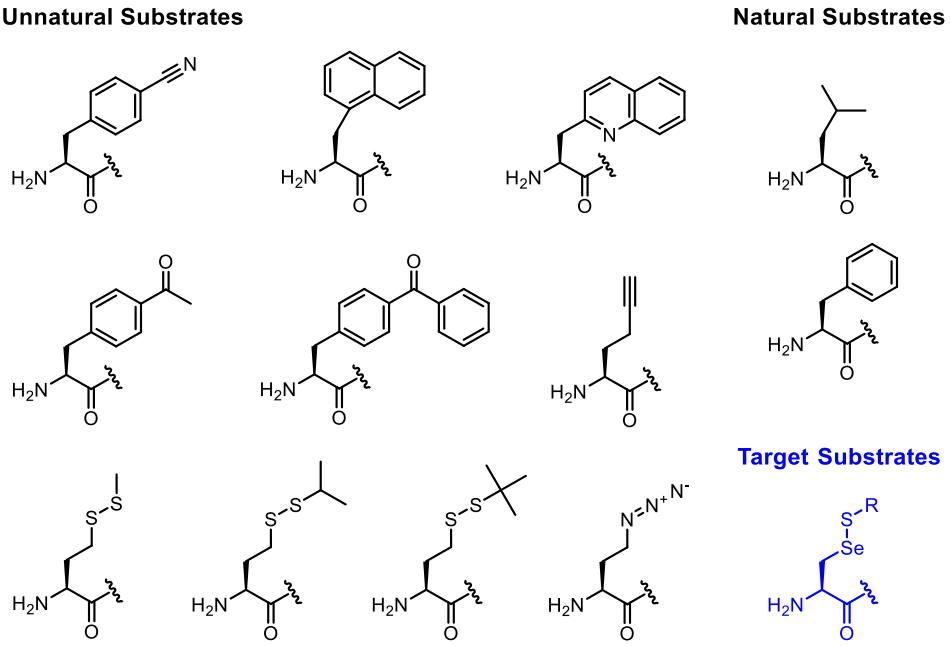


Figure 4-4. Selected Natural and Unnatural Substrates of AaT.

See *Introduction* chapter for a complete summary of known AaT substrates. We have shown previously that un-protected Cys and homocysteine (Hcs) are not viable substrates[269].

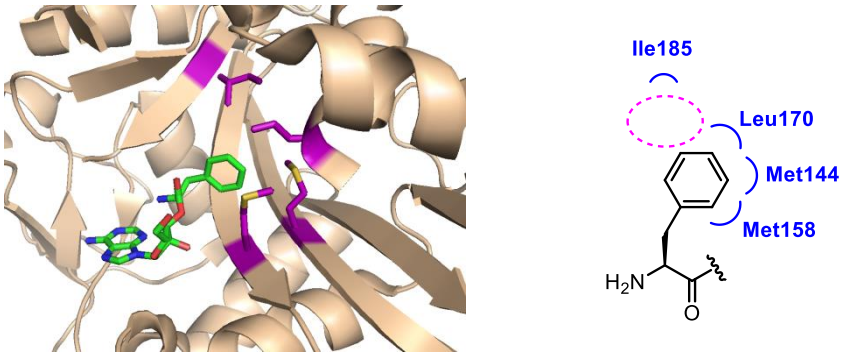


Figure 4-5. Crystal Structure and Contact Scheme of AaT Substrate Binding Pocket.

Crystal structure adapted from Watanabe *et al.* (PDB: 2Z3K)[261]; residues in direct contact with substrate Phe side chain (within 5 Å) are highlighted in purple; certain non-contact segments are omitted for clarity. A small void space is present next to the *ε*-carbon (pink dotted circle).

of disulfide protected homocysteine and polyaromatic amino acids, we also expect good conformational flexibility of this binding pocket for Sec(SR) side chains

It is worth noting that while the Sec amino acid can be naturally incorporated with a special cellular machinery[298], a post-translational modification approach is still necessary for traceless ligation. The Sec amino acid is encoded by the UGA stop codon, which is re-interpreted as a Sec codon by a special elongation factor EFsec, a selenocysteine insertion sequence (SECIS) in the 3'-untranslated region of the mRNA, and a SECIS/EFsec-binding protein SBP2[245]. Current attempts to recombinantly express Sec-containing proteins had limited success, where Trp mis-incorporation (Trp codon: UGG) and low yields due to chain termination were cited as the most common complications[252]. In addition, all natural and recombinant Sec-proteins are expressed with a Sec-Xaa_n-Cys motif (n = 2 ~ 4) to sequester the highly reactive selenol side chain as an intramolecular hemiselenide with the adjacent Cys during expression and handling[251,252]. If we choose the recombinant expression approach with Sec/Cys motifs, it would completely defeat the purpose of ligation site expansion through traceless NCL by re-introducing the need for a native Cys residue (or a non-native Cys mutation). Another approach is the metabolic replacement of the Cys sulfur with selenium by supplying nutrients with a certain percentage of selenites in the *E. coli* growth media[240]; this results in global replacement of all Cys, which is also undesirable for our applcations. In comparison, the AaT post-translational modification approach directly incorporates a single Sec with hemiselenide protection in a site-specific manner, enabling the expansion of traceless NCL to expressed protein fragments without the Cys complication.

4.2 Results and Discussion

To screen hemiselenide protected Sec derivatives for activity as AaT substrates, we first needed to chemically synthesize the corresponding "minimal" adenosine (Ade) donors. Using a similar strategy as previously applied to other amino acids[268,269], we developed a 6-step synthesis of H-Sec(SR)-Ade donors **6** from *L*-selenocystine **1** (Figure 4-7). Hemi-selenides have unique thiol exchange properties – similar to the well-known disulfide exchange, the Se−S bond is in constant equilibrium with other thiol or selenol species in their surrounding environment, and can be readily substituted by supplying large excess of the desired thiol (Figure 4-6)[226,220]. Taking advantage of this property, we were able to greatly simplify and streamline our synthesis − an H-Sec(*S*-*t*Bu)-Ade donor **6c** was first prepared in bulk, and then derivatized through a simple one-step hemiselenide exchange to generate a library of adenosine donors **6a-e**. We note that these exchange reactions are highly versatile – interconversion among the H-Sec(SR)-Ade donors can be achieved by simply incubating one donor with excess of the desired thiol (Figure 4-14).

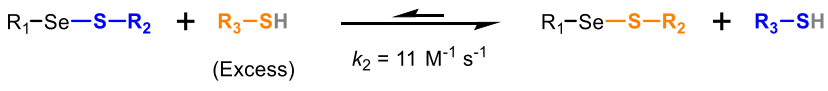


Figure 4-6. Thiol Exchange Equilibrium of Hemiselenides.

To assess the chemoenzymatic activity of our Sec-adenosine donor library, we subject these donors to an *in vitro* transfer assay using a model peptide LysAlaAcm **7**. If the hemiselenide protected Sec amino acids are viable substrates for AaT, we should observe

transfer of the Sec(SR) cargo from adenosine donor onto the N-terminus of LysAlaAcm. Indeed, the formation of Sec(SR)-LysAlaAcm products **8a-e** was clearly observed for all

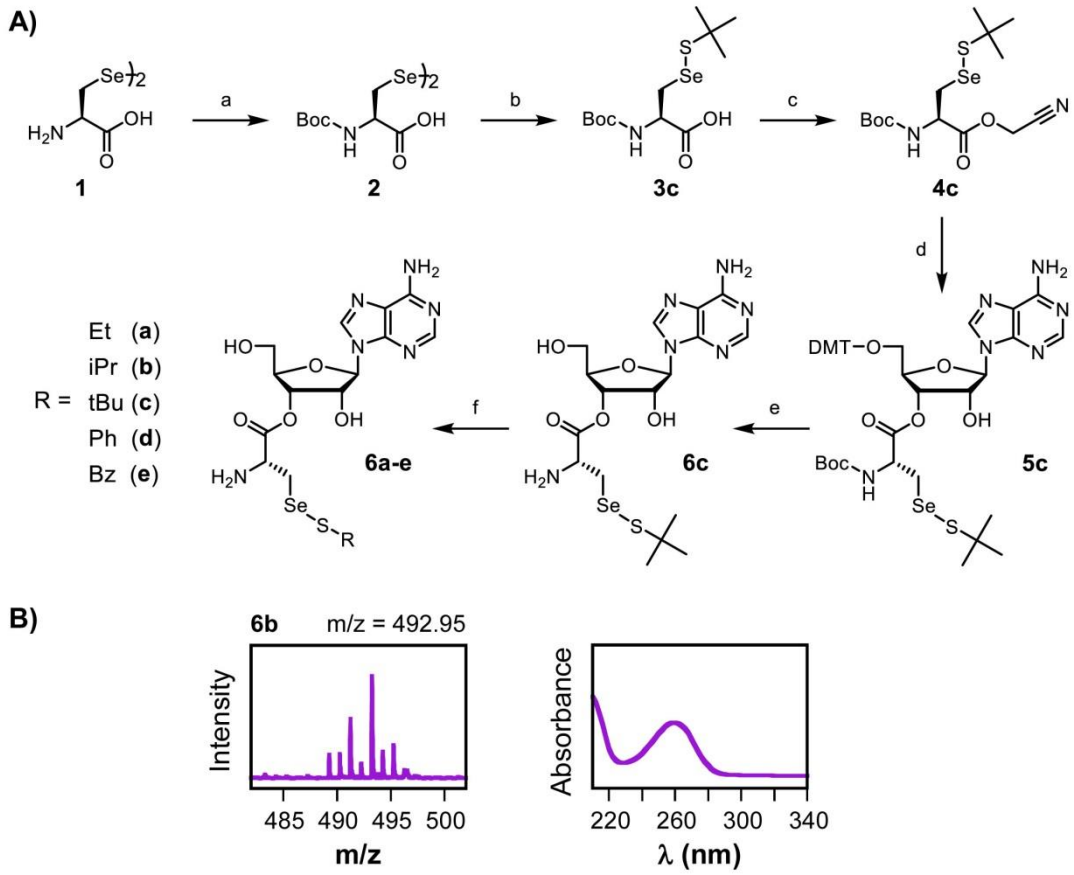


Figure 4-7. Synthesis of Hemiselenide-Protected Sec-Ade Donors from *L*-Selenocystine.

A) Synthesis scheme from *L*-selenocystine **1** to desired adenosine donors **6a-e**. B) MALDI-TOF MS and UV-Vis absorption characterization using **6b** as an example. Selenium isotopic pattern can be clearly observed in final product by MALDI-TOF MS; presence of adenosine is also evident in UV-Vis absorption. See Figure 4-12 for characterization of other products. Conditions: a) $(Boc)_2O$, DIPEA, 1:1 H_2O/1,4-dioxane; b) tBuSH, DIPEA, THF; c) $ClCH_2CN$, DIPEA, THF; d) (DMT)-A, DIPEA, catalytic NBu_4Ac, THF; e) TFA, tBuSH, TIPS, THF; f) RSH, H_2O. Boc = *tert*-butoxy carbamate; DIPEA = *N*,*N*-diisopropylethylamine, THF = tetrahydrofuran; DMT = 4,4'-dimethoxytrityl; (DMT)-A = 5'-*O*-DMT-adenosine; NBu_4Ac = tetrabutylammonium acetate; TFA = trifluoroacetic acid; TIPS = triisopropylsilane; Et = ethyl; iPr = isopropyl; Ph = phenyl; Bz = benzyl.

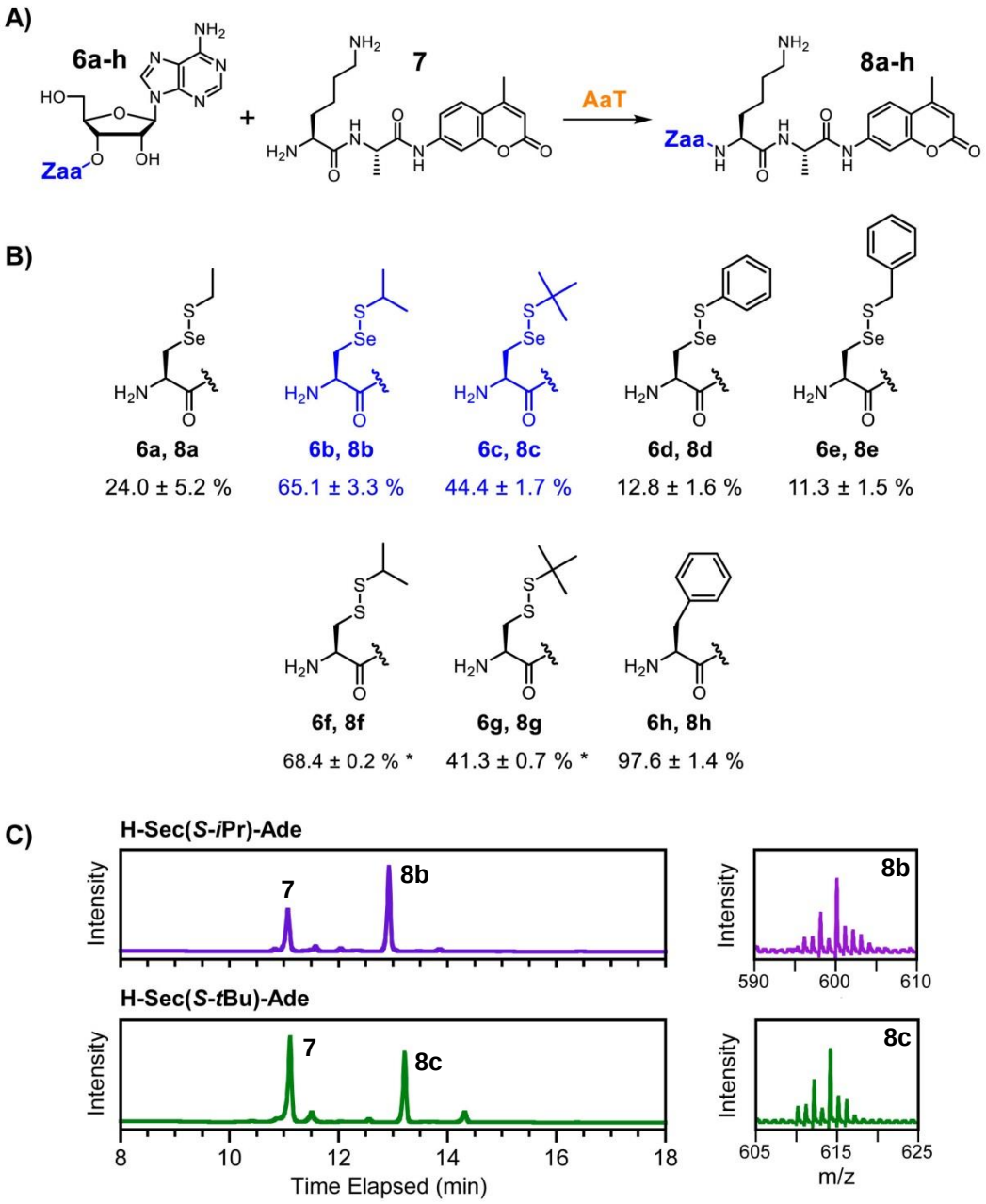


Figure 4-8. Screening of Hemiselenide Protected Sec-Ade Donors as AaT Substrates.

A) Schematic representation of the AaT activity assay for substrate screening. B) Transfer efficiencies of various amino acids by wildetype AaT. The natural substrate Phe is also tested as H-Phe-Ade donor for positive control. C) HPLC and MALDI-TOF MS characterization of the transfer reactions. Chromatogram monitored at 325 nm; integrity of selenium in isolated products can be clearly observed in the MALDI-TOF spectra. See Table 4-4 for a complete characterization of all test reactions. Acm = 7-aminocoumarin.* Data on protected Cys were acquired by Tomohiro Tanaka and Anne M. Wagner.

Sec-Ade donors tested (Figure 4-8). The isopropyl protected Sec(S-*i*Pr) and *t*-butyl protected Sec(S-*t*Bu) exhibited the highest transfer efficiencies, presumably due to the better steric accommodation of the small, branched side chains by the AaT substrate binding pocket (see Figure 4-5). As a further validation, we also synthesized the Cys counterparts H-Cys(*S*-*i*Pr)-Ade **6f** and H-Cys(*S*-*t*Bu)-Ade **6g** of the two most effective Sec-Ade donors, and observed nearly identical transfer efficiencies as the Sec substrates **6b** and **6c**. This gave us confidence that selenium in the hemiselenide side chains did not interfere with our AaT-mediated chemoenzymatic transformation. Lastly, we showed that the reaction was tolerant of ambient oxygen – nearly identical transfer efficiencies were observed in degassed or non-degassed reaction buffers (Figure 4-15).

To demonstrate the utility of this chemoenzymatic strategy in incorporating Sec onto the N-termini of expressed protein fragments, we generated $\alpha S_{6\text{-}140}$ **9** (a truncated version of αS with an N-terminal Lys) through cellular expression and successfully transferred Sec(*S*-*i*Pr) onto its N-terminus from adenosine donor **6b** (Figure 4-9). The integrity of Sec can be further confirmed in a subsequent deselenization test – when **10** was treated with *tris*(2-carboxyethyl)phosphine (TCEP), a 155 Da mass shift was clearly observed, corresponding to the characteristic deselenization of Sec into an Ala (Figure 4-16).

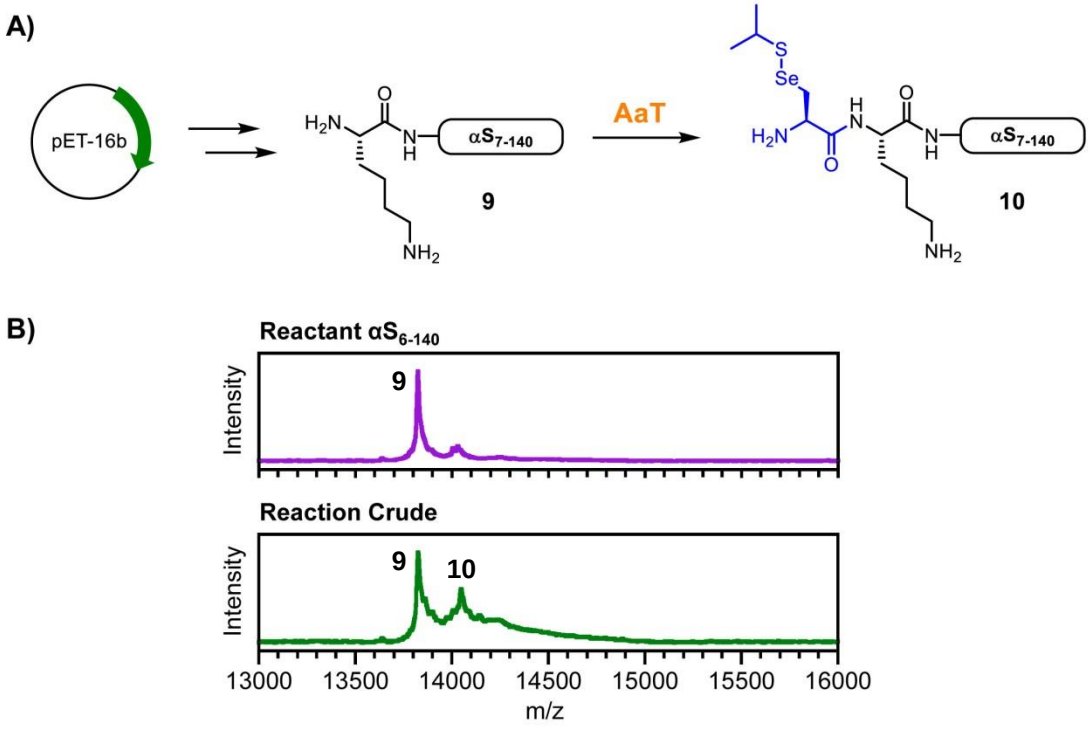


Figure 4-9. Chemoenzymatic Incorporation of Sec(*S*-*i*Pr) onto Express $\alpha S_{6\text{-}140}$.

A) Schematic representation of the AaT chemoenzymatic modification reaction. $\alpha S_{6\text{-}140}$ was prepared through cellular expression, and then subject to AaT mediated N-terminal modification using **6b** as adenosine donor. B) MALDI-TOF MS characterization. [**9** + H]$^+$, expected 13828.95, found 13824.14; [**10** + H]$^+$, expected 14053.92, found 14049.13.

4.3 Conclusion

We have developed an efficient chemoenzymatic approach to incorporate hemi-selenide protected selenocysteine (Sec) site-specifically onto the N-termini of expressed proteins with *E. coli* aminoacyl transferase (AaT). Using minimal adenosine donors, we identified that small, branched side chain protecting groups were preferred as AaT substrates, with the isopropyl protected Sec(*S*-*i*Pr) and *t*-butyl protected Sec(*S*-*t*Bu) being the most effective. As a proof-of-concept example, we successfully transferred Sec(*S*-*i*Pr) onto the N-terminus of an expressed protein, $\alpha S_{6\text{-}140}$, from the corresponding adenosine donor. Combining this chemoenzymatic approach with solid phase peptide synthesis (SPPS) for short peptides, we are now capable of generating N-terminal Sec peptide and protein fragments of any size (given an appropriate N-terminus, see discussion below), enabling thioamide incorporation at any desired regions in the protein sequence through traceless native chemical ligation.

4.4 Future Directions

To fully realize our vision of efficiently installing any traceless ligation handle at any desired position, we need to achieve three additional milestones: 1) quantitative transfer efficiency of the desired ligation handle; 2) ability to utilize β-thiol analogs as AaT substrates; 3) expansion of the N-terminal recognition residue to those other than Lys or Arg.

To achieve quantitative transfer efficiency, we would need to either evolve a more efficient AaT for our desired substrates, or reinstate the fully enzymatic route for AaT-mediated N-terminal modification (see Figure 4-3). We chose the second approach because it is more feasible − by re-introducing a RS into the reaction, the charged tRNA substrate can be continuously regenerated. It does not require either the AaT or the RS to achieve quantitative transfer/aminoacylation efficiency; with the appropriate reaction time, the substrate regeneration process *per se* is sufficient to drive the transfer reaction to completion. The hurdle is, of course, to identify/generate a synthetase that would recognize our desired ligation handle. As a preliminary trial, we screened *S. cerevisiae* phenylalaninyl synthetase (PheRS), *E. coli* methioninyl synthetase (MetRS), *E. coli* leucyl synthetase (LeuRS) and their rationally designed mutants, but unfortunately were not able to identify a positive candidate. (See *Materials and Methods* for details.) With the vast number of potential variables, we will next resort to directed evolution to generate a viable synthetase.

For the second goal of expanding AaT substrate scope to β-thiol analogs, we will utilize directed evolution and computational design to evolve an appropriate AaT mutant. In wildtype AaT, there are tight van der Waals contacts between the β-carbon of the substrate and nearby residues, precluding the incorporation of β-branched substrates. Rationally designed mutations have been attempted by other members in our group, but did not yield much success. Work is currently underway in our group to use a combination of surface display screening and *in silico* design to generate a mutant AaT for β-branched substrates.

To expand the selection of N-terminal recognition residues, we can either mutate the AaT peptide recognition pocket or explore other transferases with different recognition properties. For *E. coli* AaT, the preference for N-terminal Arg or Lys is dictated by the presence of a negatively charged Glu156 residue in the peptide binding pocket, with the assistance of Tyr42, Tyr120 and Gln188 as hydrogen bonding partners[261]. Conceivably, these residues can be mutated through directed evolution to alter the steric and electrostatic selectivity of the binding pocket, thus expanding the scope of N-terminal residue recognition. Another option is to adopt other transferases – for example, arginyl transferase (ATE-1) in *S. cerevisiae* and *H. sapiens* recognizes an N-terminal Asp, Glu, and Cys, and then transfers an Arg onto the N-terminus of these proteins[299,300]. Using a similar minimal adenosine donor strategy, we can quickly characterize the substrate promiscuity of these enzymes, and if necessary, evolve them for our desired ligation handles.

4.5 Materials and Methods

General Information *L*-Selenocystine was purchased from Sigma-Aldrich (St. Louis, MO). 5′-*O*-(4,4′-Dimethoxytrityl) adenosine ((DMT)-A) was purchased as a custom order from ChemGenes Corporation (Wilmington, MA). Lysylalanylamino-methylcoumarin (LysAlaAcm) was purchased from Bachem (Torrence, CA). The pEG6 plasmid, containing His_{10}-tagged *E. coli* AaT, was a gift from Alexander Varshavsky (California Institute of Technology). QuikChange® site-directed mutagenesis kits were purchased from Stratagene (La Jolla, CA). DNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA). Bradford reagent assay kits were purchased from BioRAD (Hercules, CA). Amicon Ultra centrifugal filter units (3k Da MWCO) were purchased from EMD Millipore (Billerica, MA). All other reagents and solvents were purchased from Fisher Scientific (Waltham, MA) or Sigma-Aldrich (St. Louis, MO) unless otherwise specified.

High resolution electrospray ionization mass spectra (ESI-HRMS) were collected with a Waters LCT Premier XE liquid chromatograph/mass spectrometer (Milford, MA). Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer. Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA). UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (Agilent Technologies, Santa Clara, CA). Matrix assisted laser desorption/ionization with time-of-flight

detector (MALDI-TOF) mass spectra were acquired on a Bruker Ultraflex III instrument (Billerica, MA). Analytical HPLC was performed on an Agilent 1100 Series HPLC system (Santa Clara, CA). Preparative HPLC was performed on a Varian Prostar HPLC system (Agilent Technologies, Santa Clara, CA). HPLC columns were purchased from W. R. Grace & Compnay (Columbia, MD).

Synthesis of *N*,*N*'-di-Boc-*L*-Selenocystine (2) **2** was synthesized by Boc protect-tion of *L*-selenocystine **1** as described in Chapter 3.

Synthesis of Boc-Sec(*S*-*t*Bu)-OH (3c) Tetrahydrofuran (THF) and triethylamine (Et_3N) were freshly degassed by the freeze-pump-thaw method. **2** (0.2072 g, 0.4 mmol, 1 equiv) was dissolved in 6.5 mL of degassed THF, and then Et_3N (541 μL, 4 mmol, 10 equiv) and 2-methyl-2-propanethiol (*t*BuSH, 1.1 mL, 10 mmol, 25 equiv) were added. The reaction was allowed to proceed at room temperature under an argon atmosphere for 4 h. Upon completion, the reaction mixture was concentrated by rotary evaporation. The residue was brought up in 10 mL of H_2O, acidified with 1 mL of 3 M HCl, and then extracted with ethyl acetate (10 mL × 3). The organic layers were combined and concen-trated by rotary evaporation, and then purified by flash chromatography in 5:5 ethyl acetate/petroleum ether with 0.1% acetic acid. A pale yellow crystalline solid was isolated as the final product (0.2523 g, 0.7 mmol, 91% yield). R_f = 0.25 in 5:5 ethyl acetate/ petroleum ether with 0.1% acetic acid. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 11.16 (s, 1H), 5.44 (d, *J* = 7.1 Hz, 1H), 4.64 (d, *J* = 3.7 Hz, 1H), 3.30 (dd, *J* = 11.9, 4.1 Hz, 1H), 3.22 (dd, *J* = 12.1, 5.8 Hz, 1H), 1.44 (s, 9H), 1.34 (s, 9H). ^{13}C-NMR (500 MHz, $CDCl_3$):

δ 175.61, 155.69, 82.52, 53.78, 46.57, 34.39, 30.84, 28.61. ESI$^-$-HRMS: calculated for $C_{12}H_{22}NO_4SSe^-$: 356.0435; found $[M - H]^-$: 356.0442.

Synthesis of Boc-Sec(*S*-*t*Bu)-OCH$_2$CN (4c) Tetrahydrofuran (THF) and diisopropylethylamine (DIPEA) were freshly degassed by the freeze-pump-thaw method. **3c** (0.2194 g, 0.64 mmol, 1 equiv) was dissolved in degassed THF (7.7 mL), and chilled to 0 ˚C on ice. DIPEA (335 μL, 1.9 mmol, 3 equiv) was added, followed by $ClCH_2CN$ (2.0 mL, 32 mmol, 50 equiv). The reaction was allowed to proceed on ice for 2 h, and then at room temperature overnight (≥ 12 hr) under an argon atmosphere. Upon completion, the reaction mixture was concentrated by rotary evaporation, and then purified by flash chromatography in 5:5 ethyl acetate/petroleum ether. The reactant was recovered in 5:5 ethyl acetate/petroleum ether with 0.1% acetic acid, and then subject to two additional rounds of $ClCH_2CN$ activation. A pale yellow solid was isolated as the product (0.1396 g, 0.35 mmol, 57% yield). R_f = 0.74 in 5:5 ethyl acetate/ petroleum ether. ^{1}H-NMR (500 MHz, $CDCl_3$): δ 5.39 (d, *J* = 6.9 Hz, 1H), 4.79 (s, 2H), 4.72 (dd, *J* = 12.4, 5.7 Hz, 1H), 3.24 (dd, *J* = 13.2, 4.8 Hz, 1H), 3.20 (dd, *J* = 13.4, 5.7 Hz, 1H), 1.45 (s, 9H), 1.36 (s, 9H). ^{13}C-NMR (500 MHz, $CDCl_3$): δ 170.04, 155.29, 114.15, 81.01, 53.53, 49.48, 47.00, 33.55, 30.83, 28.59. ESI$^+$-HRMS: calculated for $C_{14}H_{24}N_2O_4SSeNa^+$: 419.0520; found $[M + Na]^+$: 419.0519.

Synthesis of Boc-Sec(*S*-*t*Bu)-(5'-*O*-DMT)Adenosine (5c) Tetrahydrofuran (THF) was freshly degassed by the freeze-pump-thaw method. **4c** (0.1396 g, 0.35 mmol, 2 equiv) was dissolved in THF (4 mL). 5′-*O*-(4,4′-Dimethoxytrityl) adenosine (0.1006 g,

0.18 mmol, 1 equiv) was added, followed by catalytic amount of tetrabutylammonium acetate (0.0027 g, 9 μmol, 0.05 equiv). The reaction was allowed to proceed overnight (≥ 12 h) under an argon atmosphere. Upon completion, the reaction was concentrated by rotary evaporation, and then purified by flash chromatography in 1:19 methanol/ethyl acetate. A white foam was isolated as the final product (0.1029 g, 0.11 mmol, 65% yield). R_f = 0.42 in 1:19 methanol/ethyl acetate. ^{1}H-NMR and ^{13}C-NMR sepctra (500 MHz, d^8-THF) are shown below (Figure 4-10). ESI$^+$-HRMS: calculated for $C_{43}H_{53}N_6O_9SSe^+$: 909.2760; found $[M + H]^+$: 909.2767.

Synthesis of H-Sec(*S*-*t*Bu)-Ade (**6c**) Tetrahydrofuran (THF), CH_2Cl_2 and H_2O were freshly degassed by the freeze-pump-thaw method. **5c** (0.0200 g, 22 μmol, 1 equiv) was dissolved in degassed THF (1 mL) and chilled to 0 °C on ice. Triisopropylsilane (23 μL, 110 μmol, 5 equiv) was added, followed by tBuSH (205 μL, 2.2 mmol, 100 equiv). Trifluoroacetic acid (1 mL) was added dropwise, and then the reaction was allowed to proceed at room temperature for 2 h under an argon atmosphere. Upon completion, the reaction was dried by rotary evaporation. The residue was brought up in degassed CH_2Cl_2 (1 mL), and then extracted with degassed H_2O (1 mL × 3) to recover the product. The crude was directly purified by reverse phase HPLC using a binary solvent system of water/acetonitrile with 0.1% trifluoroacetic acid (Table 4-1 through Table 4-3).

Synthesis of Other H-Sec(SR)-Ade Donors (**6a-e**) All other adenosine donors were derivatized from **6c** through thiol exchange: H_2O was freshly degassed by the freeze-pump-thaw method. **6c** (0.0056 g, 11 μmol, 1 equiv) was dissolved in 1 mL of degassed

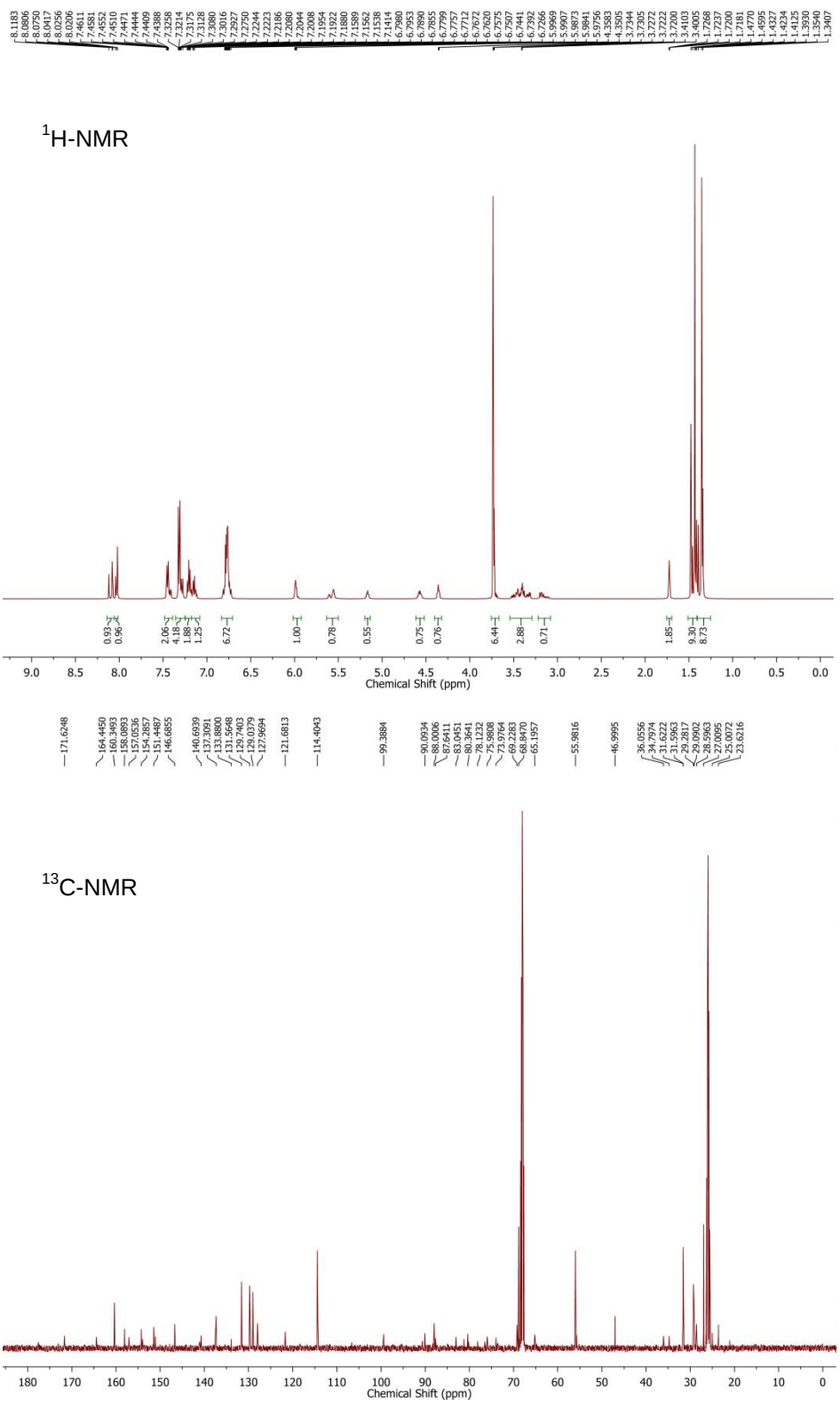


Figure 4-10. ^{1}H and ^{13}C NMR Characterization of Boc-Sec(*S*-*t*Bu)-(5'-*O*-DMT)Ade (**5c**)

H_2O. 100 µL of an appropriate thiol (ethanethiol for **6a**; 2-propanethiol for **6b**; thiophenol for **6d**; benzyl mercaptan for **6e**) was added, and the reaction was allowed to proceed at room temperature for 2 h under an argon atmosphere. Upon completion, excess thiol was evaporated by a stream of argon. The aqueous crude was directly injected onto reverse phase HPLC for purification (Table 4-1 through Table 4-3).

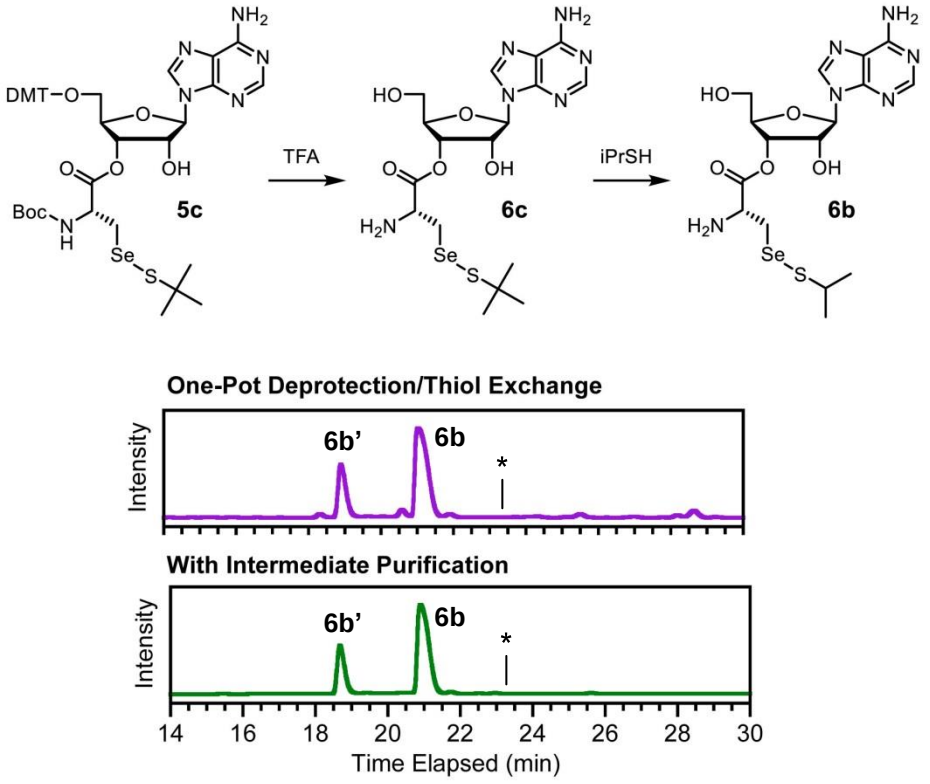


Figure 4-11. *i*PrSH Exchange with or without Intermediate Purification.

HPLC analysis of crude reaction mixtures monitored at 260 nm on a YMC-Pack Pro C8 semi-prep column using gradient **1**.In one-pot deprotection/thiol exchange, **5c** was deprotected, lyophilized, and then directly subject to thiol exchange without purification of **6c**. **6b'** denotes the 2'-acylated isomer of **6b** (see later sections); asterisk denotes the absence of **6c**.

We note that no major difference was observed when the thiol exchange reaction was conducted on either HPLC purified **6c**, or crude **6c** that had simply been lyophilized after

TFA deprotection of **5c** (Figure 4-11). Therefore, all bulk preparations of the adenosine donors were performed in a one-pot manner without the intermediate HPLC purification step, both to simplify the process and to avoid material loss during purification. MALDI-TOF MS and UV-Vis characterization of final products **6a-e** are shown in Figure 4-12.

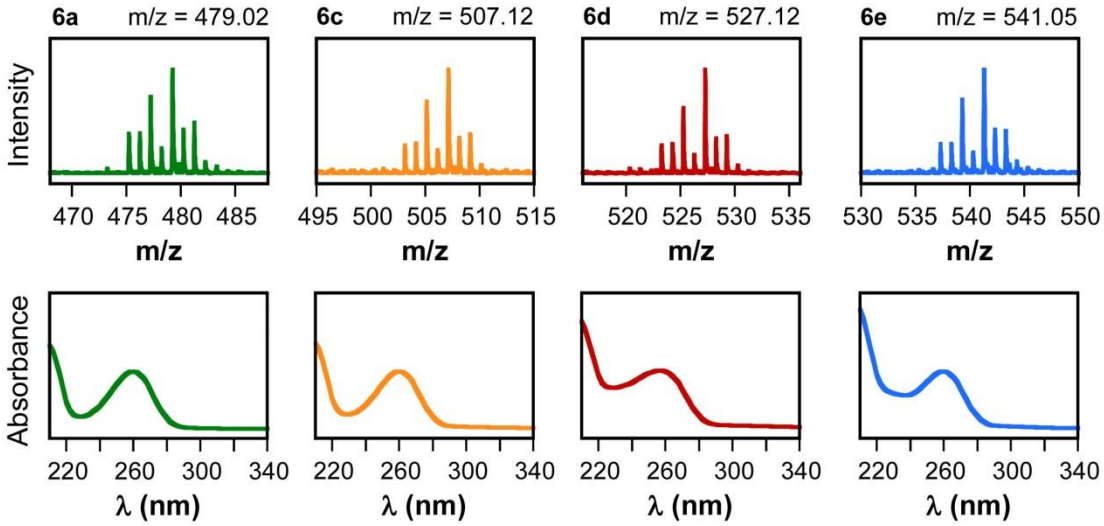


Figure 4-12. MALDI-TOF MS and UV-Vis Characterization of Adenosine Donor **6a-e**.

See Figure 4-7 for characterization of **6b**, and Table 4-3 for a complete list of expected and observes m/z values. The integrity of selenium and adenosine can be clearly observed in all products. **6d** and **6e** showed additional absorption around 230 nm, which is characteristic for the side chain phenyl/benzyl substitution.

Table 4-1. Adenosine Donor Purification Methods and Retention Time.

Compound	Gradient	Retention Time*	Column
H-Sec(*S*-Et)-Ade (**6a**)	**1**	16.2 / 18.2 min	YMC-Pack Pro C8 Semi-prep
H-Sec(*S*-*i*Pr)-Ade (**6b**)	**1**	19.2 / 20.8 min	YMC-Pack Pro C8 Semi-prep
H-Sec(*S*-*t*Bu)-Ade (**6c**)	**1**	21.1 / 23.4 min	YMC-Pack Pro C8 Semi-prep
H-Sec(*S*-Ph)-Ade (**6d**)	**1**	22.0 / 24.5 min	YMC-Pack Pro C8 Semi-prep
H-Sec(*S*-Bz)-Ade (**6e**)	**1**	22.7 / 25.0 min	YMC-Pack Pro C8 Semi-prep

* All adenosine donors are isolated as a mixture of interconverting 2'- and 3'-isomers.

Table 4-2. HPLC Gradients for Purification and Characterization.

No.	Time (min)	%B	**No.**	Time (min)	**%B**	**No.**	Time (min)	**%B**
1	0:00	2	**2**	0:00	1	**3**	0:00	2
	5:00	2		5:00	1		5:00	2
	10:00	10		10:00	30		25:00	20
	30:00	30		15:00	40		30:00	100
	35:00	100		20:00	100		35:00	100
	40:00	100		25:00	100		45:00	2
	45:00	2		27:00	1			
				30:00	1			

* Solvent A: 0.1% trifluoroacetic acid in water; Solvent B: 0.1% trifluoroacetic acid in acetonitrile

Table 4-3. MALDI-TOF MS Characterization of Purified Adenosine Donors.

Compound	**$[M + H]^+$**		**$[M + Na]^+$**	
	Calculated	**Found**	**Calculated**	**Found**
H-Sec(*S*-Et)-Ade (**6a**)	479.05	479.02	501.04	500.98
H-Sec(*S*-*i*Pr)-Ade (**6b**)	493.07	492.95	515.06	514.91
H-Sec(*S*-*t*Bu)-Ade (**6c**)	507.09	507.12	529.08	529.09
H-Sec(*S*-Ph)-Ade (**6d**)	527.05	527.12	-	-
H-Sec(*S*-Bz)-Ade (**6e**)	541.07	541.05	563.06	563.05

Characterization of 2'- and 3'-Isomers of Adenosine Donors As noted in our previous work[268], all adenosine donors were isolated as a mixture of 2'- and 3'-acylated isomers. To ascertain this observation in our hemiselenide protected Sec donors, we isolated the early and late peaks (assigned as 2'- and 3'-isomers based on the known thermodynamic preference for 3'-acylation of adenosine[301]) of H-Sec(*S*-*i*Pr)-Ade, re-injected them separately, and observed interconversion of the two species (Figure 4-13). Since these two isomers co-exist in rapid equilibrium, we reasoned that it would be unnecessary to obtain one pure isomer − as the 3'-isomer is consumed by AaT in chemo-

enzymatic reaction, the 2'-isomer will convert into the 3'-isomer by equilibrating, thus supplementing substrates for the reaction.

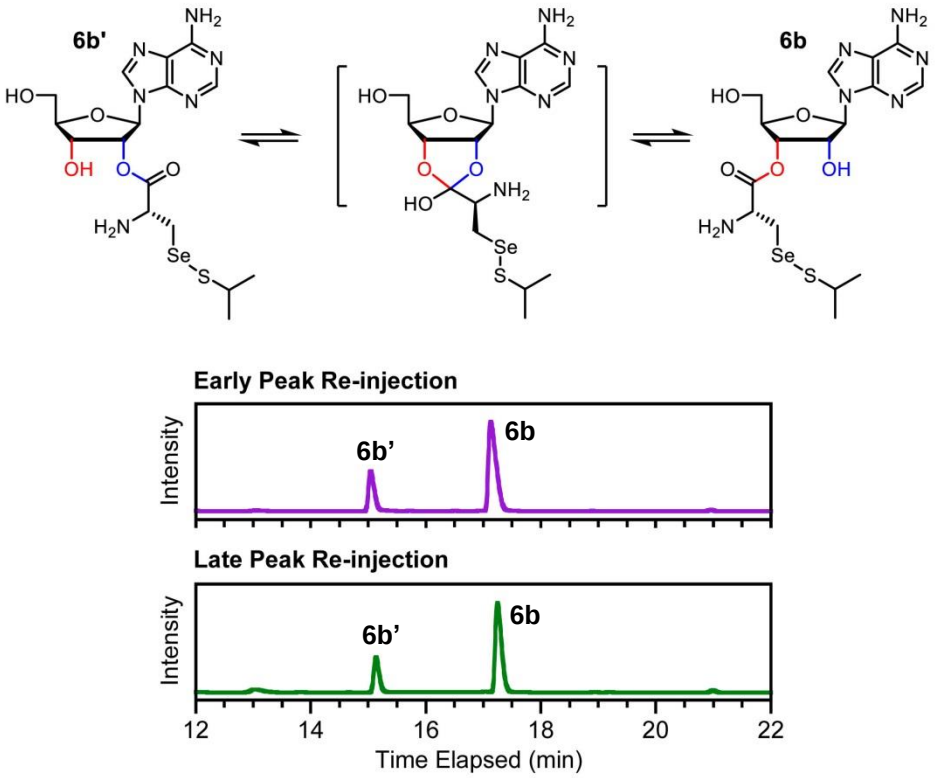


Figure 4-13. Interconversion of 2'- and 3'-Isomers of H-Sec(S-*i*Pr)-Ade **6b**.

Early peak (assigned as **6b'**) and late peak (assigned as **6b**) were isolated, brought up in H_2O, and re-injected onto reverse phase HPLC. Chromotogram monitored at 260 nm on a Jupiter C18 analytical column using gradient **1**. In both cases, the interconversion rapidly reached thermodynamic equilibrium of 1:3 **6b'**/**6b** under ambient conditions. MALDI-TOF MS: [**6b** + H]$^+$, expected 493.07, found 493.06; [**6b'** + H]$^+$, expected 493.07, found 493.05.

Versatility of Thiol Exchange on Hemiselenides To demonstrate the versatility of the thiol exchange reaction in derivatizing hemiselenides, we performed thiol exchange on H-Sec(*S*-*t*Bu)-Ade **6c** and H-Sec(*S*-Bz)-Ade **6e**, and showed that the two species could be interconverted simply by adding different thiols in excess (Figure 4-14).

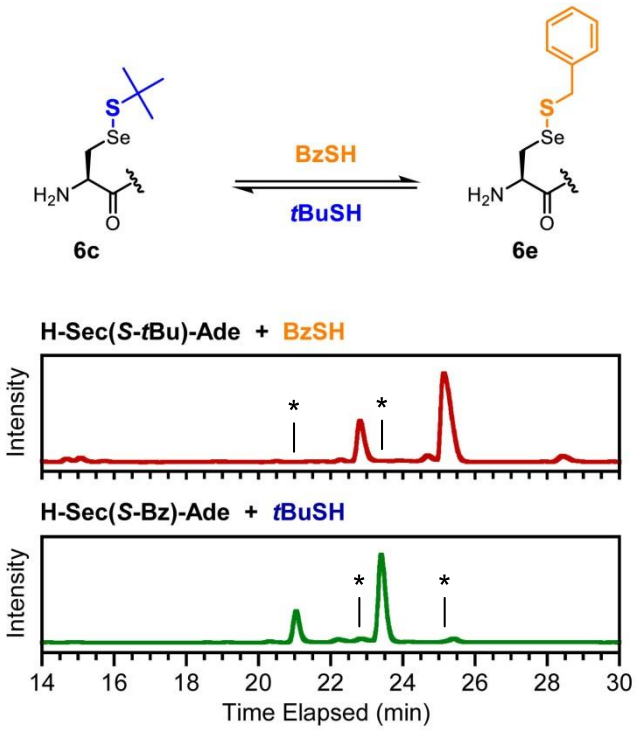


Figure 4-14. Interconversion of Adenosine Analogs as Driven by Excess of Different Thiols.

6c was readily converted into **6e** by treatment with excess BzSH; *vice versa*, **6e** could be generated from **6c** by treatment with excess *t*BuSH. HPLC chromatograms of crude reaction mixtures monitored at 260 nm on a YMC-Pack Pro C8 semi-prep column using gradient **1**. MALDI-TOF MS: [**6c** + H]$^+$, expected 507.12, found 507.08; [**6e** + H]$^+$, expected 541.05, found 541.07. Star signs indicate the absence of the original adenosine donor added.

Expression and Purification of *E. coli* Aminoacyl Transferase (AaT) His_{10}-tagged *E. coli* amino acyl transferase (AaT) was expressed from a pEG6 plasmid (ampicillin resistant) in *E. coli* BL21-Gold (DE3) cells using a procedure adapted from Graciet *et al.*[302] In particular, no β-mercaptoethanol (BME) was used in our purification, to avoid complications with the hemiselenide side chains. For protein expression, *E. coli* BL21-Gold (DE3) cells were transformed with the plasmid, and selected on an LB plate in the presence of ampicillin (Amp, 100 μg/mL) by overnight growth at 37 °C. 5 mL LB media was inoculated with a single colony, and then grown at 37 °C in the presence of Amp

(100 μg/mL) until OD600 ≥ 0.5. The primary culture was diluted into 1 L of LB media with Amp (100 mg/L), and grown at 37 °C until OD600 = 0.6. 100 mg of isopropyl *β*-D-thiogalactoside (IPTG, 0.1 mM final concentration) was added to induced AaT expression, and then the cells were grown at 25 °C for another 16 ~ 18 h.

For purification, cells were harvested at 6,000 rpm using a GS3 rotor on a Sorvall RC-5 centrifuge. Cell pellets were resuspended in Ni-NTA binding buffer (50 mM Tris, 10 mM imidazole, 300 mM KCl, pH 8.0 with protease inhibitor cocktail, 1 mM phenylmethanesulfonyl fluoride and 10 units/mL DNAse1–Grade II), and then lysed by sonication. The crude was centrifuged at 13,200 rpm for 15 min, and then supernatant was recovered and incubated with Ni-NTA resin for 1 h on ice with gentle shaking. The resin was collected by filtration, rinsed with 4 volumes of wash buffer (50 mM Tris, 50 mM imidazole, 300 mM KCl, pH 8.0), and then eluted with 8 volumes of elution buffer (50 mM Tris, 250 mM imidazole, 300 mM KCl, pH 8.0). Pure fractions of AaT were identified by SDS-PAGE, and then collected and dialyzed into a storage buffer (50 mM Tris, 30 % glycerol, 120 mM $(NH_4)_2SO_4$, pH 8.0) at 4 °C overnight. After concentration determination by Bradford assay[303], the purified enzymes were stored at -80 °C until use.

AaT Activity Assay for Adenosine Donor Screening H_2O, acetone, a 10x buffer stock (500 mM Tris, 1.5 M KCl, 100 mM $MgCl_2$, pH 8.0), and a LysAlaAcm **7** stock (10 mM in water) were freshly degassed by the freeze-pump-thaw method. Sec adenosine donor was brought up in degassed H_2O, and adjusted to 10 mM based on UV-Vis absorption (ε_{260} = 15,400 M^{-1} cm^{-1}). Under an argon atmosphere, H_2O (41.6 μL), buffer stock

(6.25 µL), LysAlaAcm stock (0.625 µL) and Sec adenosine donor stock (6.25 µL) were thoroughly mixed, and then AaT (7.81 µL, 0.8 mg/mL) was added to initiate the reaction. The reaction was allowed to proceed at 37 ˚C for 4 h under an argon atmosphere. Upon completion, the reaction was quenched with 1% acetic acid in degassed H_2O (187.5 µL). Degassed acetone (1 mL) was added, and then the mixture was kept at -20 °C for 1 h to precipitate the enzyme. Supernatant was recovered after centrifugation at 13,200 rpm and 4 ˚C for 20 min, and then excess acetone was removed by rotary evaporation. The residue was diluted into degassed H_2O (800 µL), and then analyzed by reverse phase HPLC on a Jupiter C18 analytical column using gradient **2**.

For quantification, peak identities were assigned based on MALDI-TOF MS (Table 4-4), and then transfer efficiencies were calculated from peak areas monitored at 325 nm. For each adenosine donor, the transfer efficiency was reported as the average value with standard deviation from three independent trials.

Table 4-4. Retention Time and MALDI-TOF MS Characterization of LysAlaAcm Peptides.

Peptide	Retention Time	$[M + H]^+$		$[M + Na]^+$	
		Calc'd	**Found**	**Calc'd**	**Found**
LysAlaAcm (**7**)	11.1 min	375.20	375.29	397.19	
Sec(*S*-Et)-LysAlaAcm (**8a**)	12.6 min	586.16	586.38	-	-
Sec(*S*-*i*Pr)-LysAlaAcm (**8b**)	13.1 min	600.17	600.14	622.16	622.11
Sec(*S*-*t*Bu)-LysAlaAcm (**8c**)	13.2 min	614.18	614.13	636.17	636.10
Sec(*S*-Ph)-LysAlaAcm (**8d**)	13.4 min	634.15	634.33	656.14	656.23
Sec(*S*-Bz)-LysAlaAcm (**8e**)	13.3 min	648.17	648.48	-	-
Phe-LysAlaAcm (**8h**)	12.4 min	522.26	522.33	544.25	544.31

* Retention time obtained on a Jupiter C18 analytical column using gradient **2**.

Oxygen Tolerance Reactions were conducted similarly as the standard procedure, except that non-degassed water and buffers were used (Figure 4-15). Transfer efficiencies were reported as the average value from three independent trials.

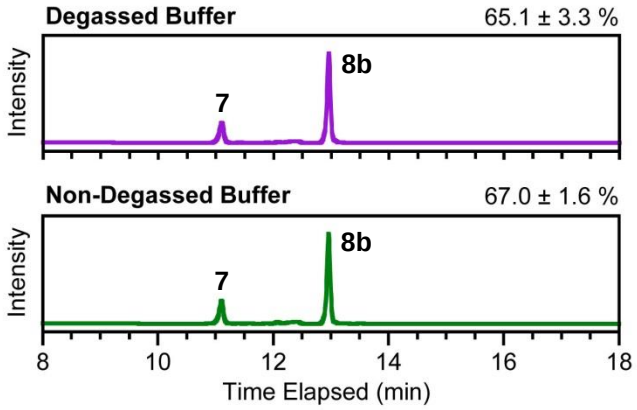


Figure 4-15. Oxygen Tolerance of Chemoenzymatic AaT Reaction with H-Sec(*S*-*i*Pr)-Ade.

Expression and Purification of $\alpha S_{6\text{-}140}$ Expression and purification were performed by Anne M. Wagner as detailed in *Angew. Chem. Int. Ed*. **2013**, *52*, 6210-6213. Briefly, a pET-16b plasmid encoding for His_{Tag}-$\alpha S_{6\text{-}140}$ was transformed into *E. coli* BL21 DE3 cells, and over-expressed with isopropyl-*β*-D-thiogalactopyranoside (IPTG) induction. After lysis of the cells, target protein was pulled down by Ni-NTA resin, and treated with Factor Xa to remove the His_{Tag}. The protein was dialyzed into storage buffer (20 mM Tris, 150 mM KCl, 10 mM Mg_2Cl_2, pH 8.0), quantified by BCA assay (Pierce), and stored at -80 °C until use.

AaT-Mediated Chemoenzymatic Transfer of Sec(*S*-*i*Pr) onto $\alpha S_{6\text{-}140}$ H-Sec(*S*-*i*Pr)-Ade **6b** was brought up in H_2O, and adjusted to 10 mM based on UV-Vis absorption

(ε_{260} = 15,400 M^{-1} cm^{-1}). H_2O (35.2 µL), 10x reaction buffer (6.25 µL, 500 mM HEPES, 1.5 M KCl, 100 M $MgCl_2$, pH 8.0), $\alpha S_{6\text{-}140}$ stock (6.94 µL, 0.9 mg/mL) and **6b** stock (6.25 µL) were thoroughly mixed. AaT stock (7.81 uL, 0.8 mg/mL) was added to initiate the reaction. The reaction was allowed to proceed at 37 ˚C for 4 h; an additional dose of **6b** stock (6.25 uL) was added each hour. Upon completion, the reaction was buffer-exchanged into H_2O using Amicon Ultra centrifugal filter units (3k Da MWCO). The resulting crude mixture was then directly analyzed by MALDI-TOF MS.

Deselenization of Sec(*S*-*i*Pr)-$\alpha S_{6\text{-}140}$ To confirm the presence of Sec, 50 µL of the above reaction crude was mixed with 150 µL of a TCEP stock (50 mM Tris, 20 mM TCEP, 10 mM DTT, pH 8.0). The reaction was allowed to proceed at 37 ˚C for 6 h under an argon atmosphere. Upon completion, the reaction was buffer-exchanged into H_2O using Amicon Ultra centrifugal filter units (3k Da MWCO), and analyzed by MALDI-TOF MS. TCEP = *tris*(2-carboxyethyl)phosphine; DTT = dithiothreitol.

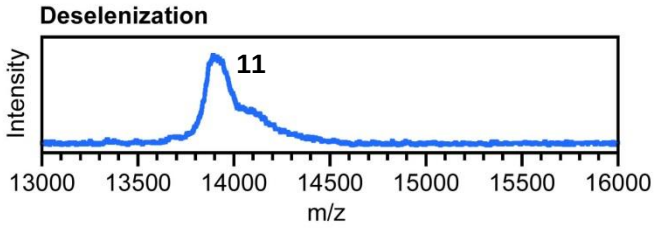


Figure 4-16. Deselenization of Sec(*S*-*i*Pr)-$\alpha S_{6\text{-}140}$ into Ala- $\alpha S_{6\text{-}140}$ **11**. MALDI-TOF MS: [**11** + H]$^+$, expected 13899.99, found 13893.67.

Synthesis of H-Sec(*S*-*i*Pr)-OH and H-Sec(*S*-Et)-OH (S1, S2) Tetrahydrofuran (THF) was freshly degassed by the freeze-pump-thaw method. *N*,*N*'-di-Boc-*L*-seleno-

cystine **2** (0.0730 g, 0.14 mmol, 1 equiv) was dissolved in degassed THF (1.6 mL), and then ethanethiol (246 μL, 3.4 mmol, 25 equiv) or 2-propanethiol (316 μL, 3.4 mmol, 25 equiv) was added to initiate the thiolysis. The reaction was allowed to proceed at room temperature for 4 h under an argon atmosphere. Upon completion, trifluoroacetic acid (2 mL) was added dropwise to initiate the Boc deprotection. The reaction was allowed to proceed for 45 min under an argon atmosphere, and then the crude product was recovered by ether precipitation. The residue was brought up in 1:10 CH_3CN/H_2O, and then purified by reverse phase HPLC on a Vydac C8 semi-prep column using gradient **3**. Retention time: 21.3 min for **S1**; 14.2 min for **S2**. ESI^+-LRMS: calc'd for $C_6H_{14}NO_2SSe^+$ 243.98, found [**S1** + H]$^+$ 244.01; calc'd for $C_5H_{12}NO_2SSe^+$ 229.97, found [**S2** + H]$^+$ 230.02

Expression and Purification of Synthetases Expression and purification of His_6-tagged *S. cerevisiae* phenylalaninyl synthetase (PheRS) and *E. coli* leucyl synthetase (LeuRS, with a T252F mutation to block the editing site[304]) were conducted using a similar procedure as AaT expression and purification, except that different buffers were used. For PheRS: binding buffer, 50 mM Tris, 300 mM KCl, pH 8.0; wash buffer, 50 mM Tris, 20 mM imidazole, 300 mM KCl, pH 8.0; elution buffer, 50 mM Tris, 40 mM imidazole, 300 mM KCl, pH 8.0; storage buffer, 25 mM Tris, 50% glycerol, pH 7.6. For LeuRS: binding buffer, 50 mM Tris, 10 mM imidazole, 300 mM KCl, pH 8.0; wash buffer, 25 mM Tris, 20 mM imidazole, 300 mM KCl, pH 8.0; elution buffer, 50 mM Tris, 100 mM imidazole, 300 mM KCl, pH 8.0; storage buffer, 50 mM Tris, 50% glycerol, pH 7.6. *E. coli* methionine synthetase (MetRS) and its mutant MetRS L13G were expressed and purified by John B. Warner following literature precedence by the Tirrell group[264].

Synthetase Activity Profiling in AaT-Mediate N-Terminal Modification Synthetase activities were tested in a fully enzymatic transfer assay using LysAlaAcm **7** as the model peptide and H-Sec(*S*-*i*Pr)-OH **S1** as the amino acid substrate; in the case of MetRS and its L13G mutant, H-Sec(*S*-Et)-OH **S2** was also tested since it more closely resembles the natural substrate Met. For each synthetase/mutant, a positive control with its natural substrate and two negative controls (withholding amino acid or tRNA) were also included as reference points (Figure 4-17). For a typical assay: a 15 mM H-Sec(SR)-OH amino acid stock was freshly prepared from **S1** or **S2** in H_2O. Milli-Q H_2O (21.8 µL), 10x buffer stock (6.25 µL, 500 mM HEPES, 1.5 M KCl, 100 mM Mg_2Cl, pH 8.0), LysAlaAcm **7** stock (0.625 µL, 10 mM in H_2O), Sec amino acid stock (4.17 µL) and ATP stock (3.13 µL, 50 mM in H_2O) were thoroughly mixed. tRNA (12.5 µL, 200 µM), AaT (7.81 µL, 0.8 mg/mL) and synthetase (6.25 µL, 1 mg/mL) were then added to initiate the reaction. The reaction was allowed to proceed at 37 ˚C for 4 h under an argon atmosphere, and then worked up and analyzed similarly as the standard AaT assay. Unfortunately, none of the synthetases showed activity towards the H-Sec(SR)-OH amino acids tested.

Rationally Designed LeuRS Mutants Based on crystal structure[305] and previous mutagenesis work by Schultz *et al.*[306], we rationally designed five LeuRS mutants with various permutations of the amino acid binding pocket and tested them for activity in AaT-mediated N-terminal modification assays. Unfortunately, none of these mutants showed the desired activity. Mutants and primers are listed in Table 4-5 and Table 4-6.

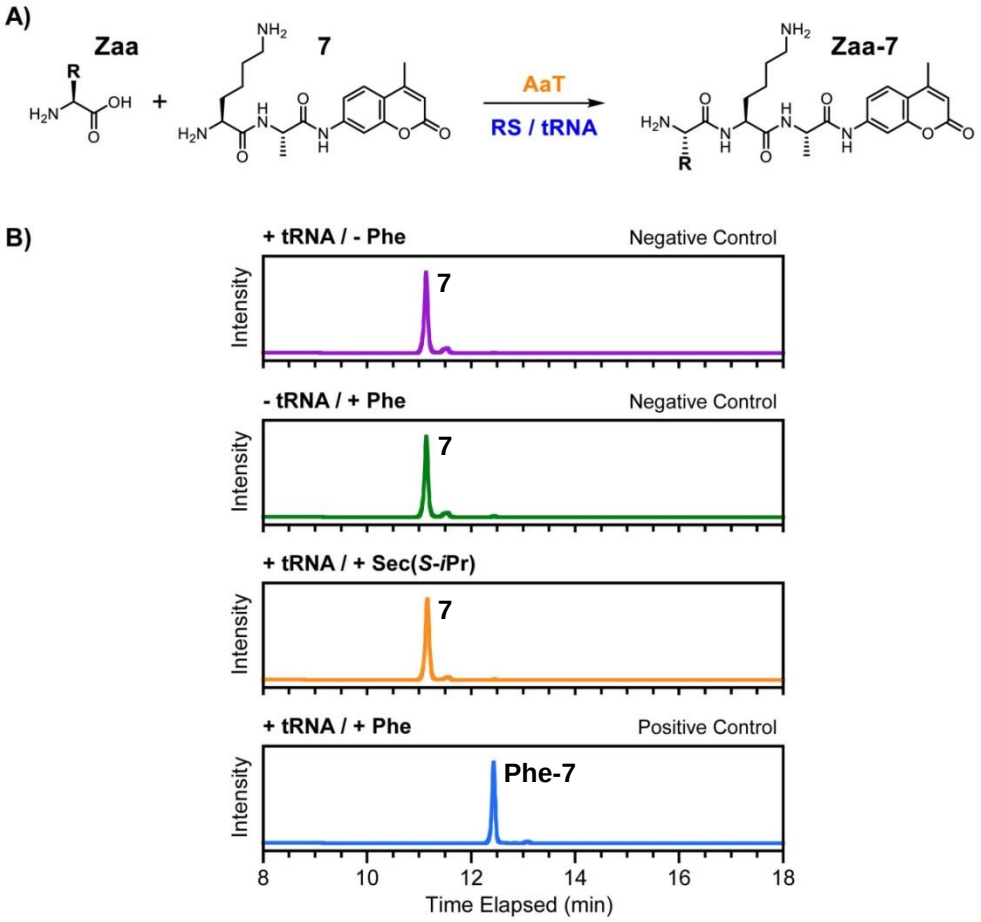


Figure 4-17. Synthetase (RS) Activity Profiling Assay using PheRS as an Example.

A) Assay scheme. B) HPLC chromatograms monitored at 325 nm. MALDI-TOF MS: [**7** + H]$^+$, expected 375.20, found 375.20; [**Phe-7** + H]$^+$, expected 522.26, found 522.25.

Table 4-5. Rationally Designed LeuRS Mutants.

Mutant	M40	T252	Y499	Y527	H533	H537	Primers Used
pWT	-	T252F	-	-	-	-	-
1	-	T252F	Y499S	-	-	H537G	2, 4
2	M40A	T252F	-	-	H533A	H537G	1, 5
3	-	T252F	-	Y527T	H533A	H537G	3, 5
4	-	T252F	Y499S	Y527T	H533A	-	2, 3, 4
5	-	T252F	Y499S	Y527T	H533A	H537G	2, 3, 5

* A T252Y mutation was included in all mutants to block the editing site[304].

Table 4-6. Primers for LeuRS Mutagenesis.

No.	Mutation		Primer Sequence
1	M40A	Forward	5'- G AAG TAT TAC TGC CTG TCT **GCG** CTT CCC TAT CCT TCT GGT C -3'
		Reverse	5'- G ACC AGA AGG ATA GGG AAG **CGC** AGA CAG GCA GTA ATA CTT C -3'
2	Y499S	Forward	5'- TTT ATG GAG TCC TCC TGG **TCT** TAT GCG CGC TAC ACT TGC -3'
		Reverse	5'- GCA AGT GTA GCG CGC ATA **AGA** CCA GGA GGA CTC CAT AAA -3'
3	Y527T	Forward	5'-TAC TGG CTG CCG GTG GAT ATC **ACC** ATT GGT GGT ATT GAA-3'
		Reverse	5'-TTC AAT ACC ACC AAT **GGT** GAT ATC CAC CGG CAG CCA GTA-3'
4	H533/H537G	Forward	5'- ATT GAA CAC GCC ATT ATG **GGC** CTG CTC TAC TTC CGC TTC -3'
		Reverse	5'- GAA GCG GAA GTA GAG CAG **GCC** CAT AAT GGC GTG TTC AAT -3'
5	H533A/H537G	Forward	5'- C ATT GGT GGT ATT GAA **GCG** GCC ATT ATG **GGC** CTG CTC TAC TTC CGC -3'
		Reverse	5'- GCG GAA GTA GAG CAG **GCC** CAT AAT GGC **CGC** TTC AAT ACC ACC AAT G -3'

* Desired mutations are highlighted in bold.

4.6 Acknowledgement

This work was supported by funding from the University of Pennsylvania, Searle Scholars Program (10-SSP-214 to EJP), and the National Institute of Health (NIH NS081033). Instruments supported by the National Science Foundation and National Institutes of Health include: HRMS (NIH RR-023444) and MALDI-MS (NSF MRI-0820996). We thank Dr. Rakesh Kohli for assistance with HRMS analysis, and Michael Nicastri for assistance with LeuRS mutagenesis during his rotation.

BIBLIOGRAPHY

1. Voet, D.; Voet, J. G., *Biochemistry*. John Wiley & Sons: Hoboken, NJ, 2011.

2. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P., *Molecular Biology of the Cell*. Garland Science: New York, NY, 2007.

3. Creighton, T. E., *Proteins*. W. H. Freeman: New York, NY, 1990.

4. Crick, F., Central dogma of molecular biology. *Nature* **1970,** *227* (5258), 561–563.

5. Anfinsen, C. B., Principles that govern the folding of protein chains. *Science* **1973,** *181* (4096), 223–230.

6. Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G., Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997,** *48* (1), 545−600.

7. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **2009,** *5* (11), 789–796.

8. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964−972.

9. Li, C.; Liu, M., Protein dynamics in living cells studied by in-cell NMR spectroscopy. *FEBS Lett.* **2013,** *587* (8), 1008−1011.

10. Stefani, M., Protein folding and misfolding on surfaces. *Int. J. Mol. Sci.* **2008,** *9* (12), 2515−2542.

11. Valastyan, J. S.; Lindquist, S., Mechanisms of protein-folding diseases at a glance. *Dis. Models Mech.* **2014,** *7* (1), 9−14.

12. Soto, C., Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.* **2003,** *4* (1), 49−60.

13. Mukherjee, A.; Morales-Scheihing, D.; Butler, P. C.; Soto, C., Type 2 diabetes as a protein misfolding disease. *Trends Mol. Med.* **2015,** *21* (7), 439−449.

14. Gregersen, N.; Bross, P.; Vang, S.; Christensen, J. H., Protein misfolding and human disease. *Ann. Rev. Genom. Hum. G.* **2006,** *7* (1), 103−124.

15. Herczenik, E.; Gebbink, M. F. B. G., Molecular and cellular aspects of protein misfolding and disease. *FASEB J.* **2008,** *22* (7), 2115−2133.

16. Uversky, V. N., Fink, Anthony, *Protein Misfolding, Aggregation and Conformational Diseases*. Springer: Singapore, 2007.

17. Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M., The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15* (6), 384−396.

18. Jucker, M.; Walker, L. C., Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **2013,** *501* (7465), 45−51.

19. Nordstedt, C.; Näslund, J.; Tjernberg, L. O.; Karlström, A. R.; Thyberg, J.; Terenius, L., The Alzheimer Aβ peptide develops protease resistance in association with its polymerization into fibrils. *J. Biol. Chem.* **1994,** *269* (49), 30773−30776.

20. de Lau, L. M. L.; Breteler, M. M. B., Epidemiology of Parkinson's disease. *Lancet Neurol.* **2006,** *5* (6), 525–535.

21. Parkinson, J., *An Essay on the Shaking Palsy*. Whittingham and Rowland: London, UK, 1817.

22. Longo, D.; Fauci, A.; Kasper, D.; Hauser, S.; Jameson, J.; Loscalzo, J., *Harrison's Principles of Internal Medicine*. McGraw-Hill Education: New York, NY, 2015.

23. Hornykiewicz, O., A brief history of levodopa. *J. Neurol.* **2010,** *257* (2), 249−252.

24. Rodrigues e Silva, A. M.; Geldsetzer, F.; Holdorff, B.; Kielhorn, F. W.; Balzer-Geldsetzer, M.; Oertel, W. H.; Hurtig, H.; Dodel, R., Who was the man who discovered the "Lewy bodies"? *Movement Disord.* **2010,** *25* (12), 1765−1773.

25. Polymeropoulos, M. H.; Lavedan, C.; Leroy, E.; Ide, S. E.; Dehejia, A.; Dutra, A.; Pike, B.; Root, H.; Rubenstein, J.; Boyer, R.; Stenroos, E. S.; Chandrasekharappa, S.; Athanassiadou, A.; Papapetropoulos, T.; Johnson, W. G.; Lazzarini, A. M.; Duvoisin, R. C.; Di Iorio, G.; Golbe, L. I.; Nussbaum, R. L., Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science* **1997,** *276* (5321), 2045−2047.

26. Braak, H.; Tredici, K. D.; Rüb, U.; de Vos, R. A. I.; Jansen Steur, E. N. H.; Braak, E., Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **2003,** *24* (2), 197–211.

27. Jellinger, K. A., A critical evaluation of current staging of α-synuclein pathology in Lewy body disorders. *BBA - Mol. Basis Dis.* **2009,** *1792* (7), 730−740.

28. Stefanis, L., α-Synuclein in Parkinson's disease. *Cold Spring Harb. Perspect. Med.* **2012,** *2* (2), a009399.

29. Burré, J.; Sharma, M.; Tsetsenis, T.; Buchman, V.; Etherton, M. R.; Südhof, T. C., α-Synuclein promotes SNARE-complex assembly *in vivo* and *in vitro*. *Science* **2010,** *329* (5999), 1663–1667.

30. Fortin, D. L.; Nemani, V. M.; Voglmaier, S. M.; Anthony, M. D.; Ryan, T. A.; Edwards, R. H., Neural activity controls the synaptic accumulation of α-synuclein. *J. Neurosci.* **2005,** *25* (47), 10913−10921.

31. Pfefferkorn, C. M.; Jiang, Z.; Lee, J. C., Biophysics of α-synuclein membrane interactions. *BBA - Biomembranes* **2012,** *1818* (2), 162−171.

32. Kruger, R.; Kuhn, W.; Muller, T.; Woitalla, D.; Graeber, M.; Kosel, S.; Przuntek, H.; Epplen, J. T.; Schols, L.; Riess, O., Ala30Pro mutation in the gene encoding α-synuclein in Parkinson's disease. *Nat. Genet.* **1998,** *18* (2), 106−108.

33. Zarranz, J. J.; Alegre, J.; Gómez-Esteban, J. C.; Lezcano, E.; Ros, R.; Ampuero, I.; Vidal, L.; Hoenicka, J.; Rodriguez, O.; Atarés, B.; Llorens, V.; Tortosa, E. G.; del Ser, T.; Muñoz, D. G.; de Yebenes, J. G., The new mutation, E46K, of α-synuclein causes parkinson and Lewy body dementia. *Ann. Neurol.* **2004,** *55* (2), 164−173.

34. Appel-Cresswell, S.; Vilarino-Guell, C.; Encarnacion, M.; Sherman, H.; Yu, I.; Shah, B.; Weir, D.; Thompson, C.; Szu-Tu, C.; Trinh, J.; Aasly, J. O.; Rajput, A.; Rajput, A. H.; Jon Stoessl, A.; Farrer, M. J., Alpha-synuclein p.H50Q, a novel pathogenic mutation for Parkinson's disease. *Movement Disord.* **2013,** *28* (6), 811–813.

35. Lesage, S.; Anheim, M.; Letournel, F.; Bousset, L.; Honoré, A.; Rozas, N.; Pieri, L.; Madiona, K.; Dürr, A.; Melki, R.; Verny, C.; Brice, A., G51D α-synuclein mutation causes a novel Parkinsonian–pyramidal syndrome. *Ann. Neurol.* **2013,** *73* (4), 459−471.

36. Ulmer, T. S.; Bax, A.; Cole, N. B.; Nussbaum, R. L., Structure and dynamics of micelle-bound human α-synuclein. *J. Biol. Chem.* **2005,** *280* (10), 9595−9603.

37. Weinreb, P. H.; Zhen, W.; Poon, A. W.; Conway, K. A.; Lansbury, P. T., NACP, a protein implicated in Alzheimer's disease and learning, Is natively unfolded. *Biochemistry* **1996,** *35* (43), 13709−13715.

38. Serpell, L. C.; Berriman, J.; Jakes, R.; Goedert, M.; Crowther, R. A., Fiber diffraction of synthetic α-synuclein filaments shows amyloid-like cross-β conformation. *P. Natl. Acad. Sci. USA* **2000,** *97* (9), 4897−4902.

39. Jao, C. C.; Der-Sarkissian, A.; Chen, J.; Langen, R., Structure of membrane-bound α-synuclein studied by site-directed spin labeling. *P. Natl. Acad. Sci. USA* **2004,** *101* (22), 8331−8336.

40. Middleton, E. R.; Rhoades, E., Effects of curvature and composition on α-synuclein binding to lipid vesicles. *Biophys. J.* **2010,** *99* (7), 2279−2288.

41. Bartels, T.; Ahlstrom, L. S.; Leftin, A.; Kamp, F.; Haass, C.; Brown, M. F.; Beyer, K., The N-terminus of the intrinsically disordered protein α-synuclein triggers membrane binding and helix folding. *Biophys. J.* **2010,** *99* (7), 2116–2124.

42. Uversky, V. N.; Li, J.; Fink, A. L., Evidence for a partially folded intermediate in α-synuclein fibril formation. *J. Biol. Chem.* **2001,** *276* (14), 10737−10744.

43. Bertoncini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M., Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. *P. Natl. Acad. Sci. USA* **2005,** *102* (5), 1430–1435.

44. Lee, J. C.; Gray, H. B.; Winkler, J. R., Tertiary contact formation in α-synuclein probed by electron transfer. *J. Am. Chem. Soc.* **2005,** *127* (47), 16388−16389.

45. McClendon, S.; Rospigliosi, C. C.; Eliezer, D., Charge neutralization and collapse of the C-terminal tail of α-synuclein at low pH. *Protein Sci.* **2009,** *18* (7), 1531−1540.

46. Giasson, B. I.; Uryu, K.; Trojanowski, J. Q.; Lee, V. M.-Y., Mutant and wild type human α-synucleins assemble into elongated filaments with distinct morphologies *in vitro*. *J. Biol. Chem.* **1999,** *274* (12), 7619−7622.

47. Vilar, M.; Chou, H.-T.; Lührs, T.; Maji, S. K.; Riek-Loher, D.; Verel, R.; Manning, G.; Stahlberg, H.; Riek, R., The fold of α-synuclein fibrils. *P. Natl. Acad. Sci. USA* **2008,** *105* (25), 8637−8642.

48. Heise, H.; Hoyer, W.; Becker, S.; Andronesi, O. C.; Riedel, D.; Baldus, M., Molecular-level secondary structure, polymorphism, and dynamics of full-length α-synuclein fibrils studied by solid-state NMR. *P. Natl. Acad. Sci. USA* **2005,** *102* (44), 15871−15876.

49. 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., Structured regions of α-synuclein fibrils include the early-onset Parkinson's disease mutation sites. *J. Mol. Biol.* **2011,** *411* (4), 881–895.

50. Gath, J.; Habenstein, B.; Bousset, L.; Melki, R.; Meier, B.; Böckmann, A., Solid-state NMR sequential assignments of α-synuclein. *Biomol. NMR Assign.* **2012,** *6* (1), 51−55.

51. Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; Jiang, L.; Arbing, M. A.; Nannenga, B. L.; Hattne, J.; Whitelegge, J.; Brewster, A. S.; Messerschmidt, M.; Boutet, S.; Sauter, N. K.; Gonen, T.; Eisenberg, D. S., Structure of the toxic core of α-synuclein from invisible crystals. *Nature* **2015,** *525* (7570), 486−490.

52. Pornsuwan, S.; Giller, K.; Riedel, D.; Becker, S.; Griesinger, C.; Bennati, M., Long-range distances in amyloid fibrils of α-synuclein from PELDOR spectroscopy. *Angew. Chem. Int. Ed.* **2013,** *52* (39), 10290−10294.

53. Eisenberg, D.; Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Sambashivan, S.; Ivanova, M. I.; Madsen, A.; Riekel, C., The structural biology of protein aggregation diseases: fundamental questions and some answers. *Accounts Chem. Res.* **2006,** *39* (9), 568–575.

54. Biancalana, M.; Koide, S., Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *BBA - Proteins Proteom.* **2010,** *1804* (7), 1405–1412.

55. Wood, S. J.; Wypych, J.; Steavenson, S.; Louis, J.-C.; Citron, M.; Biere, A. L., α-Synuclein fibrillogenesis is nucleation-dependent: implications for the pathogenesis of Parkinson's disease. *J. Biol. Chem.* **1999,** *274* (28), 19509−19512.

56. Xue, W.-F.; Hellewell, A. L.; Gosal, W. S.; Homans, S. W.; Hewitt, E. W.; Radford, S. E., Fibril fragmentation enhances amyloid cytotoxicity. *J. Biol. Chem.* **2009,** *284* (49), 34272−34282.

57. Mahul-Mellier, A. L.; Vercruysse, F.; Maco, B.; Ait-Bouziad, N.; De Roo, M.; Muller, D.; Lashuel, H. A., Fibril growth and seeding capacity play key roles in α-synuclein-mediated apoptotic cell death. *Cell Death Differ.* **2015**.

58. Hoyer, W.; Antony, T.; Cherny, D.; Heim, G.; Jovin, T. M.; Subramaniam, V., Dependence of α-synuclein aggregate morphology on solution conditions. *J. Mol. Biol.* **2002,** *322* (2), 383−393.

59. Outeiro, T. F.; Klucken, J.; Bercury, K.; Tetzlaff, J.; Putcha, P.; Oliveira, L. M. A.; Quintas, A.; McLean, P. J.; Hyman, B. T., Dopamine-induced conformational changes in α-synuclein. *PLoS ONE* **2009,** *4* (9), e6906.

60. Lee, H.-J.; Baek, S. M.; Ho, D.-H.; Suk, J.-E.; Cho, E.-D.; Lee, S.-J., Dopamine promotes formation and secretion of non-fibrillar α-synuclein oligomers. *Exp. Mol. Med.* **2011,** *43*, 216−222.

61. Chan, T.; Chow, A. M.; Cheng, X. R.; Tang, D. W. F.; Brown, I. R.; Kerman, K., Oxidative stress effect of dopamine on α-synuclein: electroanalysis of solvent interactions. *ACS Chem. Neurosci.* **2012,** *3* (7), 569–574.

62. Venda, L. L.; Cragg, S. J.; Buchman, V. L.; Wade-Martins, R., α-Synuclein and dopamine at the crossroads of Parkinson's disease. *Trends Neurosci.* **2010,** *33* (12), 559−568.

63. Luk, K. C.; Kehm, V.; Carroll, J.; Zhang, B.; O'Brien, P.; Trojanowski, J. Q.; Lee, V. M. Y., Pathological α-synuclein transmission initiates Parkinson-like neurodegeneration in non-transgenic mice. *Science* **2012,** *338* (6109), 949−953.

64. Volpicelli-Daley, Laura A.; Luk, Kelvin C.; Patel, Tapan P.; Tanik, Selcuk A.; Riddle, Dawn M.; Stieber, A.; Meaney, David F.; Trojanowski, John Q.; Lee, Virginia M. Y., Exogenous α-synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* **2011,** *72* (1), 57−71.

65. Luk, K. C.; Song, C.; O'Brien, P.; Stieber, A.; Branch, J. R.; Brunden, K. R.; Trojanowski, J. Q.; Lee, V. M.-Y., Exogenous α-synuclein fibrils seed the formation of Lewy body-like intracellular inclusions in cultured cells. *P. Natl. Acad. Sci. USA* **2009,** *106* (47), 20051−20056.

66. Guo, Jing L.; Covell, Dustin J.; Daniels, Joshua P.; Iba, M.; Stieber, A.; Zhang, B.; Riddle, Dawn M.; Kwong, Linda K.; Xu, Y.; Trojanowski, John Q.; Lee, Virginia M. Y., Distinct α-synuclein strains differentially promote tau inclusions in neurons. *Cell* **2013,** *154* (1), 103−117.

67. Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W.-M.; Mattson, M. P.; Tycko, R., Self-propagating, molecular-level polymorphism in Alzheimer's ß-amyloid fibrils. *Science* **2005,** *307* (5707), 262−265.

68. Krishnan, V. V.; Rupp, B., Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy. In *eLS*, John Wiley & Sons: New York, NY, 2001.

69. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. Springer: New York, NY, 2006.

70. Trexler, A. J.; Rhoades, E., Single molecule characterization of α-synuclein in aggregation-prone states. *Biophys. J.* **2010,** *99* (9), 3048−3055.

71. Cremades, N.; Cohen, Samuel I.; Deas, E.; Abramov, Andrey Y.; Chen, Allen Y.; Orte, A.; Sandal, M.; Clarke, Richard W.; Dunne, P.; Aprile, Francesco A.; Bertoncini, Carlos W.; Wood, Nicholas W.; Knowles, Tuomas P.; Dobson, Christopher M.; Klenerman, D., Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell* **2012,** *149* (5), 1048–1059.

72. Demchenko, A. P., *Advanced Fluorescence Reporters in Chemistry and Biology*. Springer: New York, NY, 2010.

73. Choudhary, A.; Raines, R. T., An evaluation of peptide-bond isosteres. *ChemBioChem* **2011,** *12* (12), 1801–1807.

74. Petersson, E. J.; Goldberg, J. M.; Wissner, R. F., On the use of thioamides as fluorescence quenching probes for tracking protein folding and stability. *Phys. Chem. Chem. Phys.* **2014,** *16* (15), 6827−6837.

75. Bondi, A., van der Waals volumes and radii. *J. Phys. Chem. - US* **1964,** *68* (3), 441–451.

76. Truter, M. R., An accurate determination of the crystal structure of thioacetamide. *J. Chem. Soc.* **1960**, 997−1007.

77. Lippert, E., The strengths of chemical bonds. *Angew. Chem.* **1960,** *72* (16), 602.

78. Lee, C. M.; Kumler, W. D., The dipole moment and structure of thiolactams. *J. Org. Chem.* **1962,** *27* (6), 2052−2054.

79. Wiberg, K. B.; Rablen, P. R., Why does thioformamide have a larger rotational barrier than formamide? *J. Am. Chem. Soc.* **1995,** *117* (8), 2201−2209.

80. Pauling, L., *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. Cornell University Press: 1960.

81. Dudek, E. P.; Dudek, G. O., Proton magnetic resonance spectra of thiocarboxamides. *J. Org. Chem.* **1967,** *32* (3), 823–824.

82. Min, B. K.; Lee, H.-J.; Choi, Y. S.; Park, J.; Yoon, C.-J.; Yu, J.-A., A comparative study on the hydrogen bonding ability of amide and thioamide using near IR spectroscopy. *J. Mol. Struct.* **1998,** *471* (1–3), 283−288.

83. Lee, H.-J.; Choi, Y.-S.; Lee, K.-B.; Park, J.; Yoon, C.-J., Hydrogen bonding abilities of thioamide. *J. Phys. Chem. A* **2002,** *106* (30), 7010−7017.

84. Bordwell, F. G.; Algrim, D. J.; Harrelson, J. A., The relative ease of removing a proton, a hydrogen atom, or an electron from carboxamides versus thiocarboxamides. *J. Am. Chem. Soc.* **1988,** *110* (17), 5903–5904.

85. Wipf, P.; Hayes, G. B., Synthesis of oxazines and thiazines by cyclodehydration of hydroxy amides and thioamides. *Tetrahedron* **1998,** *54* (25), 6987−6998.

86. Misra, S.; Tewari, U., Complexing behaviour of aromatic thioamides (ArCSNHCOR). Transition metal complexes of N-carboethoxy-4-chlorobenzene- and N-carboethoxy-4-bromobenzene thioamide ligands. *Transit. Metal Chem.* **2002,** *27* (1), 120-125.

87. Judge, R. H.; Moule, D. C.; Goddard, J. D., Thioamide spectroscopy: long path length absorption and quantum chemical studies of thioformamide vapour, $CHSNH_2/CHSND_2$. *Can. J. Chemistry* **1987,** *65* (9), 2100−2105.

88. Reiner, A.; Wildemann, D.; Fischer, G.; Kiefhaber, T., Effect of thioxopeptide bonds on α-helix structure and stability. *J. Am. Chem. Soc.* **2008,** *130* (25), 8079−8084.

89. Miwa, J. H.; Pallivathucal, L.; Gowda, S.; Lee, K. E., Conformational stability of helical peptides containing a thioamide linkage. *Org. Lett.* **2002,** *4* (26), 4655−4657.

90. Pan, M.; Mabry, T. J.; Beale, J. M.; Mamiya, B. M., Nonprotein amino acids from *Cycas revoluta*. *Phytochemistry* **1997,** *45* (3), 517−519.

91. Abas, S. A.; Hossain, M. B.; van der Helm, D.; Schmitz, F. J.; Laney, M.; Cabuslay, R.; Schatzman, R. C., Alkaloids from the Tunicate *Polycarpa aurata* from Chuuk Atoll. *J. Org. Chem.* **1996,** *61* (8), 2709–2712.

92. Kim, H. J.; Graham, D. W.; DiSpirito, A. A.; Alterman, M. A.; Galeva, N.; Larive, C. K.; Asunskis, D.; Sherwood, P. M. A., Methanobactin, a copper-acquisition compound from methane-oxidizing bacteria. *Science* **2004,** *305* (5690), 1612−1615.

93. Hayakawa, Y.; Sasaki, K.; Adachi, H.; Furihata, K.; Nagai, K.; Shin-ya, K., Thioviridamide, a novel apoptosis inducer in transformed cells from *Streptomyces olivoviridis*. *J. Antibiot.* **2006,** *59* (1), 1−5.

94. Lincke, T.; Behnken, S.; Ishida, K.; Roth, M.; Hertweck, C., Closthioamide: An Unprecedented Polythioamide Antibiotic from the Strictly Anaerobic Bacterium Clostridium cellulolyticum. *Angew. Chem. Int. Ed.* **2010,** *49* (11), 2011-2013.

95. Behnken, S.; Hertweck, C., Anaerobic bacteria as producers of antibiotics. *Appl. Microbiol. Biotechnol.* **2012,** *96* (1), 61–67.

96. Kloss, F.; Pidot, S.; Goerls, H.; Friedrich, T.; Hertweck, C., Formation of a dinuclear copper(I) complex from the *Clostridium*-derived antibiotic closthioamide. *Angew. Chem. Int. Ed.* **2013,** *52* (41), 10745−10748.

97. Ermler, U.; Grabarse, W.; Shima, S.; Goubeaud, M.; Thauer, R. K., Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation. *Science* **1997,** *278* (5342), 1457−1462.

98. Acton, Q. A., *Amides — Advances in Research and Application*. Scholarly Editions: Atlanta, GA, 2013.

99. Jones, W. C.; Nestor, J. J.; Du Vigneaud, V., Synthesis and some pharmacological properties of [1-deamino-9-thioglycine]oxytocin. *J. Am. Chem. Soc.* **1973,** *95* (17), 5677−5679.

100. Clausen, K.; Spatola, A. F.; Lemieux, C.; Schiller, P. W.; Lawesson, S.-O., Evidence of a peptide backbone contribution toward selective receptor recognition for leucine enkephalin thioamide analogs. *Biochem. Bioph. Res. Co.* **1984,** *120* (1), 305–310.

101. Okano, A.; James, R. C.; Pierce, J. G.; Xie, J.; Boger, D. L., Silver(I)-promoted conversion of thioamides to amidines: divergent synthesis of a key series of vancomycin aglycon residue 4 amidines that clarify binding behavior to model ligands. *J. Am. Chem. Soc.* **2012,** *134* (21), 8790−8793.

102. Manna, D.; Roy, G.; Mugesh, G., Antithyroid drugs and their analogues: synthesis, structure, and mechanism of action. *Accounts Chem. Res.* **2013,** *46* (11), 2706−2715.

103. Bond, M. D.; Holmquist, B.; Vallee, B. L., Thioamide substrate probes of metal-substrate interactions in carboxypeptidase A catalysis. *J. Inorg. Biochem.* **1986,** *28* (2–3), 97–105.

104. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T., n→π* Interactions of Amides and Thioamides: Implications for Protein Stability. *J. Am. Chem. Soc.* **2013,** *135* (21), 7843-7846.

105. Bartlett, G. J.; Newberry, R. W.; VanVeller, B.; Raines, R. T.; Woolfson, D. N., Interplay of hydrogen bonds and n→π* interactions in proteins. *J. Am. Chem. Soc.* **2013,** *135* (49), 18682–18688.

106. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T., n→π* interactions of amides and thioamides: implications for protein stability. *J. Am. Chem. Soc.* **2013,** *135* (21), 7843−7846.

107. Miwa, J. H.; Patel, A. K.; Vivatrat, N.; Popek, S. M.; Meyer, A. M., Compatibility of the thioamide functional group with β-sheet secondary structure: incorporation of a thioamide linkage into a β-hairpin peptide. *Org. Lett.* **2001,** *3* (21), 3373−3375.

108. Culik, R. M.; Jo, H.; DeGrado, W. F.; Gai, F., Using thioamides to site-specifically interrogate the dynamics of hydrogen bond formation in β-sheet folding. *J. Am. Chem. Soc.* **2012,** *134* (19), 8026–8029.

109. Newberry, R. W.; VanVeller, B.; Raines, R. T., Thioamides in the collagen triple helix. *Chem. Comm.* **2015,** *51* (47), 9624−9627.

110. Wildemann, D.; Schiene-Fischer, C.; Aumüller, T.; Bachmann, A.; Kiefhaber, T.; Lücke, C.; Fischer, G., A nearly isosteric photosensitive amide-backbone substitution allows enzyme activity switching in ribonuclease S. *J. Am. Chem. Soc.* **2007,** *129* (16), 4910−4918.

111. Goldberg, J. M.; Batjargal, S.; Petersson, E. J., Thioamides as fluorescence quenching probes: minimalist chromophores to monitor orotein dynamics. *J. Am. Chem. Soc.* **2010,** *132* (42), 14718−14720.

112. Goldberg, J. M.; Wissner, R. F.; Klein, A. M.; Petersson, E. J., Thioamide quenching of intrinsic protein fluorescence. *Chemical Communications* **2012,** *48*, 1550-1552.

113. Rehm, D.; Weller, A., Kinetics of fluorescence quenching by electron and H-atom transfer. *Isr. J. Chem.* **1970,** *8* (2), 259−271.

114. Goldberg, J. M.; Batjargal, S.; Chen, B. S.; Petersson, E. J., Thioamide quenching of fluorescent probes through photoinduced electron transfer: mechanistic studies and applications. *J. Am. Chem. Soc.* **2013,** *135* (49), 18651−18658.

115. Goldberg, J. M.; Speight, L. C.; Fegley, M. W.; Petersson, E. J., Minimalist probes for studying protein dynamics: thioamide quenching of selectively excitable fluorescent amino acids. *J. Am. Chem. Soc.* **2012,** *134* (14), 6088−6091.

116. Goldberg, J. M.; Chen, X.; Meinhardt, N.; Greenbaum, D. C.; Petersson, E. J., Thioamide-based fluorescent protease sensors. *J. Am. Chem. Soc.* **2014,** *136* (5), 2086−2093.

117. von Meyer, E., *A History of Chemistry from Earliest Times to the Present Day*. Macmillan and Company: London, UK, 1891.

118. Metzner, P., Thiocarbonyl Compounds as Specific Tools for Organic Synthesis. In *Organosulfur Chemistry 1*, Springer: New York, NY, 1999.

119. Hurd, R. N.; DeLaMater, G., The preparation and chemical properties of thionamides. *Chem. Rev.* **1961,** *61* (1), 45−86.

120. Ozturk, T.; Ertas, E.; Mert, O., A Berzelius reagent, phosphorus decasulfide (P_4S_{10}), in organic syntheses. *Chem. Rev.* **2010,** *110* (6), 3419−3478.

121. Scheeren, J. W.; Ooms, P. H. J.; Nivard, R. J. F., A general procedure for the conversion of a carbonyl group into a thione group with tetraphosphorus decasulfide. *Synthesis* **1973,** *1973* (03), 149−151.

122. Cava, M. P.; Levinson, M. I., Thionation reactions of lawesson's reagents. *Tetrahedron* **1985,** *41* (22), 5061–5087.

123. Jesberger, M.; Davis, T. P.; Barner, L., Applications of Lawesson's reagent in organic and organometallic syntheses. *Synthesis* **2003,** *2003* (13), 1929−1958.

124. Shalaby, M. A.; Grote, C. W.; Rapoport, H., Thiopeptide synthesis. α-amino thionoacid derivatives of nitrobenzotriazole as thioacylating agents. *J. Org. Chem.* **1996,** *61* (25), 9045−9048.

125. Mukherjee, S.; Verma, H.; Chatterjee, J., Efficient site-specific incorporation of thioamides into peptides on a solid support. *Org. Lett.* **2015,** *17* (12), 3150−3153.

126. Durek, T.; Becker, C. F. W., Protein semi-synthesis: new proteins for functional and structural studies. *Biomol. Eng.* **2005,** *22* (5–6), 153–172.

127. Verzele, D.; Madder, A., Patchwork protein chemistry: a practitioner's treatise on the advances in synthetic peptide stitchery. *ChemBioChem* **2013,** *14* (9), 1032−1048.

128. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B., Synthesis of proteins by native chemical ligation. *Science* **1994,** *266* (5186), 776–779.

129. Okamoto, R.; Izumi, M.; Kajihara, Y., Expanding the scope of native chemical ligation in glycopeptide synthesis. *Int. J. Pept. Res. Ther.* **2010,** *16* (3), 191−198.

130. Hackenberger, C. P. R.; Schwarzer, D., Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed.* **2008,** *47* (52), 10030−10074.

131. Fauvet, B.; Butterfield, S. M.; Fuks, J.; Brik, A.; Lashuel, H. A., One-pot total chemical synthesis of human α-synuclein. *Chem. Comm.* **2013,** *49* (81), 9254−9256.

132. Hejjaoui, M.; Butterfield, S.; Fauvet, B.; Vercruysse, F.; Cui, J.; Dikiy, I.; Prudent, M.; Olschewski, D.; Zhang, Y.; Eliezer, D.; Lashuel, H. A., Elucidating the role of C-terminal post-translational modifications using protein semisynthesis strategies: α-synuclein phosphorylation at tyrosine 125. *J. Am. Chem. Soc.* **2012,** *134* (11), 5196−5210.

133. Haj-Yahya, M.; Fauvet, B.; Herman-Bachinsky, Y.; Hejjaoui, M.; Bavikar, S. N.; Karthikeyan, S. V.; Ciechanover, A.; Lashuel, H. A.; Brik, A., Synthetic polyubiquitinated α-synuclein reveals important insights into the roles of the ubiquitin chain in regulating its pathophysiology. *P. Natl. Acad. Sci. USA* **2013,** *110* (44), 17726−17731.

134. Burai, R.; Ait-Bouziad, N.; Chiki, A.; Lashuel, H. A., Elucidating the role of site-specific nitration of α-synuclein in the pathogenesis of Parkinson's disease *via* protein semisynthesis and mutagenesis. *J. Am. Chem. Society.* **2015,** *137* (15), 5041–5052.

135. Müller, M. M.; Muir, T. W., Histones: at the crossroads of peptide and protein chemistry. *Chem. Rev.* **2015,** *115* (6), 2296−2349.

136. Johnson, E. C. B.; Kent, S. B. H., Insights into the mechanism and catalysis of the native chemical ligation Reaction. *J. Am. Chem. Soc.* **2006,** *128* (20), 6640−6646.

137. Bang, D.; Pentelute, B. L.; Kent, S. B. H., Kinetically controlled ligation for the convergent chemical synthesis of proteins. *Angew. Chem. Int. Ed.* **2006,** *45* (24), 3985–3988.

138. Shah, N. H.; Muir, T. W., Inteins: nature's gift to protein chemists. *Chem. Sci.* **2014,** *5* (1), 446−461.

139. Klabunde, T.; Sharma, S.; Telenti, A.; Jacobs, W. R.; Sacchettini, J. C., Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat. Struct. Mol. Biol.* **1998,** *5* (1), 31−36.

140. Muir, T. W.; Sondhi, D.; Cole, P. A., Expressed protein ligation: a general method for protein engineering. *P. Natl. Acad. Sci. USA* **1998,** *95* (12), 6705−6710.

141. Zettler, J.; Schütz, V.; Mootz, H. D., The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* **2009,** *583* (5), 909−914.

142. Lockless, S. W.; Muir, T. W., Traceless protein splicing utilizing evolved split inteins. *P. Natl. Acad. Sci. USA* **2009,** *106* (27), 10999−11004.

143. Topilina, N. I.; Mills, K. V., Recent advances in *in vivo* applications of intein-mediated protein splicing. *Mob. DNA* **2014,** *5*, 5.

144. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T., Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.* **2000,** *2* (13), 1939−1941.

145. Tam, A.; Soellner, M. B.; Raines, R. T., Water-soluble phosphinothiols for traceless Staudinger ligation and integration with expressed protein ligation. *J. Am. Chem. Soc.* **2007,** *129* (37), 11421−11430.

146. Saxon, E.; Bertozzi, C. R., Cell surface engineering by a modified staudinger reaction. *Science* **2000,** *287* (5460), 2007−2010.

147. Kolakowski, R. V.; Shangguan, N.; Sauers, R. R.; Williams, L. J., Mechanism of thioacid/azide amidation. *J. Am. Chem. Soc.* **2006,** *128* (17), 5695−5702.

148. Shangguan, N.; Katukojvala, S.; Greenberg, R.; Williams, L. J., The reaction of thioacids with azides: a new mechanism and new synthetic applications. *J. Am. Chem. Soc.* **2003,** *125* (26), 7754−7755.

149. Rohmer, K.; Mannuthodikayil, J.; Wittmann, V., Application of the thioacid-azide ligation (TAL) for the preparation of glycosylated and fluorescently labeled amino acids. *Isr. J. Chem.* **2015,** *55* (3−4), 437−446.

150. Muhlberg, M.; Siebertz, K. D.; Schlegel, B.; Schmieder, P.; Hackenberger, C. P. R., Controlled thioamide *vs.* amide formation in the thioacid-azide reaction under acidic aqueous conditions. *Chem. Comm.* **2014,** *50* (35), 4603−4606.

151. Bode, J. W.; Fox, R. M.; Baucom, K. D., Chemoselective amide ligations by decarboxylative condensations of *N*-alkylhydroxylamines and α-ketoacids. *Angew. Chem. Int. Ed.* **2006,** *45* (8), 1248–1252.

152. Pusterla, I.; Bode, J. W., The mechanism of the α-ketoacid–hydroxylamine amide-forming ligation. *Angew. Chem. Int. Ed.* **2012,** *51* (2), 513−516.

153. Pattabiraman, V. R.; Ogunkoya, A. O.; Bode, J. W., Chemical protein pynthesis by chemoselective α-ketoacid–hydroxylamine (KAHA) ligations with 5-oxaproline. *Angew. Chem. Int. Ed.* **2012,** *51* (21), 5114−5118.

154. Pusterla, I.; Bode, J. W., An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations. *Nat. Chem.* **2015,** *7* (8), 668−672.

155. Carey, F. A.; Sundberg, R. J., *Advanced Organic Chemistry: Part B: Reaction and Synthesis*. Springer: New York, NY, 2010.

156. von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M., Peptide α-thioester formation using standard Fmoc-chemistry. *Tetrahedron Lett.* **2003,** *44* (17), 3551−3554.

157. Nagalingam, A. C.; Radford, S. E.; Warriner, S. L., Avoidance of epimerization in the synthesis of peptide thioesters using Fmoc protection. *Synlett* **2007,** *2007* (16), 2517−2520.

158. Bang, D.; Pentelute, B. L.; Gates, Z. P.; Kent, S. B., Direct on-resin synthesis of peptide α-thiophenylesters for use in native chemical ligation. *Org. Lett.* **2006,** *8* (6), 1049–1052.

159. Raz, R.; Rademann, J., Fmoc-based synthesis of peptide thioesters for native chemical ligation employing a *tert*-butyl thiol linker. *Org. Lett.* **2011,** *13* (7), 1606−1609.

160. Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A., Solid phase synthesis of peptide C-terminal thioesters by Fmoc/*t*-Bu chemistry. *J. Am. Chem. Soc.* **1999,** *121* (49), 11369−11374.

161. Tofteng, A. P.; Sørensen, K. K.; Conde-Frieboes, K. W.; Hoeg-Jensen, T.; Jensen, K. J., Fmoc solid-phase synthesis of C-terminal peptide thioesters by formation of a backbone pyroglutamyl imide moiety. *Angew. Chem. Int. Ed.* **2009,** *48* (40), 7411−7414.

162. Blanco-Canosa, J. B.; Dawson, P. E., An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angewandte Chemie* **2008,** *120* (36), 6957-6961.

163. Mahto, S. K.; Howard, C. J.; Shimko, J. C.; Ottesen, J. J., A reversible protection strategy to improve Fmoc-SPPS of peptide thioesters by the *N*-acylurea approach. *ChemBioChem* **2011,** *12* (16), 2488−2494.

164. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E., Chemical protein synthesis using a second-generation N-acylurea linker for the preparation of peptide-thioester precursors. *J. Am. Chem. Soc.* **2015,** *137* (22), 7197–7209.

165. Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C.-H., Solid-phase synthesis of peptide and glycopeptide thioesters through side-chain-anchoring strategies. *Chem. - Eur. J.* **2008,** *14* (12), 3620−3629.

166. Nakamura, K.; Hanai, N.; Kanno, M.; Kobayashi, A.; Ohnishi, Y.; Ito, Y.; Nakahara, Y., Design and synthesis of silyl ether-based linker for solid-phase synthesis of glycopeptides. *Tetrahedron Lett.* **1999,** *40* (3), 515−518.

167. Scott, P., *Linker Strategies in Solid-Phase Organic Synthesis*. John Wiley & Sons: New York, NY, 2009.

168. Arnesen, T., Towards a functional understanding of protein N-terminal acetylation. *PLoS Biol.* **2011,** *9* (5), e1001074.

169. Kawakami, T.; Sumida, M.; Nakamura, K. i.; Vorherr, T.; Aimoto, S., Peptide thioester preparation based on an N-S acyl shift reaction mediated by a thiol ligation auxiliary. *Tetrahedron Lett.* **2005,** *46* (50), 8805−8807.

170. Nakamura, K. i.; Kanao, T.; Uesugi, T.; Hara, T.; Sato, T.; Kawakami, T.; Aimoto, S., Synthesis of peptide thioesters via an N–S acyl shift reaction under mild acidic conditions on an N-4,5-dimethoxy-2-mercaptobenzyl auxiliary group. *J. Pept. Sci.* **2009,** *15* (11), 731−737.

171. Kawakami, T.; Aimoto, S., The use of a cysteinyl prolyl ester (CPE) autoactivating unit in peptide ligation reactions. *Tetrahedron* **2009,** *65* (19), 3871−3877.

172. Kawakami, T.; Aimoto, S., Peptide ligation using a building block having a cysteinyl prolyl ester (CPE) autoactivating unit at the carboxy terminus. *Chem. Lett.* **2007,** *36* (1), 76−77.

173. Kawakami, T.; Aimoto, S., Sequential peptide ligation by using a controlled cysteinyl prolyl ester (CPE) autoactivating unit. *Tetrahedron Lett.* **2007,** *48* (11), 1903−1905.

174. Nagaike, F.; Onuma, Y.; Kanazawa, C.; Hojo, H.; Ueki, A.; Nakahara, Y.; Nakahara, Y., Efficient microwave-assisted tandem N-to-S acyl transfer and thioester exchange for the preparation of a glycosylated peptide thioester. *Org. Lett.* **2006,** *8* (20), 4465−4468.

175. Hojo, H.; Murasawa, Y.; Katayama, H.; Ohira, T.; Nakahara, Y.; Nakahara, Y., Application of a novel thioesterification reaction to the synthesis of chemokine CCL27 by the modified thioester method. *Org. Biomol. Chem.* **2008,** *6* (10), 1808−1813.

176. Ohta, Y.; Itoh, S.; Shigenaga, A.; Shintaku, S.; Fujii, N.; Otaka, A., Cysteine-derived S-protected oxazolidinones: potential chemical devices for the preparation of peptide thioesters. *Org. Lett.* **2006,** *8* (3), 467−470.

177. Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A., N→S acyl-transfer-mediated synthesis of peptide thioesters using anilide derivatives. *Org. Lett.* **2009,** *11* (4), 823−826.

178. Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A., *N*-Sulfanylethylanilide peptide as a crypto-thioester peptide. *ChemBioChem* **2011,** *12* (12), 1840−1844.

179. Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O., A one-pot three-segment ligation strategy for protein chemical synthesis. *Angew. Chem. Int. Ed.* **2012,** *51* (1), 209−213.

180. Jung, M. E.; Piizzi, G., gem-Disubstituent effect: theoretical basis and synthetic applications. *Chem. Rev.* **2005,** *105* (5), 1735−1766.

181. Burlina, F.; Papageorgiou, G.; Morris, C.; White, P. D.; Offer, J., In situ thioester formation for protein ligation using α-methylcysteine. *Chem. Sci.* **2014,** *5* (2), 766–770.

182. Zheng, J.-S.; Chang, H.-N.; Wang, F.-L.; Liu, L., Fmoc synthesis of peptide thioesters without post-chain-assembly manipulation. *J. Am. Chem. Soc.* **2011,** *133* (29), 11080−11083.

183. Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas, A. F.; Aucagne, V., A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.* **2015**, 10.1039/C5SC02630J.

184. Adams, A. L.; Macmillan, D., Investigation of peptide thioester formation via N→Se acyl transfer. *J. Pept. Sci.* **2013,** *19* (2), 65–73.

185. Botti, P.; Villain, M.; Manganiello, S.; Gaertner, H., Native chemical ligation through in situ O to S acyl shift. *Org. Lett.* **2004,** *6* (26), 4861–4864.

186. George, E. A.; Novick, R. P.; Muir, T. W., Cyclic peptide inhibitors of *Staphylococcal* virulence prepared by Fmoc-based thiolactone peptide synthesis. *J. Am. Chem. Soc.* **2008,** *130* (14), 4914−4924.

187. Liu, F.; Mayer, J. P., An Fmoc compatible, O to S shift-mediated procedure for the preparation of C-terminal thioester peptides. *J. Org. Chem.* **2013,** *78* (19), 9848−9856.

188. Tofteng, A. P.; Jensen, K. J.; Hoeg-Jensen, T., Peptide dithiodiethanol esters for *in situ* generation of thioesters for use in native ligation. *Tetrahedron Lett.* **2007,** *48* (12), 2105−2107.

189. Zheng, J.-S.; Cui, H.-K.; Fang, G.-M.; Xi, W.-X.; Liu, L., Chemical protein dynthesis by kinetically controlled ligation of peptide O-esters. *ChemBioChem* **2010,** *11* (4), 511−515.

190. Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J., Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc.* **2004,** *126* (21), 6576−6578.

191. Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L., Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem. Int. Ed.* **2011,** *50* (33), 7645−7649.

192. Zheng, J.-S.; Tang, S.; Qi, Y.-K.; Wang, Z.-P.; Liu, L., Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nat. Protocols* **2013,** *8* (12), 2483−2495.

193. Smith, H. B.; Hartman, F. C., Restoration of activity to catalytically deficient mutants of ribulosebisphosphate carboxylase/oxygenase by aminoethylation. *J. Biol. Chem.* **1988,** *263* (10), 4921−4925.

194. Kochendoerfer, G. G.; Chen, S.-Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C. L.; Low, D. W.; Cagle, E. N.; Carnevali, M.; Gueriguian, V.; Keogh, P. J.; Porter, H.; Stratton, S. M.; Wiedeke, M. C.; Wilken, J.; Tang, J.; Levy, J. J.; Miranda, L. P.; Crnogorac, M. M.; Kalbag, S.; Botti, P.; Schindler-Horvat, J.; Savatski, L.; Adamson, J. W.; Kung, A.; Kent, S. B. H.; Bradburne, J. A., Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* **2003,** *299* (5608), 884−887.

195. Bochar, D. A.; Tabernero, L.; Stauffacher, C. V.; Rodwell, V. W., Aminoethylcysteine can replace the function of the essential active site lysine of *Pseudomonas mevalonii* 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Biochemistry* **1999,** *38* (28), 8879–8883.

196. Heinrikson, R. L., The selective S-methylation of sulfhydryl groups in proteins and peptides with methyl-*p*-nitrobenzenesulfonate. *J. Biol. Chem.* **1971,** *246*, 4090–4096.

197. Roelfes, G.; Hilvert, D., Incorporation of selenomethionine into proteins through selenohomocysteine-mediated ligation. *Angew. Chem. Int. Ed.* **2003,** *42* (20), 2275−2277.

198. Wong, C. T. T.; Tung, C. L.; Li, X., Synthetic cysteine surrogates used in native chemical ligation. *Mol. BioSyst.* **2013,** *9* (5), 826−833.

199. Yan, L. Z.; Dawson, P. E., Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001,** *123* (4), 526−533.

200. Wan, Q.; Danishefsky, S. J., Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007,** *46* (48), 9248−9252.

201. Crich, D.; Banerjee, A., Native chemical ligation at phenylalanine. *J. Am. Chem. Soc.* **2007,** *129* (33), 10064–10065.

202. Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J., Native chemical ligation at valine: a contribution to peptide and glycopeptide synthesis. *Angew. Chem. Int. Ed.* **2008,** *47* (44), 8521–8524.

203. Haase, C.; Rohde, H.; Seitz, O., Native chemical ligation at valine. *Angew. Chem. Int. Ed.* **2008,** *47* (36), 6807−6810.

204. Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X.-W.; Liu, C.-F., Dual native chemical ligation at lysine. *J. Am. Chem. Soc.* **2009,** *131* (38), 13592−13593.

205. Ajish Kumar, K. S.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A., Highly efficient and chemoselective peptide ubiquitylation. *Angew. Chem. Int. Ed.* **2009,** *48* (43), 8090–8094.

206. Tan, Z.; Shang, S.; Danishefsky, S. J., Insights into the finer issues of native chemical ligation: an approach to cascade ligations. *Angew. Chem. Int. Ed.* **2010,** *49* (49), 9500−503.

207. Chen, J.; Wang, P.; Zhu, J.; Wan, Q.; Danishefsky, S. J., A program for ligation at threonine sites: application to the controlled total synthesis of glycopeptides. *Tetrahedron* **2010,** *66* (13), 2277–2283.

208. Shang, S.; Tan, Z.; Dong, S.; Danishefsky, S. J., An advance in proline ligation. *J. Am. Chem. Soc.* **2011,** *133* (28), 10784−10786.

209. Siman, P.; Karthikeyan, S. V.; Brik, A., Native chemical ligation at glutamine. *Org. Lett.* **2012,** *14* (6), 1520−1523.

210. Malins, L. R.; Cergol, K. M.; Payne, R. J., Peptide ligation–desulfurization chemistry at arginine. *ChemBioChem* **2013,** *14* (5), 559−563.

211. Cergol, K. M.; Thompson, R. E.; Malins, L. R.; Turner, P.; Payne, R. J., One-pot peptide ligation–desulfurization at glutamate. *Org. Lett.* **2014,** *16* (1), 290–293.

212. Thompson, R. E.; Chan, B.; Radom, L.; Jolliffe, K. A.; Payne, R. J., Chemoselective peptide ligation–desulfurization at aspartate. *Angew. Chem. Int. Ed.* **2013,** *52* (37), 9723−9727.

213. Sayers, J.; Thompson, R. E.; Perry, K. J.; Malins, L. R.; Payne, R. J., Thiazolidine-protected β-thiol asparagine: applications in one-pot ligation–desulfurization chemistry. *Org. Lett.* **2015,** *17* (19), 4902−4905.

214. Malins, L. R.; Cergol, K. M.; Payne, R. J., Chemoselective sulfenylation and peptide ligation at tryptophan. *Chem. Sci.* **2014,** *5* (1), 260−266.

215. Sun, X.-H.; Yu, H.-Z.; Yang, M.-M.; Yang, Y.-M.; Dang, Z.-M., Relative facility of the desulfurization of amino acids and their carboxylic derivatives. *J. Phys. Org. Chem* **2015,** *28* (9), 586−590.

216. Metanis, N.; Keinan, E.; Dawson, P. E., Traceless ligation of cysteine peptides using selective deselenization. *Angew. Chem. Int. Ed.* **2010,** *49* (39), 7049−7053.

217. Townsend, S. D.; Tan, Z.; Dong, S.; Shang, S.; Brailsford, J. A.; Danishefsky, S. J., Advances in proline ligation. *J. Am. Chem. Soc.* **2012,** *134* (8), 3912−3916.

218. Malins, L. R.; Payne, R. J., Synthesis and utility of β-selenol-phenylalanine for native chemical ligation–deselenization chemistry. *Org. Lett.* **2012,** *14* (12), 3142−3145.

219. Malins, L. R.; Mitchell, N. J.; Payne, R. J., Peptide ligation chemistry at selenol amino acids. *J. Pept. Sci.* **2014,** *20* (2), 64−77.

220. Woollins, J. D.; Laitinen, R., *Selenium and Tellurium Chemistry: From Small Molecules to Biomolecules and Materials*. Springer: New York, NY, 2011.

221. Malins, L. R.; Mitchell, N. J.; McGowan, S.; Payne, R. J., Oxidative deselenization of selenocysteine: applications for programmed ligation at serine. *Angew. Chem. Int. Ed.* **2015,** *54* (43), 12716−12721.

222. Dery, S.; Reddy, P. S.; Dery, L.; Mousa, R.; Dardashti, R. N.; Metanis, N., Insights into the deselenization of selenocysteine into alanine and serine. *Chem. Sci.* **2015,** *6* (11), 6207–6212.

223. Wessjohann Ludger, A.; Schneider, A.; Abbas, M.; Brandt, W., Selenium in chemistry and biochemistry in comparison to sulfur. *Biol. Chem.* **2007,** *388* (10), 997−1006.

224. Allen, F., The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystall. B - Stru.* **2002,** *58*, 380–388.

225. Huber, R. E.; Criddle, R. S., Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. *Arch. Biochem. and Biophys.* **1967,** *122* (1), 164−173.

226. Nauser, T.; Steinmann, D.; Koppenol, W., Why do proteins use selenocysteine instead of cysteine? *Amino Acids* **2012,** *42* (1), 39−44.

227. Jacob, C.; Giles, G. I.; Giles, N. M.; Sies, H., Sulfur and selenium: the role of oxidation state in protein structure and function. *Angew. Chem. Int. Ed.* **2003,** *42* (39), 4742−4758.

228. Zhong, L.; Holmgren, A., Essential role of selenium in the catalytic activities of mammalian thioredoxin reductase revealed by characterization of recombinant enzymes with selenocysteine mutations. *J. Biol. Chem.* **2000,** *275* (24), 18121−18128.

229. Chaudiere, J.; Wilhelmsen, E. C.; Tappel, A. L., Mechanism of selenium-glutathione peroxidase and its inhibition by mercaptocarboxylic acids and other mercaptans. *J. Biol. Chem.* **1984,** *259* (2), 1043–1050.

230. Johansson, L.; Gafvelin, G.; Arnér, E. S. J., Selenocysteine in proteins − properties and biotechnological use. *BBA - Gen. Subjects* **2005,** *1726* (1), 1−13.

231. Hondal, R. J.; Nilsson, B. L.; Raines, R. T., Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **2001,** *123* (21), 5140−5141.

232. McGrath, N. A.; Raines, R. T., Chemoselectivity in chemical biology: acyl transfer reactions with sulfur and selenium. *Accounts Chem. Res.* **2011,** *44* (9), 752−761.

233. Berry, S. M.; Gieselman, M. D.; Nilges, M. J.; van der Donk, W. A.; Lu, Y., An engineered azurin variant containing a selenocysteine copper ligand. *J. Am. Chem. Soc.* **2002,** *124* (10), 2084–2085.

234. Okumura, M.; Shimamoto, S.; Hidaka, Y., A chemical method for investigating disulfide-coupled peptide and protein folding. *FEBS J.* **2012,** *279* (13), 2283−2295.

235. Pegoraro, S.; Fiori, S.; Cramer, J.; Rudolph-Böhner, S.; Moroder, L., The disulfide-coupled folding pathway of apamin as derived from diselenide-quenched analogs and intermediates. *Protein Sci.* **1999,** *8* (8), 1605-1613.

236. Walewska, A.; Zhang, M.-M.; Skalicky, J. J.; Yoshikami, D.; Olivera, B. M.; Bulaj, G., Integrated oxidative folding of cysteine/selenocysteine containing peptides: improving chemical synthesis of conotoxins. *Angew. Chem. Int. Ed.* **2009,** *48* (12), 2221−2224.

237. Luthra, N. P.; Costello, R. C.; Odom, J. D.; Dunlap, R. B., Demonstration of the feasibility of observing nuclear magnetic resonance signals of ^{77}Se covalently attached to proteins. *J. Biol. Chem.* **1982,** *257* (3), 1142−1144.

238. Gettins, P.; Wardlaw, S. A., NMR relaxation properties of ^{77}Se-labeled proteins. *J. Biol. Chem.* **1991,** *266* (6), 3422−3426.

239. House, K. L.; Dunlap, R. B.; Odom, J. D.; Wu, Z. P.; Hilvert, D., Structural characterization of selenosubtilisin by selenium-77 NMR spectroscopy. *J. Am. Chem. Soc.* **1992,** *114* (22), 8573−8579.

240. Schaefer, S. A.; Dong, M.; Rubenstein, R. P.; Wilkie, W. A.; Bahnson, B. J.; Thorpe, C.; Rozovsky, S., ^{77}Se Enrichment of proteins expands the biological NMR toolbox. *J. Mol. Biol.* **2013,** *425* (2), 222−231.

241. Gieselman, M. D.; Xie, L.; van der Donk, W. A., Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.* **2001,** *3* (9), 1331−1334.

242. Flemer, S., Fmoc-Sec(Xan)-OH: synthesis and utility of Fmoc selenocysteine SPPS derivatives with acid-labile sidechain protection. *J. Pept. Sci.* **2015,** *21* (1), 53−59.

243. Schroll, A. L.; Hondal, R. J.; Flemer, S., The use of 2,2′-dithiobis(5-nitropyridine) (DTNP) for deprotection and diselenide formation in protected selenocysteine-containing peptides. *J. Pept. Sci.* **2012,** *18* (3), 155−162.

244. Böck, A.; Forchhammer, K.; Heider, J.; Leinfelder, W.; Sawers, G.; Veprek, B.; Zinoni, F., Selenocysteine: the 21st amino acid. *Mol. Microbiol.* **1991,** *5* (3), 515–520.

245. Turanov, A. A.; Xu, X.-M.; Carlson, B. A.; Yoo, M.-H.; Gladyshev, V. N.; Hatfield, D. L., Biosynthesis of selenocysteine, the 21st amino acid in the genetic code, and a novel pathway for cysteine biosynthesis. *Adv. Nutr.* **2011,** *2* (2), 122−128.

246. Ganichkin, O. M.; Xu, X.-M.; Carlson, B. A.; Mix, H.; Hatfield, D. L.; Gladyshev, V. N.; Wahl, M. C., Structure and catalytic mechanism of eukaryotic selenocysteine synthase. *J. Biol. Chem.* **2008,** *283* (9), 5849−5865.

247. Gonzalez-Flores, J. N.; Gupta, N.; DeMong, L. W.; Copeland, P. R., The selenocysteine-specific elongation factor contains a novel and multi-functional domain. *Journal of Biological Chemistry* **2012,** *287* (46), 38936−38945.

248. Fletcher, J. E.; Copeland, P. R.; Driscoll, D. M.; Krol, A., The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2. *RNA* **2001,** *7* (10), 1442−1453.

249. Stoytcheva, Z.; Tujebajeva, R. M.; Harney, J. W.; Berry, M. J., Efficient incorporation of multiple selenocysteines involves an inefficient decoding step serving as a potential translational checkpoint and ribosome bottleneck. *Mol. Cell. Biol.* **2006,** *26* (24), 9177−9184.

250. Kryukov, G. V.; Castellano, S.; Novoselov, S. V.; Lobanov, A. V.; Zehtab, O.; Guigó, R.; Gladyshev, V. N., Characterization of mammalian selenoproteomes. *Science* **2003,** *300* (5624), 1439−1443.

251. Reeves, M. A.; Hoffmann, P. R., The human selenoproteome: recent insights into functions and regulation. *Cell. Mol. Life Sci.* **2009,** *66* (15), 2457−2478.

252. Su, D.; Li, Y.; Gladyshev, V. N., Selenocysteine insertion directed by the 3′-UTR SECIS element in *Escherichia coli*. *Nucleic Acids Res.* **2005,** *33* (8), 2486−2492.

253. Novoselov, S. V.; Lobanov, A. V.; Hua, D.; Kasaikina, M. V.; Hatfield, D. L.; Gladyshev, V. N., A highly efficient form of the selenocysteine insertion sequence element in protozoan parasites and its use in mammalian cells. *P. Natl. Acad. Sci. USA* **2007,** *104* (19), 7857β7862.

254. Schaefer-Ramadan, S.; Thorpe, C.; Rozovsky, S., Site-specific insertion of selenium into the redox-active disulfide of the flavoprotein augmenter of liver regeneration. *Arch. Biochem. and Biophys.* **2014,** *548*, 60−65.

255. Jiang, Z.; Arnér, E. S. J.; Mu, Y.; Johansson, L.; Shi, J.; Zhao, S.; Liu, S.; Wang, R.; Zhang, T.; Yan, G.; Liu, J.; Shen, J.; Luo, G., Expression of selenocysteine-containing glutathione S-transferase in *Escherichia coli*. *Biochem. Bioph. Res. Co.* **2004,** *321* (1), 94−101.

256. Mogk, A.; Schmidt, R.; Bukau, B., The *N*-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol.* **2007,** *17* (4), 165−172.

257. Kaji, A.; Kaji, H.; Novelli, G. D., Soluble amino acid-incorporating system I. Preparation of system and nature of reaction. *J. Biol. Chem.* **1965,** *240*, 1185−1191.

258. Leibowitz, M. J.; Soffer, R. L., Enzymatic modification of proteins: III. purification and properties of a leucyl, phenylalanyl transfer ribonucleic acid-protein transferase from *Escherichia coli*. *J. Biol. Chem.* **1970,** *245* (8), 2066−2073.

259. Soffer, R. L., Peptide acceptors in the leucine, phenylalanine transfer reaction. *J. Biol. Chem.* **1973,** *248* (24), 8424−8428.

260. Tobias, J. W.; Shrader, T. E.; Rocap, G.; Varshavsky, A., The N-end rule in bacteria. *Science* **1991,** *254* (5036), 1374−1377.

261. Watanabe, K.; Toh, Y.; Suto, K.; Shimizu, Y.; Oka, N.; Wada, T.; Tomita, K., Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature* **2007,** *449* (7164), 867−871.

262. Fung, A. W.; Ebhardt, H. A.; Abeysundara, H.; Moore, J.; Xu, Z.; Fahlman, R. P., An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation. *J. Mol. Biol.* **2011,** *409* (4), 617−629.

263. Abramochkin, G.; Shrader, T. E., Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J. Biol. Chem.* **1996,** *271* (37), 22901–22907.

264. Connor, R. E.; Piatkov, K.; Varshavsky, A.; Tirrell, D. A., Enzymatic N-terminal addition of noncanonical amino acids to peptides and proteins. *ChemBioChem* **2008,** *9* (3), 366–369.

265. Taki, M.; Kuno, A.; Matoba, S.; Kobayashi, Y.; Futami, J.; Murakami, H.; Suga, H.; Taira, K.; Hasegawa, T.; Sisido, M., Leucyl/phenylalanyl-tRNA-protein transferase-mediated chemoenzymatic coupling of N-terminal Arg/Lys units in post-translationally processed proteins with non-natural amino acids. *ChemBioChem* **2006,** *7* (11), 1676−1679.

266. Taki, M.; Sisido, M., Leucyl/phenylalanyl(L/F)-tRNA-protein transferase-mediated aminoacyl transfer of a nonnatural amino acid to the N-terminus of peptides and proteins and subsequent functionalization by bioorthogonal reactions. *Peptide Sci.* **2007,** *88* (2), 263−271.

267. Taki, M.; Kuroiwa, H.; Sisido, M., Chemoenzymatic transfer of fluorescent non-natural amino acids to the N-terminus of a protein/peptide. *ChemBioChem* **2008,** *9* (5), 719−722.

268. Wagner, A. M.; Fegley, M. W.; Warner, J. B.; Grindley, C. L. J.; Marotta, N. P.; Petersson, E. J., N-terminal protein modification using simple aminoacyl transferase substrates. *J. Am. Chem. Soc.* **2011,** *133* (38), 15139−15147.

269. Tanaka, T.; Wagner, A. M.; Warner, J. B.; Wang, Y. J.; Petersson, E. J., Expressed protein ligation at methionine: N-terminal attachment of homocysteine, ligation, and masking. *Angew. Chem. Int. Ed.* **2013,** *52* (24), 6210−6213.

270. Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G., Methods and protocols of modern solid phase peptide synthesis. *Mol. Biotechnol.* **2006,** *33* (3), 239–254.

271. Brinkley, M., A brief survey of methods for preparing protein conjugates with dyes, haptens and crosslinking reagents. *Bioconjugate Chem.* **1992,** *3* (1), 2–13.

272. England, P. M., Unnatural amino acid mutagenesis: a precise tool for probing protein structure and function. *Biochemistry* **2004,** *43* (37), 11623–11629.

273. Tzvetkova, S.; Kluger, R., Biomimetic aminoacylation of ribonucleotides and RNA with aminoacyl phosphate esters and lanthanum salts. *J. Am. Chem. Soc.* **2007,** *129* (51), 15848−15854.

274. Maini, R.; Dedkova, L. M.; Paul, R.; Madathil, M. M.; Chowdhury, S. R.; Chen, S.; Hecht, S. M., Ribosome-mediated incorporation of dipeptides and dipeptide analogues into proteins *in vitro*. *J. Am. Chem. Soc.* **2015,** *137* (35), 11206−11209.

275. Chen, J.; Warren, J. D.; Wu, B.; Chen, G.; Wan, Q.; Danishefsky, S. J., A route to cyclic peptides and glycopeptides by native chemical ligation using in situ derived thioesters. *Tetrahedron Lett.* **2006,** *47* (12), 1969–1972.

276. Blanco-Canosa, J. B.; Dawson, P. E., An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem. Int. Ed.* **2008,** *47* (36), 6851–6855.

277. Pátek, M.; Lebl, M., Safety-catch and multiply cleavable linkers in solid-phase synthesis. *Peptide Sci.* **1998,** *47* (5), 353−363.

278. Botti, P.; Villain, M.; Manganiello, S.; Gaertner, H., Native Chemical Ligation through in Situ O to S Acyl Shift. *Organic Letters* **2004,** *6* (26), 4861-4864.

279. Breydo, L.; Wu, J. W.; Uversky, V. N., α-Synuclein misfolding and Parkinson's disease. *BBA - Mol. Basis Dis.* **2012,** *1822* (2), 261–285.

280. Goldberg, J. M.; Wissner, R. F.; Klein, A. M.; Petersson, E. J., Thioamide quenching of intrinsic protein fluorescence. *Chem. Comm.* **2012,** *48* (10), 1550−1552.

281. Lorenzen, N.; Nielsen, S. B.; Buell, A. K.; Kaspersen, J. D.; Arosio, P.; Vad, B. S.; Paslawski, W.; Christiansen, G.; Valnickova-Hansen, Z.; Andreasen, M.; Enghild, J. J.; Pedersen, J. S.; Dobson, C. M.; Knowles, T. P. J.; Otzen, D. E., The role of stable α-synuclein oligomers in the molecular events underlying amyloid formation. *J. Am. Chem. Soc.* **2014,** *136* (10), 3859−3868.

282. Ireland, R. E.; Liu, L., An improved procedure for the preparation of the Dess-Martin periodinane. *J. Org. Chem.* **1993,** *58* (10), 2899−2899.

283. Batjargal, S.; Wang, Y. J.; Goldberg, J. M.; Wissner, R. F.; Petersson, E. J., Native chemical ligation of thioamide-containing peptides: development and application

to the synthesis of labeled α-Synuclein for misfolding studies. *J. Am. Chem. Soc.* **2012,** *134* (22), 9172–9182.

284. Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M., Multiparametric fluorescence detection of early stages in the amyloid protein aggregation of pyrene-labeled α-synuclein. *J. Mol. Biol.* **2008,** *378* (5), 1064−1073.

285. Yushchenko, D. A.; Fauerbach, J. A.; Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M., Fluorescent ratiometric MFC probe sensitive to early stages of α-synuclein aggregation. *J. Am. Chem. Soc.* **2010,** *132* (23), 7860−7861.

286. Echols, N.; Harrison, P.; Balasubramanian, S.; Luscombe, N. M.; Bertone, P.; Zhang, Z.; Gerstein, M., Comprehensive analysis of amino acid and nucleotide composition in eukaryotic genomes, comparing genes and pseudogenes. *Nucleic Acids Res.* **2002,** *30* (11), 2515–2523.

287. Zhou, W.; Freed, C. R., Tyrosine-to-cysteine modification of human α-synuclein enhances protein aggregation and cellular toxicity. *J. Biol. Chem.* **2004,** *279* (11), 10128−10135.

288. Wissner, R. F.; Wagner, A. M.; Warner, J. B.; Petersson, E. J., Efficient, traceless semi-synthesis of α-synuclein labeled with a fluorophore/thioamide FRET Pair. *Synlett* **2013,** *24* (18), 2454−2458.

289. Trost, B. M.; Fleming, I., *Comprehensive Organic Synthesis: Selectivity, Strategy, and Efficiency in Modern Organic Chemistry*. Elsevier: Oxford, UK, 1991.

290. Kornfeld, E. C., Raney nickel hydrogenolysis of thioamides: a new amine synthesis. *J. Org. Chem.* **1951,** *16* (1), 131−138.

291. Hendrickson, T. L.; Imperiali, B., Metal ion dependence of oligosaccharyl transferase: implications for catalysis. *Biochemistry* **1995,** *34* (29), 9444−9450.

292. Engman, L.; Stern, D., Thiol/diselenide exchange for the generation of benzeneselenolate ion. Catalytic reductive ring-opening of α,β-epoxy ketones. *J. Org. Chem.* **1994,** *59* (18), 5179–5183.

293. Ji, S.; Cao, W.; Yu, Y.; Xu, H., Dynamic diselenide bonds: exchange reaction induced by visible light without catalysis. *Angew. Chem. Int. Ed.* **2014,** *53* (26), 6781−6785.

294. Thompson, R. E.; Liu, X.; Alonso-García, N.; Pereira, P. J. B.; Jolliffe, K. A.; Payne, R. J., Trifluoroethanethiol: an additive for efficient one-pot peptide ligation-desulfurization chemistry. *J. Am. Chem. Soc.* **2014,** *136* (23), 8161−8164.

295. Cox, B. G., *Acids and Bases: Solvent Effects on Acid-Base Strength*. OUP Oxford: Oxford, UK, 2013.

296. Fu, Y.; Lin, B.-L.; Song, K.-S.; Liu, L.; Guo, Q.-X., Substituent effects on the S−H bond dissociation energies of thiophenols. *J. Chem. Soc. Perk. T. 2* **2002,** (7), 1223−1230.

297. Luo, Y.-R., BDEs of S−, Se−, Te−, Po−X bonds. In *Comprehensive Handbook of Chemical Bond Energies*, CRC Press: 2007; pp 424-454.

298. Caban, K.; Copeland, P. R., Size matters: a view of selenocysteine incorporation from the ribosome. *Cell. Mol. Life Sci.* **2006,** *63* (1), 73–81.

299. Balzi, E.; Choder, M.; Chen, W. N.; Varshavsky, A.; Goffeau, A., Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **1990,** *265* (13), 7464–7471.

300. Varshavsky, A., The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **2011,** *20* (8), 1298−1345.

301. Taiji, M.; Yokoyama, S.; Miyazawa, T., Transacylation rates of aminoacyladenosine moiety at the 3'-terminus of aminoacyl-transfer ribonucleic acid. *Biochemistry* **1983,** *22* (13), 3220−3225.

302. Graciet, E.; Hu, R.-G.; Piatkov, K.; Rhee, J. H.; Schwarz, E. M.; Varshavsky, A., Aminoacyl-transferases and the *N*-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *P. Natl. Acad. Sci. USA* **2006,** *103* (9), 3078−3083.

303. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976,** *72* (1), 248–254.

304. Tang, Y.; Tirrell, D. A., Attenuation of the editing activity of the *Escherichia coli* leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins *in vivo*. *Biochemistry* **2002,** *41* (34), 10635−10645.

305. Palencia, A.; Crépin, T.; Vu, M. T.; Lincecum, T. L.; Martinis, S. A.; Cusack, S., Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. *Nat. Struct. Mol. Biol.* **2012,** *19* (7), 677−684.

306. Brustad, E.; Bushey, M. L.; Brock, A.; Chittuluru, J.; Schultz, P. G., A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. *Bioorg. Med. Chem. Lett.* **2008,** *18* (22), 6004–6006.