Semi-Synthesis of Thioamide Containing Proteins

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID:	OB-PER-02-2015-000224.R1
Article Type:	Perspective
Date Submitted by the Author:	15-Mar-2015

Complete List of Authors:	Wang, Yanxin; University of Pennsylvania, Chemistry Szantai-Kis, Daniel; University of Pennsylvania School of Medicine, Biochemistry and Molecular Biophysics Petersson, E. James; University of Pennsylvania, Chemistry

Organic & Biomolecular Chemistry

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/obc

PERSPECTIVE

Semi-Synthesis of Thioamide Containing Proteins

Yanxin J. Wang, D. Miklos Szantai-Kis, and E. James Petersson

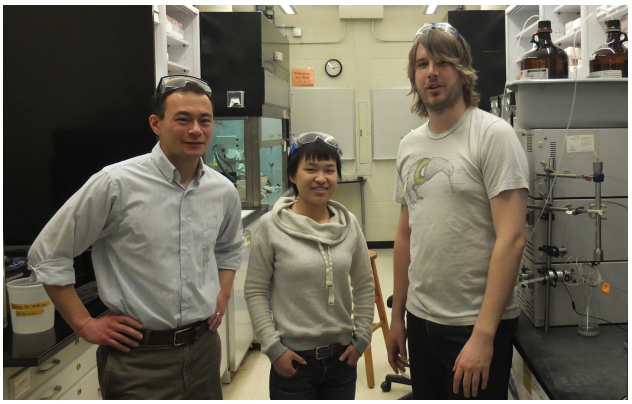
Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

Our laboratory has shown that the thioamide, a single atom O-to-S substitution, can be a versatile fluorescence quenching probe that is minimally-perturbing when placed at many locations in a protein sequence. In order to make these and other thioamide experiments applicable to full-sized proteins, we have developed methods for incorporating thioamides by generating thiopeptide fragments through solid phase synthesis and ligating them to protein fragments expressed in *E. coli*. To install donor fluorophores, we have adapted unnatural amino acid mutagenesis methods, including the generation of new tRNA synthetases for the incorporation of small, intrinsically fluorescent amino acids. We have used a combination of these two methods, as well as chemoenzymatic protein modification, to efficiently install sidechain and backbone modifications to generate proteins labeled with fluorophore/thioamide pairs.

Introduction

A thioamide is a single-atom deviation from the canonical amide functional group (herein referred to as "oxoamide"), where the carbonyl oxygen is replaced with sulfur.[1] While sterically similar,[2-4] these two groups confer very different chemical and physical properties (Table 1) - thioamides exhibit higher reactivity as nucleophiles,[5] greater affinity for soft metals,[6] altered hydrogen bonding abilities,[7] as well as unique spectroscopic features.[8, 9] Although thioamides are known to undergo desulfurization under harsh conditions,[10, 11] the thioamide group has been repeatedly shown to be stable in aqueous buffer.[9, 12] Therefore, thioamides are excellent isosteric replacements for oxoamides found in proteins,[13] where their unique properties can be leveraged without major steric perturbations or stability concerns.

In nature, thioamides are found in small molecules and at least one folded protein, and are typically key contributors to their biological activity. Closthioamide, a polyamide isolated from *Clostridium cellulolyticum*, derives its antibiotic activity solely from the presence of six thioamide bonds.[14] Methyl-coenzyme M reductase, an enzyme involved in biological methane formation, was characterized by X-ray crystallography to possess a thioamide bond near its active site, which presumably facilitates the oxidation and reduction of cofactor and substrate.[15] In synthetic organic chemistry, thioamides have also been explored for similar applications, especially in the derivatization of natural products for structure-activity relationship studies, where an oxoamide was converted to either a thioamide or a subsequent derivative – thiazole, thiazolene, thiazine, or amidine – and assessed for biological activity.[16-18]



Yanxin J. Wang (Center) graduated from the Hong Kong University of Science and Technology in 2010 with a B.S. degree in Chemistry. She then joined Prof. Petersson's lab at the University of Pennsylvania to pursue her Ph.D. degree. She has worked on several projects to enable the incorporation of thioamide into peptides and proteins.

D. Miklos Szantai-Kis (Right) earned his Diplom in Chemistry in 2012 from the Karlsruhe Institute of Technology, Karlsruhe, Germany, where he worked for Prof. Katja Schmitz. After working with Prof. Ralph Mazitschek at Harvard University for a year, he started pursuing a Ph.D. in Biochemistry and Molecular Biophysics at the University of Pennsylvania in 2013. He currently is a graduate student in Prof. Petersson's group, working on biological incorporation of sidechain thioamides.

E. James Petersson (Left) was educated at Dartmouth College, where he worked in the laboratory of David Lemal. He then studied under Dennis Dougherty at the California Institute of

Technology as an NIH Predoctoral Fellow. After obtaining his Ph.D. in 2005, he worked as an NIH Postdoctoral Fellow at Yale University with Alanna Schepartz. He was appointed as Assistant Professor in the Department of Chemistry at the University of Pennsylvania in 2008 and in the Biochemistry and Molecular Biophysics group in the Perelman School of Medicine in 2013. He has been the recipient of several awards, including a Sloan Fellowship, an NSF CAREER award, the JPOC Early Excellence in Physical Organic Chemistry award, and recognition as a Searle Scholar.

In peptide and protein chemistry, thioamides are particularly appealing modifications, given the abundance of peptide bonds and side chain oxoamides in these biomolecules. One could envision placing thioamides at any residue along the protein sequence as probes for mechanistic or structural studies. In one early example, Du Vigneaud introduced a thioglycine residue (Gly^S, in this perspective thioamide residues will be denoted by the correspoding amino acid code with a superscript S) into the peptide hormone oxytocin.[19] Another example showed the use of thioamide-modified short peptide substrates to evaluate the implications of metal-thioamide binding in carboxypeptidase A catalysis.[20] More recently, Raines and co-workers have utilized the different electronic properties between C=S and C=O bonds to experimentally measure the strength of the n→π* interactions between two adjacent peptide bonds.[21] Other groups have incorporated thioamides into short α-helices and β-strands, and demonstrated that thioamides can exert subtle changes in folding pathways while being generally well-tolerated in these secondary structures.[9, 12, 22, 23]

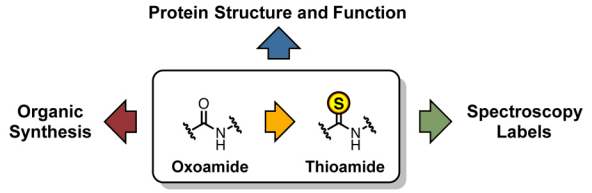


Table 1 Selected Properties and Applications of Thioamides.[24]

Property	Oxo	Thio	Example Application
vdW radius (Å)[2]	1.40	1.85	---
C=X bond length (Å)[25]	1.23	1.71	---
Electronegativity[26]	3.44	2.58	Metal-substrate, n→π* interactions[20, 21]
C=X···H–N BDE (kcal mol^{-1})[27]	6.1	4.8	Peptide folding[9, 22, 23]
π–π* absorption (nm)[28]	200	270	Photoswitch, FRET-based quencher[8, 29]
E_{Ox} (V *vs*. S.H.E.)[30]	3.29	1.21	PET-based quencher[31, 32]

vdW = van der Waals, BDE = bond dissociation energy, FRET = Förster resonance energy transfer; PET = photoinduced electron transfer, S.H.E. = standard hydrogen electrode

With the work of our group and others, thioamides are now emerging as valuable, minimally-perturbing spectroscopic labels for the study of protein folding and stability. Earlier experiments showed that thioamides exhibit unique circular dichroism signature[9] and photo-isomerization properties[29] in a background of oxoamides, allowing the specific examination and manipulation of the thioamide label. We have extensively explored thioamides as fluorescence quenching probes, where two types of fluorophore partners were identified based on quenching mechanism. UV wavelength dyes, such as *p*-cyanophenylalanine (Cnf) and the natural amino acid Tyr, have spectral overlap of their emission with thioamide absorption and are quenched through space *via* Förster Resonance Energy Transfer (FRET).[8, 33] Near UV and visible wavelength fluorophores, including acridone and other commonly used dyes such as Alexa Fluor 488 and BODIPY FL, are quenched upon near-van der Waals contact with thioamides *via* photo-induced electron transfer (PET).[31, 32] Both types of quenching interactions are distance-dependent, allowing us to track the relative movements of the thioamide and dye label by tracking the fluorescence change in a doubly labeled protein. This approach has been validated in multiple peptide and protein systems for various applications, including protein misfolding of relevance to Parkinson's disease[34] and calpain protease activity in cell lysates.[35] The photophysics and biological applications of thioamides have recently been reviewed elsewhere.[24] In this perspective, we will focus on our methodologocal studies to enable the efficient synthesis of fluorophore/thioamide dually labeled proteins. It is also worth noting that the methods discussed below can conceivably be applied to selenopeptides, where a selenocarbonyl is inserted in place of the carbonyl. Such peptides have proven useful in biophysical assays,[36] but can be tricky to synthesize and handle.[37] In the following discussion, we will only focus on thioamides.

Incorporation of thioamides into peptides and proteins

A crucial step towards exploiting the properties of thioamides was establishing a reliable method for their incorporation at specific sites in proteins of interest. Prior to our work, the applications of thioamides were largely restricted to small molecules and short peptides (≤ 35 aa).[9, 21] Using a modification of the benzotriazole thioamide precursor strategy developed by Rappaport, we have successfully incorporated thioamide versions of the following amino acids through solid phase peptide synthesis (SPPS, Fig. 1): Ala, Phe, Ile, Leu, Pro, and Val using benzotriazoles made from singly *N*-Boc-protected 1,2-phenylenediamine, and Asp, Glu, Arg, and Ser using the *p*-deactivated 4-nitro-1,2-phenylenediamine.[23, 38, 39]

Use of a pre-activated thioacyl benzotriazole avoids standard *in situ* activation requiring nucleophilic substitution reaction of an *O*-alkyl thioester, which is unstable under ambient conditions. Thioacyl benzotriazole monomers (Fig. 1, R = H) are relatively stable toward hydrolysis and can be isolated by chromatography and even cyrstallized, whereas the corresponding nitrobenzotriazoles (Fig. 1, R = NO_2) are typically used without further purification in peptide couplings. While both the Fmoc- and Boc-protected precursors can be generated, we chose Fmoc-based SPPS over the Boc-based procedure to avoid harsh HF cleavage conditions at the end of the synthesis.

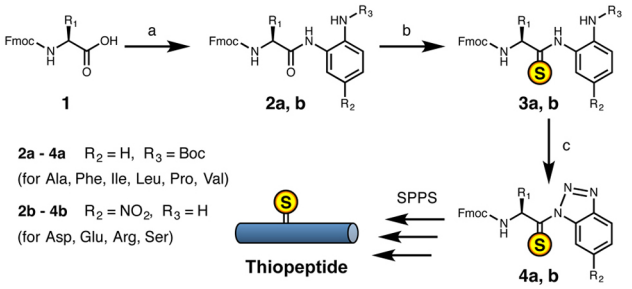


Fig. 1 Thiopeptide Synthesis. Thioamide precursors **4** are synthesized from commercially available Fmoc-protected amino acids in three steps, and then incorporated into peptides *via* solid phase peptide synthesis (SPPS). Typical conditions: a) IBCF, NMM, THF, 0 °C, 15 min; *N*-Boc-OPD or 4-nitro-OPD, overnight; b) Lawesson's Reagent or P_4S_{10}, CH_2Cl_2 or THF, reflux, 2 ~ 18 h; c) 50:50 TFA/CH_2Cl_2, 0 °C, 2 h (for **3a** only); $NaNO_2$, 95:5 AcOH/H_2O, 30 min. IBCF = isobutyl chloroformate; NMM = *N*-methylmorpholine; THF = tetrahydrofuran; OPD = *o*-phenylenediamine; TFA = trifluoroacetic acid.

In SPPS, we demonstrated that thioamides can be successfully incorportaed into thiopeptides, and are generally compatible with common Fmoc SPPS reagents; including activators, non-nucleophilic bases and trifluoroacetic acid (TFA). There can, however, be some sequence-specific issues that one must consider. For example, incorporation of the Glu^S precursor requires that one use dry CH_2Cl_2 as solvent rather than DMF and limit the amount of base used in order to avoid a side-chain mediated desulfurization reaction.[31] The most significant general limitation of thiopeptide synthesis is that prolonged exposure to TFA can lead to cleavage of the peptide backbone at the n+1 position through an Edman degradation-like mechanism. Thus, for longer peptides, there is a tradeoff between driving protecting group removal by longer TFA exposure, and minimizing this degradation by keeping TFA cleavages short. In practice, this limits the yields of fully-deprotected, intact thiopeptides. Of course, thiopeptide length is also restricted by the inherent limitations of SPPS. Therefore, we needed to develop other methods to generate full-sized thioproteins.

We devised a strategy to access full-length proteins by joining a thioamide-containing peptide fragment with the remainder of the protein *via* the native chemical ligation (NCL).[24, 34, 39] The NCL reaction (Fig. 2), developed by Kent, involves the coupling of two unprotected peptide fragments – one with a C-terminal thioester, the other with an N-terminal Cys – in aqueous buffer, leaving a native peptide bond at the junction[40-42] The mild reaction conditions and unprotected nature of reactants are of great advantage for thioamide stability, as well as compatibility with expressed protein, product refolding after ligation, and interfacing with other post-ligation labeling methods. Even for peptides/proteins of more moderate length that may be accessible through simple SPPS, the NCL approach has the benefit that one can produce a relatively short thiopeptide fragment that can be deptotected rapidly with TFA, and then couple this to other peptide fragments at neutral pH.

In our investigations of thiopeptide NCL, we showed that thioamides can be placed in either the thioester or the N-terminal Cys fragment at almost any position except when directly adjacent to the thioester.[34] They are also amenable to tandem ligations requiring deprotection of N-terminal thiazolidines with $MeONH_2$•HCl, and inert to denaturants such as Gdn·HCl, both of which may become necessary in the total synthesis of large thioproteins.

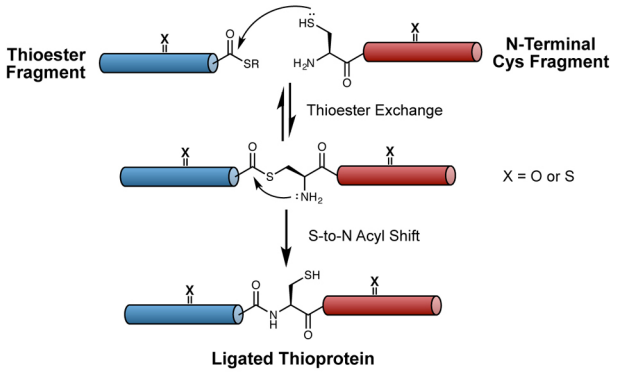


Fig. 2 General Scheme for Native Chemical Ligation (NCL). Two unprotected peptide or protein fragments are reacted in aqueous buffer at nearly neutral pH to yield a ligated product with a native peptide bond at the junction. Thioamides can be placed in either fragment at almost any position, except when directly adjacent to the thioester.

In order to minimize the synthetic efforts necessary to generate full-sized proteins, we typically synthesize a small thiopeptide fragment *via* SPPS, and then express the non-thioamide-containing fragment in *E. coli*. When expressing the N-terminal Cys fragment, we can either generate a fusion protein with an N-terminal protease regonition sequence (most commonly IEGR for Factor Xa) and then reveal the Cys through post-expression proteolysis, or take advantage of the endogenous Met aminopeptidase to cleave the N-terminal Met during expression and then expose the Cys by treating with $MeONH_2$•HCl.[34, 41]

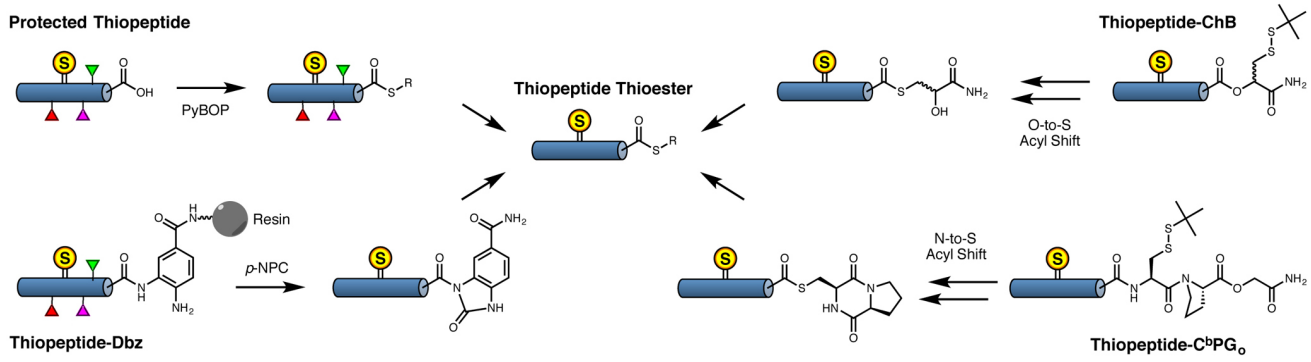


Fig. 3 Synthesis of Thiopeptide Thioesters. In PyBOP activation, a protected peptide is converted to a thioester, and then deprotected to yield the final product. In latent thioester strategies (C^bPG_o, ChB and Nbz), an unprotected peptide is directly converted to a thioester in aqueous buffer.

Organic & Biomolecular Chemistry

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/obc

PERSPECTIVE

When expressing the thioester fragment, we adapted the expressed protein ligation methods of Muir and Cole,[43] where a C-terminal intein fusion was expressed with the desired protein fragment and then converted to a thioester upon treatment with a small molecule thiol, most typically mercaptoethanesulfonate (MESNA).[44] As proof-of-concept, we prepared α-synuclein (αS), a 140 amino acid protein implicated in Parkinson's disease, from an N-terminal thiopeptide fragment synthesized through SPPS and a C-terminal Trp-containing protein fragment expressed in *E. coli*.[34] The Trp/thioamide dual labeling allowed us to monitor αS aggregation by fluorescence quenching, obtaining site-specfic information on misfolding as it occurred *in situ*.

It is worth noting that when placing thioamides in the two NCL components, an N-terminal Cys thiopeptide is relatively straightforward to generate, whereas a thiopeptide thioester is non-trivial to obtain when using Fmoc-based synthesis, and incompatible with Boc-based thioesterification methods. In fact, it was once speculated that ligation of thioamide-containing peptides was "not suitable because the presence of [thioamide] bonds is not compatible with the subsequent synthesis of the thioester moiety."[29] We investigated a variety of thioesterification methods, and showed that thioamides are sufficiently chemically distinct to be left intact in these procedures (Fig. 3). Such methods include conventional solution phase PyBOP activation[34] and more recent *in situ* thioesterification strategies making use of latent thioesters, such as C^bPG_o (N-to-S acyl shift),[34, 45] ChB (O-to-S acyl shift),[46, 47] and Nbz, which promotes thioester formation by activating the C-terminal carbonyl.[39, 48] While the latter three methods generate stereochemically pure products in contrast to the C-terminal residue epimerization commonly observed in PyBOP activation, their yields are generally lower. In some cases, we have used *in situ* thioesterification to identify the desired epimer (when separable by HPLC), and then used PyBOP activation for bulk preparation of the thioester fragment.

Thioamides in complex ligations

NCL reactions allowed us to synthesize larger thioproteins, but two significant barriers remained to having real freedom in what labeled proteins we might make: the ability to attach an arbitrary fluorophore at a specific site of our choosing, and the need for Cys residues at the ligation sites. We have addressed these two challenges by making use of complementary protein semi-synthesis methods: unnatural amino acid (Uaa) mutagenesis and chemoenzymatic protein modification.

In addition to using Trp as a donor fluorophore, we can also make proteins labeled with fluorophore/thioamide PET pairs by conjugating commercially available fluorophore-maleimides at the Cys sites that result from the NCL reaction. While this approach offers site-specific attachment of a wide selection of fluorophores (for proteins with no solvent-exposed Cys other than the ligation site), it limits labeling to the ligation site, which may not be optimal for studying protein folding. Although the fluorophore could be incorporated into proteins *via* NCL along with the thioamide; because the fluorophore is a sidechain modification, we can simplify the SPPS steps and increase protein yields using the Uaa mutagenesis techniques developed by Schultz.[49] Briefly, Uaa mutagenesis is accomplished by transforming *E. coli* with two plasmids, one coding for an unnatural amino acid tRNA synthetase (UaaRS) with $tRNA_{CUA}$ (recognizing a UAG stop codon) and one coding for the protein of interest with a TAG codon replacement at the desired position (Fig. 4A). We have made use of existing synthetases for azide or alkyne amino acids, which we use to attach fluorescent labels *via* bioorthogonal "click" chemistry reactions.[50, 51] This allows us to site-specifically attach a fluorophore at virually any location of the protein, even in the presence of endogenous Cys residues.

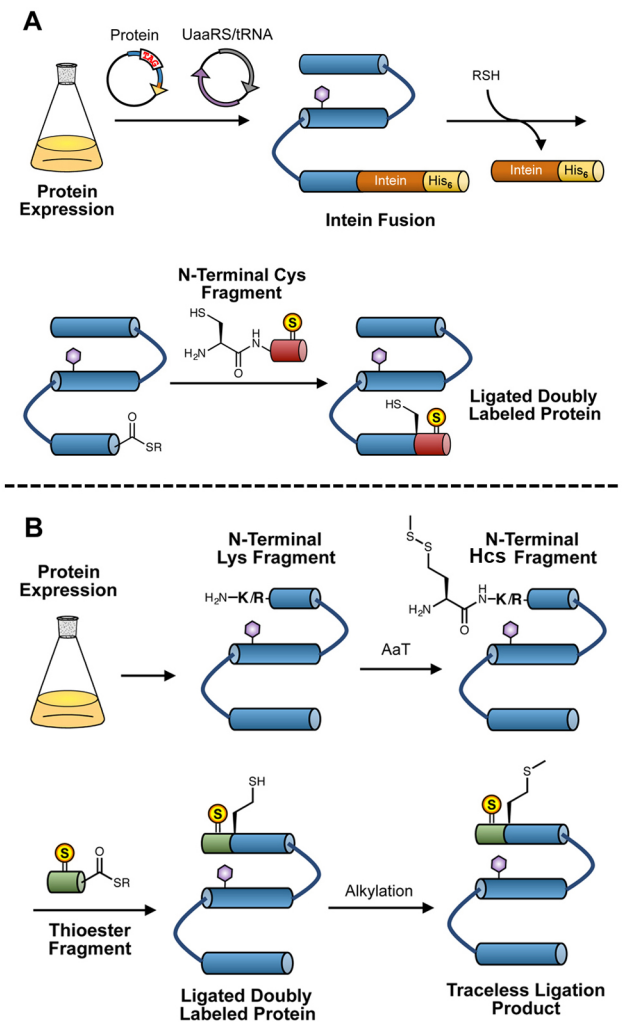


Fig. 4 Complex Ligations. A: Scheme for Uaa/C-terminal thioamide-double labeling. A Uaa-containing intein fusion protein is expressed using a plasmid encoding the protein of interest and a plasmid encoding the requisite UaaRS and tRNA. The intein fusion construct is then purified from *E. coli* cells. Treatment with a thiol reagent converts the intein fusion into a thioester fragment that can be ligated to the thioamide-containing synthetic peptide. B: Scheme for Uaa/N-terminal thioamide double labeling with traceless ligation. A Uaa-containing fragment is expressed and purified from *E. coli* cells. Aminoacyl transferase (AaT) transfers homocysteine (Hcs) to the N-terminus of this fragment. After ligation to the thioamide-containing synthetic thioester fragment, the resulting Hcs residue is selectively methylated to yield Met at the ligation site.

Although post-translational modification offers a great variety of commercially-available fluorophores that can be attached to the protein, it limits labeling to surface-exposed sites or to proteins that can be unfolded and refolded. Intrinsically-fluorescent amino acids are inserted co-translationally, and can therefore be placed at any site where they are tolerated by the protein fold. We have therefore made use of UaaRSs for fluorescent amino acids like Cnf and a hydroxycoumarin amino acid.[52, 53] Since there are relatively few fluorescent Uaas,[54] we have also developed an efficient synthesis of the blue wavelength fluorophore acridon-2-ylalanine (Acd), and through collaboration with Ryan Mehl, generated a synthetase to enable Acd incorporation into proteins expressed in *E. coli*.[55] We are currently working to develop UaaRSs for other fluorescent amino acids as well. Use of these Uaas allows us to produce doubly labeled proteins with complete control of fluorophore location.

To address the other challenge, the need for Cys at the site of ligation, we have adapted approaches that had previously existed in peptide NCL for use with expressed protein fragments. Many investigators have sought to solve the "Cys problem" by functionalizing peptide N-termini with Uaas that can be used in ligation, and then converted to other amino acids.[56] Danishefsky and others have developed strategies for the total chemical synthesis of proteins using synthetic Cys analogs that are converted to Leu, Ile, Phe, Val, Thr, Glu, Trp, Arg, Asp, Gln, Pro, and Lys by desulfurization.[57-63] Homocysteine (Hcs) has been used for ligation and then converted to Met by selective methylation.[64-67] Since these ligation handles needed to be incorporated synthetically, they could be used in total syntheses of proteins, but not with expressed protein fragments. We have made use of ligation site masking by both alkylation and desulfurization, and we have worked to expand the use of these erasable ligation handles by finding ways to attach them to the N-termini of cellularly expressed proteins.

One enzyme that naturally functionalizes protein N-termini is the aminoacyl transferase (AaT) enzyme from *E. coli*, which transfers Phe, Leu, or Met from an aminoacyl tRNA.[68, 69] AaT transfers these amino acids to N-terminal Arg or Lys residues regardless of the adjacent sequence.[70] Previous work from Tirrell and Sisido, as well as our own discovery of small molecule AaT substrates, has shown that AaT could transfer unnatural amino acids as well.[71, 72] We have used AaT to functionalize proteins with erasable Cys analogs for ligation. AaT can deliver disulfide-protected Hcs to the N-termini of various proteins under mild conditions, where it is used in a ligation reaction, and then converted to Met by alkylation.[73] We utilized this strategy for N-terminal thioamide labeling of αS, where protected Hcs was transferred onto the N-terminus of a fragment of αS by AaT. This fragment was then used in NCL with a thiopeptide, and after ligation, alkylation was shown to be selective for the Cys sulfur over the thioamide sulfur (when reaction times were limited). Moreover, we have combined AaT functionalization and NCL with Uaa mutagenesis, to synthesize αS variants labeled with Cnf and a thioamide with a Met left at the ligation site, a so-called traceless synthesis (Fig. 4B).[74]

Future directions

The methodological developments made in the early years of our research group have enabled us to make sufficient quantities of full-sized thioproteins to begin studying questions of biological interest. However, there are certainly many imrpovements that can still be made in order to improve the efficiency of thioprotein synthesis and the scope of proteins and labeling locations that one can consider in a folding experiment.

Organic & Biomolecular Chemistry

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/obc

PERSPECTIVE

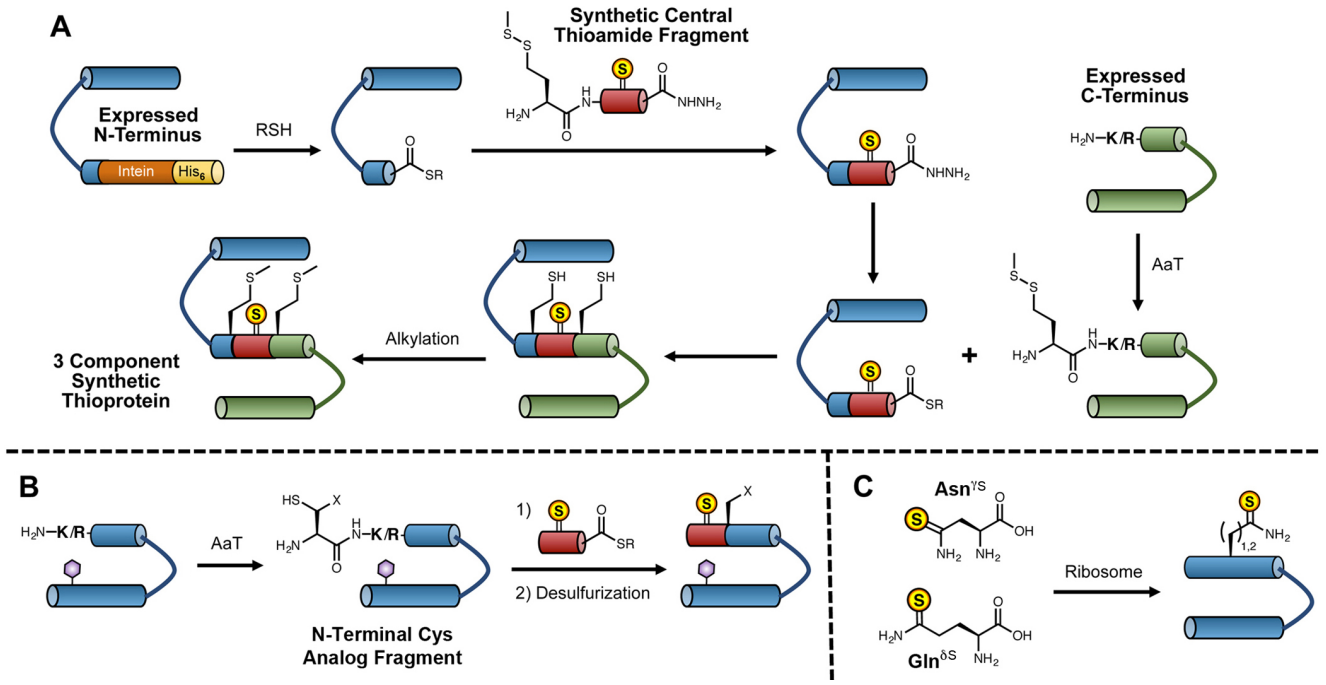


Fig. 5 Future Directions. A: Strategy for central thioamide labeling. The synthetic thioamide fragment contains an N-terminal Hcs and C-terminal acyl hydrazide. After ligating the synthetic thioamide fragment to the expressed N-terminus fragment, the acyl hyrdazide can be converted to a thioester, allowing a second ligation with the expressed C-terminus fragment. Selective methylation after ligation yields the final traceless ligation product. B: Alternative strategy for traceless ligation of thioamide containing proteins. A mutant AaT transfers a *β*-thiol analogue of an amino acid onto the N-terminus of an expressed protein fragment. After ligating the expressed protein fragment to a synthetic thioamide-containing peptide, selective desulfurization removes the *β*-thiol at the ligation site, yielding a traceless ligation product. C: Alternative thioamide probes. Side-chain thioamide analogs of Gln and Asn that have the potential for ribosomal incorporation into proteins.

To date, our laboratory has only ligated thioamide-containing synthetic peptides to the N- or C-terminal regions of proteins. In order to incorporate thioamides into the middle of a protein, a three part ligation would be necessary. Obtaining a doubly-labeled protein with a central thioamide requires all of the previously mentioned procedures. There are many possible operations for carrying out such a synthesis, but one currently under investigation in our laboratory is as follows (Fig. 5A). First, an expressed C-terminal protein with a Lys or Arg residue on the N-terminus would be functionalized with an erasable Cys analog (e.g., Hcs). In parallel, the N-terminal fragment would be expressed as an intein fusion, and ligated to a thiopeptide with a C-terminal acyl hydrazide. After this first ligation, the acyl hydrazide of the product would be activated and converted to a thioester using the method of Liu.[75] This thioester would then be ligated to the Hcs-functionalized C-terminal fragment, and both Hcs residues alkylated to form Met and complete the thioprotein synthesis. The initial results of our three-component synthesis of thioamide-modified calmodulin show this stratgey to be promising.[76]

We are also working to expand the use of thioamide-compatible erasable ligation handles beyond Hcs. As noted above, a common way to increase the number of potential ligation sites for NCL is to perform a desulfurization reaction after ligation. Our preliminary data show that selective desulfurization of Cys and other *β*-thiols in the presence of thioamides is possible, and leaves the thioamide intact.[76] Given our success in transferring Hcs, we are exploring the possibility of transferring these Cys analogs with AaT (Fig. 5B). Wild type AaT is not able to transfer these *β*-thiol containing amino acids, therefore we have undetaken the evolution of an AaT mutant that is capable of transferring *β*-thiol containing amino acids. Such an enzyme would not only enable more complex thiopeptide ligations in our own laboratory, but would be beneficial to the whole protein ligation community.

Finally, our laboratory has thus far focused on incorporation of backbone thioamides into proteins, since the ubiquity of the peptide bond makes their use as spectroscopic probes very general. However, there is certainly also utility in thioamide-containing sidechain derivatives of glutamine ($Gln^{\delta S}$) and asparagine ($Asn^{\gamma S}$). These can serve as complementary probes, to be utilized in contexts where backbone thioamides might disturb

the natural fold of a protein. They are also more readily amenable than backbone thioamides to ribosomal incorporation into proteins because they do not require a modifed attachment to the 3´end of tRNA, as backbone thioamides would. $Gln^{\delta S}$ and $Asn^{\gamma S}$ building blocks for SPPS have been described before,[77, 78] and we have been successful in incorporating these into short model peptides for photophysical characterization. We find that there is substantial quenching of Cnf by these thioamides, as expected.[76] If these sidechain probes can be incorporated into peptides and proteins by ribosomes during *in vitro* translation, it would allow us to install them into proteins that are too large or insufficently robust to be made by NCL. Ribosomal incorportion of $Gln^{\delta S}$ and $Asn^{\gamma S}$ (Fig. 5C) would also be a first step toward the biosynthesis of proteins containing backbone thioamides.

Notes and references

Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323 USA. Fax: +1-215-573-2112; Tel: +1-215-746-2221; E-mail:ejpetersson@sas.upenn.edu

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

‡ Footnotes should appear here. These might include comments relevant to but not central to the matter under discussion, limited experimental and spectral data, and crystallographic data.

1. Note on "thioamide" nomeclature: In the literature, the thioamide functional group has historically been referred to as a thionoamide, thioxoamide, endothiopeptide, mercaptoamide, thiopeptide unit, thiodepsipeptide, and thiocarboxamide. For the sake of brevity and to be consistent with current Organic Chemistry usage patterns, we prefer thioamide. Formally, according to IUPAC recommendations, thioamide backbone modifications in peptides and proteins should be annotated systematically as ψ[CS-NH]. For the sake of brevity, we have previously used a prime symbol (´) to indicate the presence of a thioamide between two residues. However, our recent work in the protease field has led us to realize that this may be confusing, since the prime symbol is used to indicate a residue's realtionship to the scissile bond. Thus, from this publication forward, we will use the superscript S (^{S}) to indicate the presence of the thioamide. Additionally, it should be noted that we use the term thiopeptides to describe thioamide-containing peptides while this term is often also used to describe thiazole- and thiazolene-containing natural products such as thiostrepton.
2. A. Bondi, *J Phys Chem-Us*, 1964, **68**, 441-451.
3. R. Bardi, A. M. Piazzesi, C. Toniolo, O. E. Jensen, R. S. Omar and A. Senning, *Biopolymers*, 1988, **27**, 747-761.
4. T. F. M. La Cour, H. A. S. Hansen, K. I. M. Clausen and S. O. Lawesson, *Int. J. Peptide Protein Res.*, 1983, **22**, 509-512.
5. P. Wipf and G. B. Hayes, *Tetrahedron*, 1998, **54**, 6987-6998.
6. S. Misra and U. Tewari, *Transit. Metal Chem.*, 2002, **27**, 120-125.
7. E. P. Dudek and G. O. Dudek, *J. Org. Chem.*, 1967, **32**, 823-824.
8. J. M. Goldberg, S. Batjargal and E. J. Petersson, *J. Am. Chem. Soc.*, 2010, **132**, 14718-14720.
9. J. H. Miwa, L. Pallivathucal, S. Gowda and K. E. Lee, *Org. Lett.*, 2002, **4**, 4655-4657.
10. G. Kalaus, I. Greiner, M. Kajtar-Peredy, J. Brlik, L. Szabo and C. Szantay, *J. Org. Chem.*, 1993, **58**, 6076-6082.
11. T. Fukuyama, L. Li, A. A. Laird and R. K. Frank, *J. Am. Chem. Soc.*, 1987, **109**, 1587-1589.
12. A. Reiner, D. Wildemann, G. Fischer and T. Kiefhaber, *J. Am. Chem. Soc.*, 2008, **130**, 8079-8084.
13. A. Choudhary and R. T. Raines, *ChemBioChem*, 2011, **12**, 1801-1807.
14. T. Lincke, S. Behnken, K. Ishida, M. Roth and C. Hertweck, *Angew. Chem. Int. Ed.*, 2010, **49**, 2011-2013.
15. U. Ermler, W. Grabarse, S. Shima, M. Goubeaud and R. K. Thauer, *Science*, 1997, **278**, 1457-1462.
16. J. Xie, A. Okano, J. G. Pierce, R. C. James, S. Stamm, C. M. Crane and D. L. Boger, *J. Am. Chem. Soc.*, 2011, **134**, 1284-1297.
17. A. Okano, R. C. James, J. G. Pierce, J. Xie and D. L. Boger, *J. Am. Chem. Soc.*, 2012, **134**, 8790-8793.
18. P. Wipf and S. Venkatraman, *J. Org. Chem.*, 1996, **61**, 8004-8005.
19. W. C. Jones, J. J. Nestor and V. D. Vigneaud, *J. Am. Chem. Soc.*, 1973, **95**, 5677-5679.
20. M. D. Bond, B. Holmquist and B. L. Vallee, *J. Inorg. Biochem.*, 1986, **28**, 97-105.
21. R. W. Newberry, B. VanVeller, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2013, **135**, 7843-7846.
22. R. M. Culik, H. Jo, W. F. DeGrado and F. Gai, *J. Am. Chem. Soc.*, 2012, **134**, 8026-8029.
23. J. H. Miwa, A. K. Patel, N. Vivatrat, S. M. Popek and A. M. Meyer, *Org. Lett.*, 2001, **3**, 3373-3375.
24. E. J. Petersson, J. M. Goldberg and R. F. Wissner, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6827-6837.
25. M. R. Truter, *J. Chem. Soc.*, 1960, 997-1007.
26. L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, Cornell University Press, 1960.
27. H.-J. Lee, Y.-S. Choi, K.-B. Lee, J. Park and C.-J. Yoon, *J. Phys. Chem. A*, 2002, **106**, 7010-7017.
28. R. H. Judge, D. C. Moule and J. D. Goddard, *Can. J. Chem.*, 1987, **65**, 2100-2105.
29. D. Wildemann, C. Schiene-Fischer, T. Aumüller, A. Bachmann, T. Kiefhaber, C. Lücke and G. Fischer, *J. Am. Chem. Soc.*, 2007, **129**, 4910-4918.
30. F. G. Bordwell, D. J. Algrim and J. A. Harrelson, *J. Am. Chem. Soc.*, 1988, **110**, 5903-5904.
31. J. M. Goldberg, S. Batjargal, B. S. Chen and E. J. Petersson, *J. Am. Chem. Soc.*, 2013, **135**, 18651-18658.
32. J. M. Goldberg, L. C. Speight, M. W. Fegley and E. J. Petersson, *J. Am. Chem. Soc.*, 2012, **134**, 6088-6091.
33. J. M. Goldberg, R. F. Wissner, A. M. Klein and E. J. Petersson, *Chem. Commun.*, 2012, **48**, 1550-1552.
34. S. Batjargal, Y. J. Wang, J. M. Goldberg, R. F. Wissner and E. J. Petersson, *J. Am. Chem. Soc.*, 2012, **134**, 9172-9182.

35. J. M. Goldberg, X. Chen, N. Meinhardt, D. C. Greenbaum and E. J. Petersson, *J. Am. Chem. Soc.*, 2014, **136**, 2086-2093.
36. Y. Huang, G. Jahreis, C. Lücke, D. Wildemann and G. Fischer, *J. Am. Chem. Soc.*, 2010, **132**, 7578-7579.
37. T. M. Vishwanatha, N. Narendra, B. Chattopadhyay, M. Mukherjee and V. V. Sureshbabu, *The Journal of Organic Chemistry*, 2012, **77**, 2689-2702.
38. M. A. Shalaby, C. W. Grote and H. Rapoport, *J. Org. Chem.*, 1996, **61**, 9045-9048.
39. S. Batjargal, Y. Huang, Y. J. Wang and E. J. Petersson, *J. Peptide Sci.*, 2014, **20**, 87-91.
40. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776-779.
41. C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem. Int. Ed.*, 2008, **47**, 10030-10074.
42. L. Raibaut, N. Ollivier and O. Melnyk, *Chem. Soc. Rev.*, 2012, **41**, 7001-7015.
43. T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci.*, 1998, **95**, 6705-6710.
44. E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640-6646.
45. T. Kawakami and S. Aimoto, *Tetrahedron*, 2009, **65**, 3871-3877.
46. P. Botti, M. Villain, S. Manganiello and H. Gaertner, *Org. Lett.*, 2004, **6**, 4861-4864.
47. E. A. George, R. P. Novick and T. W. Muir, *J. Am. Chem. Soc.*, 2008, **130**, 4914-4924.
48. J. B. Blanco-Canosa and P. E. Dawson, *Angew. Chem. Int. Ed.*, 2008, **47**, 6851-6855.
49. L. Wang and P. G. Schultz, *Angew. Chem. Int. Ed.*, 2005, **44**, 34-66.
50. S. Batjargal, C. R. Walters and E. J. Petersson, *J. Am. Chem. Soc.*, In Press.
51. E. M. Sletten and C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2009, **48**, 6974-6998.
52. H. Taskent-Sezgin, J. Chung, V. Patsalo, S. J. Miyake-Stoner, A. M. Miller, S. H. Brewer, R. A. Mehl, D. F. Green, D. P. Raleigh and I. Carrico, *Biochemistry*, 2009, **48**, 9040-9046.
53. J. Y. Wang, J. M. Xie and P. G. Schultz, *J. Am. Chem. Soc.*, 2006, **128**, 8738-8739.
54. L. C. Speight, M. Samanta and E. J. Petersson, *Aust. J. Chem.*, 2014, **67**, 686-700.
55. L. C. Speight, A. K. Muthusamy, J. M. Goldberg, J. B. Warner, R. F. Wissner, T. S. Willi, B. F. Woodman, R. A. Mehl and E. J. Petersson, *J. Am. Chem. Soc.*, 2013, **135**, 18806-18814.
56. C. T. T. Wong, C. L. Tung and X. C. Li, *Mol. Biosys.*, 2013, **9**, 826-833.
57. J. Chen, Q. Wan, Y. Yuan, J. L. Zhu and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2008, **47**, 8521-8524.
58. J. Chen, P. Wang, J. L. Zhu, Q. Wan and S. J. Danishefsky, *Tetrahedron*, 2010, **66**, 2277-2283.
59. D. Crich and A. Banerjee, *J. Am. Chem. Soc.*, 2007, **129**, 10064-10065.
60. C. Haase, H. Rohde and O. Seitz, *Angew. Chem. Int. Ed.*, 2008, **47**, 6807-6810.
61. Z. Harpaz, P. Siman, K. S. A. Kumar and A. Brik, *Chembiochem*, 2010, **11**, 1232-1235.
62. S. D. Townsend, Z. P. Tan, S. W. Dong, S. Y. Shang, J. A. Brailsford and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3912-3916.
63. R. L. Yang, K. K. Pasunooti, F. P. Li, X. W. Liu and C. F. Liu, *J. Am. Chem. Soc.*, 2009, **131**, 13592-13593.
64. B. Aussedat, B. Fasching, E. Johnston, N. Sane, P. Nagorny and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3532-3541.
65. J. P. Tam and Q. T. Yu, *Biopolymers*, 1998, **46**, 319-327.
66. K. Pachamuthu and R. R. Schmidt, *Synlett*, 2003, 659-662.
67. A. Saporito, D. Marasco, A. Chambery, P. Botti, S. M. Monti, C. Pedone and M. Ruvo, *Biopolymers*, 2006, **83**, 508-518.
68. K. Suto, Y. Shimizu, K. Watanabe, T. Ueda, S. Fukai, O. Nureki and K. Tomita, *EMBO J.*, 2006, **25**, 5942-5950.
69. K. Watanabe, Y. Toh, K. Suto, Y. Shimizu, N. Oka, T. Wada and K. Tomita, *Nature*, 2007, **449**, 867-U865.
70. A. Varshavsky, *Nat. Struct. Mol. BIol.*, 2008, **15**, 1238-1240.
71. R. E. Connor, K. Piatkov, A. Varshavsky and D. A. Tirrell, *Chembiochem*, 2008, **9**, 366-369.
72. M. Taki, H. Kuroiwa and M. Sisido, *Chembiochem*, 2008, **9**, 719-722.
73. T. Tanaka, A. M. Wagner, J. B. Warner, Y. X. J. Wang and E. J. Petersson, *Angew. Chem. Int. Ed.*, 2013, **52**, 6210-6213.
74. R. F. Wissner, A. M. Wagner, J. B. Warner and E. J. Petersson, *Synlett*, 2013, **24**, 2454-2458.
75. G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui and L. Liu, *Angew. Chem. Int. Ed.*, 2011, **50**, 7645-7649.
76. Unpublished Results.
77. T. L. Hendrickson and B. Imperiali, *Biochemistry*, 1995, **34**, 9444-9450.
78. J. M. Sanderson, P. Singh, C. W. G. Fishwick and J. B. C. Findlay, *J. Chem. Soc. Perkin 1*, 2000, 3227-3231.