

Published in final edited form as:

*J Am Chem Soc.* 2013 May 1; 135(17): 6529–6540. doi:10.1021/ja4005943.

## Labeling Proteins with Fluorophore/Thioamide FRET Pairs by Combining Unnatural Amino Acid Mutagenesis and Native Chemical Ligation

Rebecca F. Wissner, Solongo Batjargal, Colin M. Fadzen, and E. James Petersson

University of Pennsylvania, Department of Chemistry, 231 South 34th Street, Philadelphia, Pennsylvania, 19104-6323 USA

E. James Petersson: ejpetersson@sas.upenn.edu

### Abstract

We have recently shown that *p*-cyanophenylalanine (Cnf) and a thioamide can be used as a minimally-perturbing Förster resonant energy transfer (FRET) pair to monitor protein conformation. We have also shown that thioamide analogs of natural amino acids can be incorporated into full-sized proteins through native chemical ligation. For intermolecular studies with Cnf/thioamide FRET pairs, Cnf can be incorporated into proteins expressed in *E. coli* through unnatural amino acid mutagenesis using a Cnf-specific tRNA synthetase. For intramolecular studies, a Cnf-labeled protein fragment can be expressed in *E. coli* and then ligated to a thioamide-labeled peptide synthesized on solid phase. This combination of methods allows for rapid access to double-labeled proteins with a minimum of unnecessary chemical synthesis. We demonstrate the utility of this approach by studying the binding of peptides to the protein calmodulin and by determining the orientation of the N- and C-termini in the amyloidogenic protein  $\alpha$ -synuclein.

### INTRODUCTION

Fluorescence spectroscopy can be used to monitor conformational changes in proteins by exploiting distance-dependent interactions such as Förster resonant energy transfer (FRET).<sup>1,2</sup> In a FRET experiment, a donor fluorophore is paired with an acceptor fluorophore or quencher. An increase in chromophore separation leads to an increase in donor fluorescence and a decrease in acceptor fluorescence or quenching with an  $R^{-6}$  distance dependence. The utility of the data obtained in these experiments often depends on the ability to label the proteins at appropriate locations. Bulky fluorophores are limited in the number of places that the label can be placed, and are subject to the concern that the label itself alters the observed motion.<sup>3</sup> Therefore, our laboratory has developed the thioamide – sulfur replacement of oxygen in the peptide backbone – as an extremely small label compatible with virtually any position in the protein. (Fig. 1)

Thioamides can quench fluorophores that have spectral emission overlap with thioamide absorption, such as cyanophenylalanine (Cnf) and tyrosine, through a FRET mechanism.<sup>4,5</sup>

Correspondence to: E. James Petersson, ejpetersson@sas.upenn.edu.

#### ASSOCIATED CONTENT

**Supporting Information.** Characterization of all peptides and proteins, assessment of CaM mutants by native PAGE gel and circular dichroism (CD), primary fluorescence data for CaM experiments, analysis of fluorescence data using Förster theory, HPLC analysis of NCL reactions, details of UV/Vis, fluorescence, and trypsin/MALDI MS characterization of NCL products, primary fluorescence data and PAGE gel analysis of  $\alpha$ S refolding experiments. This material is available free of charge via the Internet at <http://pubs.acs.org>. Primary data used to generate figures and tables have been digitally archived and can be obtained by emailing the corresponding author.

Other fluorophores are quenched through a photoinduced electron transfer (PET) mechanism, like tryptophan, coumarin, and acridone.<sup>5,6</sup> Here we will focus on thioamide FRET partners, for which more reliable distance information can be obtained by applying Förster theory. To use thioamide FRET quenching to study protein conformation, we generate constructs in which we label the protein with a Cnf donor fluorophore and a thioamide analog of one of the natural amino acids. We denote these analogs using either the three or one letter amino acid code with a prime symbol indicating the thiocarbonyl (thioleucine is shown in Figure 1). Distance constraints are assigned by computing quenching efficiency ( $E_Q = 1 - F_{\text{Thio}}/F_{\text{Oxo}}$ ) from the fluorescence of equimolar concentrations of Cnf/thioamide-labeled protein ( $F_{\text{Thio}}$ ) and a Cnflabeled control ( $F_{\text{Oxo}}$ ). After appropriate corrections to the Förster radius ( $R_0$ ) for each Cnf location, the Cnf/thioamide separation ( $R_{\text{FRET}}$ ) is determined according to Förster theory. (Fig. 1, and discussion below) These distance measurements are used to monitor conformational changes in the protein.

All of our initial studies used small peptides and proteins in which the thioamide was incorporated by solid phase peptide synthesis (SPPS). In order to use fluorophore/thioamide pairs to monitor structural changes in full-sized proteins, we needed to identify methods for incorporating these two modifications with high efficiency. We have recently published an account of our studies of the compatibility of thioamides with native chemical ligation (NCL) reactions.<sup>7</sup> NCL is a valuable method for synthesizing proteins by fragment condensation of a C-terminal thioester peptide and an N-terminal Cys peptide in buffered aqueous media.<sup>8,9</sup> We found that thioamides could be placed at nearly any position in the thioester or Cys peptide fragments. Furthermore, we found that thioamides could be used in NCL reactions in which the N-terminal Cys fragment was expressed as a protein construct in *E. coli* (expressed protein ligations). In general, we wish to minimize the need for manual synthesis by maximizing the role of biosynthesis in preparing labeled proteins. These protein ligations can limit the portion that must be prepared by SPPS to a few amino acids around the thioamide. SPPS is deemed necessary for the incorporation of backbone thioamide modifications, but sidechain fluorophores like Cnf can be incorporated during ribosomal biosynthesis.

To easily incorporate Cnf, we have employed unnatural amino acid (Uaa) mutagenesis methods pioneered by Schultz and coworkers.<sup>10,11</sup> These methods allow one to insert a Uaa at a specific site in the protein by using a so-called orthogonal aminoacyl tRNA synthetase (RS) that is selective for the Uaa. The UaaRS charges a tRNA that recognizes the UAG stop codon. The UAG stop codon is mutated into the protein coding sequence at the site of interest. Plasmids coding for the UaaRS, tRNA, and protein of interest are transformed into *E. coli*. After growth in the presence of the Uaa, protein containing the unnatural amino acid is purified for use in biochemical experiments or subsequent NCL reactions.

Here, we made use of Uaa mutagenesis to incorporate donor fluorophores for thioamide FRET experiments examining both protein/protein interactions and intramolecular measurements of protein folding. In intermolecular experiments, no manipulation of the Uaa-containing protein is required following purification. The protein is simply mixed with a thioamide-labeled peptide or protein and quenching observed. We use this strategy to examine a series of pairwise interactions in which Cnf is incorporated in the calcium sensor protein calmodulin (CaM) and the thioamide is incorporated in a CaM-binding peptide. The combinatorial nature of these experiments allows one to rapidly generate a matrix of  $m \times n$  FRET measurements from  $m$  Cnf-labeled proteins and  $n$  thioamide-labeled peptides. To carry out intramolecular experiments, we combined Uaa mutagenesis with NCL. Cnf was incorporated by Uaa mutagenesis into protein constructs in which the termini could be used in ligation to thiopeptides. Strategies for ligating to both the C- and N-termini are described.

The resulting double-labeled proteins can then be used in intramolecular folding experiments. Again, combinatorial assembly can be used to make  $m \times n$  double-labeled proteins from  $m$  N-terminal fragments and  $n$  C-terminal fragments. We applied our double-labeling strategy to study conformational changes in  $\alpha$ -synuclein ( $\alpha$ S), an amyloidogenic protein found to aggregate in Parkinson's Disease.<sup>12–14</sup>

Our methods should be applicable to a wide variety of proteins: essentially any proteins that express well in *E. coli* and are amenable to refolding after ligation. Moreover, since this combination of Uaa mutagenesis and NCL to label proteins on the backbone and sidechain is unprecedented (examples of sidechain ligation using Uaa mutagenesis exist),<sup>15,16</sup> the considerations here should be useful to those who wish to combine these two methods for other reasons, such as incorporating olefin or D-amino acid modifications in combination with unnatural sidechains.<sup>17–20</sup>

## RESULTS AND DISCUSSION

To begin combining Uaa mutagenesis with thioamides, we used CaM as a model system. We chose CaM because it expresses well in *E. coli* and it has been thoroughly studied from a structural and biophysical standpoint.<sup>21–27</sup> CaM has been shown to bind a variety of helical peptides derived from its endogenous protein interaction partners.<sup>28,29</sup> These peptides bind with high affinity in a central groove formed by the folding of the N- and C-terminal CaM lobes. In particular, we studied the binding of a peptide derived from the intracellular domain of an olfactory cyclic nucleotide-gated ion channel (pOCNC).<sup>30</sup> This peptide binds with a low nM affinity, and its CaM-bound structure has been determined previously by NMR.<sup>30,31</sup> Therefore, we expected that 1:1 mixtures of 10  $\mu$ M each CaM and pOCNC would be stably bound and that the NMR structures could be used to determine interatomic distances. These measurements would then be compared to the corresponding distances derived from our FRET measurements.

Since Cnf is excited at 240 nm and fluoresces with a maximum at 295 nm, we began by mutating the two endogenous Tyr residues in CaM to Phe, generating a construct hereafter referred to as CaM<sup>F</sup>.<sup>32</sup> This all Phe background ensures that Tyr fluorescence does not confound our ability to interpret Cnf data. Both the extinction coefficient and quantum yield of Phe ( $\epsilon_{240} = 65 \text{ cm}^{-1} \cdot \text{M}^{-1}$ ,  $\Phi = 0.022$ ) are substantially smaller than the values for Cnf ( $\epsilon_{240} = 10,333 \text{ cm}^{-1} \cdot \text{M}^{-1}$ ,  $\Phi = 0.110$ ), so in these proteins all fluorescence emission at 295nm should derive from Cnf.<sup>32,33</sup>

To incorporate Cnf, we used a plasmid containing a UaaRS/tRNA pair previously selected for Cnf incorporation by Mehl and coworkers.<sup>34</sup> This was paired with a plasmid encoding CaM under IPTG control. We expressed CaM<sup>F</sup> in M9 media supplemented with Cnf and purified the full-length protein by two rounds of phenyl sepharose chromatography. Four positions were chosen for incorporation of Cnf: F<sub>13</sub>, F<sub>17</sub>, F<sub>93</sub>, and Y<sub>100</sub>. Four pOCNC peptides were synthesized: an all Phe oxopeptide (pOCNC), and three thiopeptides – pOCNC-F'<sub>1</sub>, pOCNC-L'<sub>11</sub>, and pOCNC-F'<sub>16</sub>. Before performing fluorescence titration experiments, native PAGE gel analysis was used to ensure that each CaM<sup>F</sup>/pOCNC combination formed stable complexes. This allowed us to eliminate CaM<sup>F</sup>F\*<sub>17</sub> from further binding experiments (see Supporting Information).

For each of the other Cnf mutants, we performed the same set of experiments, which are described using CaM<sup>F</sup>F\*<sub>100</sub> as an example. Fluorescence wavelength scans were acquired for mixtures of 10  $\mu$ M CaM<sup>F</sup>F\*<sub>100</sub> with between 0 and 15  $\mu$ M pOCNC or pOCNC-F'<sub>1</sub>. For each concentration, the fluorescence emission from the pOCNC-F'<sub>1</sub> complex ( $F_{\text{Thio}}$ ) was normalized to the fluorescence emission from the pOCNC complex ( $F_{\text{Oxo}}$ ). For F\*<sub>100</sub> a

concentration-dependent decrease was observed for  $F_{\text{Thio}}$ , while no change was seen  $F_{\text{Oxo}}$ . Plotting  $F_{\text{Thio}}/F_{\text{Oxo}}$  (or  $1 - E_Q$ ) as a function of concentration shows that binding reaches saturation at 1:1 CaM<sup>F\*100</sup>/pOCNC-F'1 (Fig. 3). The quenching (FRET) efficiency can then be used to calculate a separation according to  $R_{\text{FRET}} = R_0(1/E_Q - 1)$ . We observed 25% quenching at saturating concentrations of pOCNC-F'1, which translates into a 20 Å separation using Förster theory. This is in excellent agreement with the 19 Å average separation computed from the 20 lowest energy structures deposited from the NMR structure determination by Contessa *et al.*<sup>31</sup> These calculations were performed using  $\Phi = 0.110$  and an orientation factor ( $\kappa^2$ ) of 2/3.

We completed comparable experiments for each pair of Cnf mutant and thiopeptide. The separation in the seven pairs spanned a distance range from 8 to 31 Å, covering the range that is accessible with the Cnf/thioamide FRET pair. In the case of CaM<sup>F\*100</sup>/pOCNC-F'1, no change in fluorescence was observed in the “oxoamide control” experiment. For other mutants, some changes in fluorescence were observed for pOCNC binding. (See Supporting Information) Regardless, these effects can be deconvoluted by calculating  $E_Q$  relative to the oxoamide control as we do. Such oxoamide control experiments are essential as many factors can influence fluorescence.<sup>35–37</sup> Therefore, we always compare any thioamide quenching measurement to the most relevant control, in this case the CaM/pOCNC oxoamide complex.

The oxoamide complex is also valuable in determining the appropriate value of  $\Phi$  to be used in FRET distance calculations. The changes in fluorescence intensity upon pOCNC binding represent a change in Cnf quantum yield as it moves into a new environment. For example, the emission of CaM<sup>F\*93</sup> starts out at 3.8% of the emission of an equimolar concentration of CaM<sup>F\*100</sup>, but it increases 3.2-fold upon pOCNC binding. Therefore, the relevant quantum yield for the CaM<sup>F\*93</sup>/pOCNC complex is 0.014. We use this value in our calculation of  $R_{\text{fret}}$  for all thiopeptide complexes with CaM<sup>F\*93</sup>. Similar corrections to the quantum yield were performed for CaM<sup>F\*13</sup>. Correction of the quantum yield of CaM<sup>F\*100</sup> was not necessary, presumably because Cnf is solvent-exposed at this position.

The distances calculated from our FRET measurements are collected in Table 1. All measurements are within 3 Å of the average value observed in the NMR structure except for F\*100/F'16. Moreover, our measurements are always correct in predicting the relative order of thioamide proximity for a given Cnf mutant. While it may be surprising that we achieve reasonable agreement between theory and experiment using a value of 2/3 for  $\kappa^2$ , deriving  $\kappa^2$  values from the 20 lowest energy NMR structures has only a small effect on this agreement. (See Supporting Information) The  $R_{\text{FRET}}$  values obtained using  $\kappa^2 = 2/3$  are more representative of the application of thioamide FRET in a structure determination experiment (i.e. one in which the target structure is unknown and  $\kappa^2$  would be difficult to fit accurately).

These CaM experiments involved intermolecular quenching, but our method, like FRET in general, is equally applicable to intramolecular conformational changes. However, obtaining Cnf/thioamide double-labeled proteins represents a biosynthetic challenge, as one ideally wishes to have complete freedom in placing the chromophores within the protein sequence. While this is easily accomplished in small peptides made by SPPS, and conceivably could be done by NCL, we wished to use a combination of NCL and Uaa mutagenesis to minimize unnecessary peptide synthesis. Therefore, we assessed the compatibility of these two methods. We chose  $\alpha$ S as a model system because we had previously shown that thioamides could be incorporated into  $\alpha$ S, and we wished to compare PET quenching of Trp in  $\alpha$ S to FRET quenching of Cnf at similar locations.<sup>7</sup>

First, we examined thioamide incorporation at the  $\alpha$ S N-terminus. For protein expression, we used the UaaRS/tRNA plasmid pDULE2-Cnf and a plasmid encoding  $\alpha$ S with an N-terminal truncation that could be revealed after proteolysis of a His<sub>10</sub> purification sequence using Factor Xa (Fig. 5A).<sup>38</sup> The proteolyzed protein contains an N-terminal Cys for use in NCL with the synthetic thioester. The Factor Xa method provides a general method for generating NCL constructs containing unnatural amino acids, but in the specific context of  $\alpha$ S, prolonged incubation with Factor Xa can lead to protein loss *via* aggregation. Therefore, we have also used constructs directly encoding  $\alpha$ S with an N-terminal truncation plus Met in a fashion similar to that described by Hejjaoui *et al.*<sup>39</sup> The N-terminal Met residue is removed *in vivo* by Met aminopeptidase to reveal the requisite Cys.<sup>40</sup> Protein harvested from *E. coli* contains an N-terminal thiazolidine ring formed by the condensation of Cys with either pyruvate or acetaldehyde. This is treated with methoxylamine hydrochloride to reveal free Cys for NCL reactions. Regardless of the method used in producing the expressed fragment, once purified, usage in subsequent NCL reactions is identical.

Synthesis of a thioamide-containing thioester peptide using Fmoc chemistry could be accomplished by one of two methods. The thiopeptide can be synthesized on Cl-Trt resin, cleaved under mild conditions (10% AcOH) to retain the sidechain protecting groups, and then activated using PyBOP to form the thioester (Fig. 5C).<sup>41,42</sup> This method can lead to epimerization of the  $\alpha$ -carbon of the C-terminal amino acid, but this can be avoided by using short (< 30 min) reaction times during PyBOP activation.<sup>43</sup> A more significant inherent limitation is the insolubility of protected peptides > 15 amino acids in most solvents.<sup>41</sup> This limits use of the PyBOP method to short peptides.

The other method makes use of a thioester that is masked during synthesis as a disulfide (Fig. 5D). Several variations of masked thioester exist. In our previous work, we used a three residue C<sup>b</sup>PG<sub>o</sub> sequence where C<sup>b</sup> is a *t*-Bu-protected Cys and G<sub>o</sub> is glycolic acid.<sup>7,44,45</sup> Deprotection of the Cys thiol initiates a cascade reaction generating a diketopiperazine thioester that is competent in NCL reactions. Synthesis of peptides with the C<sup>b</sup>PG<sub>o</sub> linker is inefficient because of ester linkage instability. Therefore we reserve its use for longer peptides where it is necessary.

All  $\alpha$ S protein constructs tested contain Tyr to Phe mutations at positions 39, 125, 133, and 136 to generate an “all Phe” background construct,  $\alpha$ S<sup>F</sup>. As was the case for CaM<sup>F</sup>, these mutations allow us to more easily interpret Cnf fluorescence data. Combinations of methods 5a or 5b with 5c or 5d were used to synthesize double-labeled  $\alpha$ S. While all combinations were viable, for  $\alpha$ S<sup>F</sup><sub>4</sub>C<sub>9</sub>F\*<sub>94</sub>, our highest yields were obtained with a combination of methods 5b and 5c. Protein expression yields were typically 3 mg/L of culture after methoxylamine treatment. Ligations were typically carried out on the 0.1  $\mu$ mol scale at pH 7.2 with 1% thiophenol. After 24 hours, 60 – 70% yields of ligated product were typically seen, and the products purified by C4 reverse phase HPLC, giving isolated yields of about 0.4 mg for a given ligation. PAGE gel, spectroscopic, and MALDI MS analyses of the ligation confirm production of homogeneous, double-labeled  $\alpha$ S. Data for other reactions are shown in Supporting Information.

The methods shown in Figure 5 permit protein double-labeling with thioamides located at the N-terminus. To insert thioamides at the C-terminus, we made use of intein fusion constructs popularized by Muir and others.<sup>9</sup> These self-splicing domains have been engineered to stall at a thioester intermediate and can be intercepted by treatment with a small molecule thiol to give a thioester suitable for NCL. We have found that the expression of intein fusion constructs is compatible with Uaa mutagenesis. In fact, yields of intein fusions of  $\alpha$ S Cnf mutants (e.g.  $\alpha$ S<sup>F</sup><sub>1-122</sub>F\*<sub>94</sub>-Int) are higher (20 mg/L culture) than comparable expressions of the full-length protein using either T7 or PRK vectors previously

used to express  $\alpha$ S mutants (3 – 5 mg/L). The synthetic thiopeptide with N-terminal Cys ( $\alpha$ S<sup>F</sup><sub>123–140</sub>C<sub>123A'</sub><sub>124</sub>) was prepared by standard SPPS.

The thioester  $\alpha$ S<sup>F</sup><sub>1–122</sub>F\*<sub>39</sub>-SR was prepared from  $\alpha$ S<sup>F</sup><sub>1–122</sub>F\*<sub>39</sub>-Int using the thiol mercaptoethane sulfonate (MESNA). This was combined with  $\alpha$ S<sup>F</sup><sub>123–140</sub>C<sub>123A'</sub><sub>124</sub> in a buffer containing thiophenol. Although some C-terminal Asn cyclization (~ 25% based on crude MALDI MS estimates, see Supporting Information) of the intein thioester occurred in competition with the desired MESNA thiolysis, this was not a major problem since the resulting cyclic succinimide was not competent for ligation. Ligations were typically carried out on the 0.3  $\mu$ mol scale. After 16 hours, 50% yields were obtained, and the products purified by FPLC using an ion-exchange column, giving isolated yields of 0.6 – 0.7 mg for a given ligation reaction. Once again PAGE gels, spectroscopy, and MALDI MS were used to analyze the ligation reaction. These data are shown in Figure 8. The data confirm that we have again produced homogeneous, full-length, double-labeled  $\alpha$ S.

Using these two methods, we should be able to label a variety of proteins with thioamides at the N- or C-terminus, with complete freedom in placement of the donor fluorophore. Since we have previously shown that thioamides are compatible with the MeONH<sub>2</sub>•HCl conditions used to deprotect N-terminal thiazolidines, the methods used here should also be compatible with internal thioamide-labeling through three component ligations. In addition, we note that the use of C-terminal inteins may be generally useful to those employing Uaa mutagenesis. One major limitation of the method is that the inherent competition with release factors at the UAG stop codon can lead to protein truncation.<sup>46</sup> Purifying the full-length protein containing the Uaa can be difficult for sites near the C-terminus, where the truncated form is only slightly shorter. The His-tagged intein provides a convenient handle that will only be present in the untruncated protein. Since the intein can be cleaved by thiolysis or hydrolysis in a traceless manner, the Uaa-labeled fragment or protein can be obtained in pure form in an efficient manner.

In order to demonstrate the value of double-labeling a protein such as  $\alpha$ S, we carried out experiments in which intramolecular conformational changes were monitored.  $\alpha$ S aggregates by first forming soluble oligomers which eventually convert into larger, insoluble fibrils.<sup>13,14,47,48</sup> Both the oligomers and fibrils have been shown to be cytotoxic, and recent evidence has shown that extracellular  $\alpha$ S can enter neurons and catalyze misfolding of  $\alpha$ S within healthy neurons.<sup>49–52</sup> Therefore, there is substantial current interest in understanding the structural dynamics of the meta-stable  $\alpha$ S monomers and the rearrangements that occur on formation of oligomers and subsequent fibrillization. Several researchers have recently shown the value of being able to construct semi-synthetic  $\alpha$ S to study the effects of post-translational modification on its folding, our modification should provide a uniquely non-perturbing method for tracking its dynamics.<sup>7,39,53–55</sup>

We envision thioamide quenching as a method for site-selectively monitoring conformational changes during  $\alpha$ S refolding. While a number of researchers have studied  $\alpha$ S conformation using other techniques such as NMR, IR, and EPR, we find fluorescence to be particularly useful because the sensitivity of the technique permits one to use low concentrations of double-labeled  $\alpha$ S which should not undergo aggregation on the timescale of our experiments.<sup>47,48,52,56–60</sup> This strategy allows us to monitor long-range (25 – 30 Å) interactions that are difficult to observe with other techniques. Moreover, FRET is easily applied to interactions involving residues in dynamic regions of proteins, such as the termini of  $\alpha$ S.

We have previously shown that conformational dynamics during aggregation can be monitored by intramolecular quenching of Trp by a thioamide.<sup>7</sup> Here, we demonstrate that

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

The NCL procedures that we use to synthesize our thioamide-labeled proteins introduce a Cys into the  $\alpha$ S sequence at position 9 or 123. However, we have previously shown that the presence of non-native Cys in the sequence does not cause aberrant aggregation when incubated with  $\beta$ -mercaptoethanol (BME) to prevent disulfide formation.<sup>7</sup> While Cys can have an effect on Cnf emission, since we include Cys mutations in our oxoamide control experiments, we can account for these changes as well.<sup>34,61</sup>

We carried out refolding experiments with the constructs described above,  $\alpha$ S<sup>FF'4</sup>C<sub>9</sub>F\*<sub>94</sub> and  $\alpha$ S<sup>FF\*39</sup>C<sub>123</sub>A' <sub>124</sub>, as well as  $\alpha$ S<sup>FF'4</sup>C<sub>9</sub>F\*<sub>39</sub>. In the “native” state, when our constructs were incubated in buffer alone, we observed varying levels of quenching relative to the equivalent oxoprotein control. (Fig. 10) For F\*<sub>94</sub>/F' <sub>4</sub> and F\*<sub>94</sub>/A' <sub>124</sub> we observe no quenching. For F\*<sub>39</sub>/F' <sub>4</sub>, we observe some quenching, which we interpret as resulting from local compaction of the N-terminus. All of these observations are consistent with previous FRET experiments performed on  $\alpha$ S. The observed N-terminal compaction is similar to the roughly 30 Å separation observed by Rhoades and coworkers using Alexa Fluor 488 and Alexa Fluor 594 dyes in single molecule FRET experiments.<sup>62</sup> Ferreon *et al.* observed FRET efficiencies consistent with separations of ~ 70 Å in single molecule FRET experiments with  $\alpha$ S labeled at positions 7 and 84 with the same dye pair.<sup>63,64</sup> This is well outside the useful range of the Cnf/thioamide FRET pair, so it is not surprising that we see no quenching for F\*<sub>94</sub>/F' <sub>4</sub>. Our results for F\*<sub>39</sub>/F' <sub>4</sub> are consistent with experiments carried out by Lee, Gray, and coworkers using Trp and nitrotyrosine as a FRET pair.<sup>65,66</sup> Our F\*<sub>39</sub>/A' <sub>124</sub> data are consistent with reports by Nath *et al.* showing a ~ 45 Å interaction between probes at positions 33 and 130, which is beyond the range of the Cnf/thioamide pair.<sup>62</sup>

When 6 M urea was used as an additive, an increase in  $F_{\text{Thio}}/F_{\text{Oxo}}$  was observed for F\*<sub>39</sub>/F' <sub>4</sub>, consistent with greater unfolding of the protein. For the other two mutants, no change in  $F_{\text{Thio}}/F_{\text{Oxo}}$  was observed, indicating that these regions were indeed less compacted. Control experiments using Cnf itself indicate that although an inherent change in Cnf quantum yield plays a role in the observed change in fluorescence, there is nonetheless an observable expansion of  $\alpha$ S once one corrects for inherent effects on Cnf fluorescence. (See Supporting Information) In contrast, an increase in quenching (i.e. FRET) was observed as a function of TMAO concentration for all three mutants. The changes in  $E_Q$  as a function of TMAO concentration can be seen in Figure 10. This indicates increased compaction of both N- and C- termini, as well as folding of the N-terminal region on itself.

To assign distance constraints based on these measurements, we converted the measured  $E_Q$  values to distances using Förster theory. As with CaM, we computed a specific  $R_0$  value for each construct using the quantum yields of Cnf for the oxoamide control proteins  $\alpha$ S<sup>F</sup>C<sub>9</sub>F\*<sub>39</sub> ( $R_0 = 15.6$  Å in 0 M TMAO),  $\alpha$ S<sup>F</sup>C<sub>9</sub>F\*<sub>94</sub> ( $R_0 = 13.6$  in Å in 0 M TMAO), or  $\alpha$ S<sup>F</sup>F\*<sub>39</sub>C<sub>123</sub> ( $R_0 = 13.5$  Å in 0 M TMAO). Förster radii for each mutant in other TMAO concentrations are given in Supporting Information. We use these rather than simply using  $R_0$  values calculated from free Cnf because the variation among the different mutants demonstrates that the protein context influences quantum yield. The  $R_{\text{FRET}}$  values that result from our calculations indicate that Cnf and the thioamide come into very close proximity in TMAO. While these results are consistent with the single molecule studies of Deniz and coworkers, it is surprising that our  $R_{\text{FRET}}$  values are so small.<sup>64</sup> This may be the result of a dominant contribution from very close contact. Ferreon *et al.* observe subpopulations with

near unity FRET efficiencies, but the Alexa Fluor FRET pair that they use has high FRET efficiency for any separation less than 30 Å.<sup>63,67</sup> Our probes have a shorter effective range (8 – 31 Å) than theirs, and therefore provide a more precise report on the close contact induced by TMAO. We will continue to study this phenomenon, not only with thioamides, but also with other short range FRET pairs such as Cnf and Trp.<sup>68,69</sup> Finally, it may be more appropriate to use polymer models rather than simple Förster theory to interpret our FRET data.<sup>70,71</sup> These models have clear value when interpreting single molecule FRET data, which may have broad and asymmetric FRET efficiency distributions. Although the value is less clear for ensemble measurements, we will explore the interpretation of our FRET data using models such as those used by Lee, Gray, Rhoades, and Plaxco.<sup>62,65,66,72,73</sup>

## CONCLUSIONS

The application of our recently developed Cnf/thioamide FRET probe pair to studies of the folding of full-sized proteins requires the ability to synthesize those labeled proteins in an efficient manner. Although we could, in theory, construct Cnf- and thioamide-labeled proteins purely through ligation, we have developed a far more expedient approach by melding the methods of unnatural amino acid mutagenesis and protein ligation. Intermolecular studies simply require that Cnf be incorporated into proteins expressed in *E. coli* through established unnatural amino acid mutagenesis techniques. Intramolecular studies require more complex semi-synthesis to obtain doubly labeled proteins, wherein a Cnf-labeled protein fragment can be expressed in *E. coli* and then ligated to a synthetic thiopeptide. This combination of methods minimizes SPPS of natural amino acid segments. The same methods can be used to synthesize proteins labeled with other donor fluorophores such as coumarin, or any fluorophore conjugated to a bioorthogonal handle, such as *p*-azidophenylalanine.<sup>74,75</sup> In fact, we have identified a synthetase for the incorporation of the fluorescent amino acid acridonylalanine (Acd), which we have previously shown to be quenched by thioamides.<sup>6,76</sup> We are currently working to optimize yields of Acd/thioamide labeled proteins using methods analogous to those described here.

The doubly labeled proteins produced using these methods can be applied to monitor protein folding and dynamics. For example, here we have studied the structure of a CaM/peptide complex and determined the average separation of residues in the disordered protein  $\alpha$ S in various buffer conditions. The benchmarking results using CaM help to validate our method and provide a foundation for the application of thioamide FRET to other systems. These experiments show clearly that meaningful interpretations of the interchromophore distances require determining changes in the donor quantum yield. This is easily accomplished using the oxoamide control experiments that we always include in order to determine quenching efficiency. Moreover, the dependence of quenching efficiency on Cnf quantum yield supports a FRET mechanism for quenching as opposed to Dexter or PET mechanisms, which were difficult to discount based on previous data.<sup>4,77</sup> The confirmation of a FRET mechanism for Cnf/thioamide interactions allows one to confidently interpret thioamide quenching using Förster's equations for future applications. Nonetheless, it is worth noting that there was an imperfect match of observed distances ( $R_{\text{FRET}}$ ) and predicted distances ( $R_{\text{NMR}}$ ) that was not resolved by considering the orientation parameter  $\kappa^2$ . This may be because simply using the 20 structures developed in solving the NMR structure does not fully account for intermediate geometries between these valleys on the potential energy surface. Calculating  $R_{\text{FRET}}$  by sampling  $\kappa^2$  over extended molecular dynamics trajectories may give improved agreement with the FRET data. Additionally, fluorescence lifetime measurements (which are concentration-independent) would remove the influence of small deviations in concentration on our determination of  $E_Q$ . We will explore both of these aspects in subsequent work.



The  $\alpha$ S experiments show how one might use our method in an intramolecular sense. The addition of TMAO results in compaction of monomeric  $\alpha$ S, consistent with previous independent observations by Deniz as well as small angle X-ray scattering experiments by Fink.<sup>64,78</sup> The short effective range of our probes allows us to obtain more precise information about close contact of regions of the protein than similar experiments with larger fluorophores. Moreover, the small size of our chromophores leaves less uncertainty in their location relative to the protein structure. Thus, the thioamide probes provide new short-range information on TMAO-induced conformations of  $\alpha$ S, complementing data obtained in single molecule experiments. We continue to apply fluorophore/thioamide pairs to the study of folding transitions in  $\alpha$ S monomers as well as to monitoring conformational changes during aggregation.

Of course, our labeling strategy is not restricted to CaM or  $\alpha$ S. Any protein which can be expressed at high levels in *E. coli* and reversibly unfolded should be amenable to double-labeling using a combination of Uaa mutagenesis and NCL. This strategy should be particularly appealing when one of the modifications cannot be incorporated ribosomally (i.e. backbone modifications such as olefin isosteres). For larger dyes that are attached post-translationally, Uaa mutagenesis at two sites, or used in combination with Cys modification is probably more efficient.<sup>79–82</sup> However, even in those cases, it can be difficult to separate labeled from unlabeled protein, whereas the ligated products of our reactions differ substantially in size and can easily be separated by HPLC or FPLC. The major current limitations of our labeling methods are that Uaa mutagenesis expression yields can be low (3 – 5 mg/L), and if these are coupled to inconsistent NCL yields, it can be difficult to obtain sufficient quantities of protein for biophysical experiments. We will continue methodological studies to improve both aspects to make these methods as accessible as possible to the general biochemical community.

## EXPERIMENTAL PROCEDURES

### General Information

Fmoc-L-4-cyanophenylalanine (Fmoc-Cnf-OH) was purchased from Peptech (Burlington, MA). Boc-L-thionophenylalanine-1-(6-nitro)benzotriazolide, Fmoc-Gln(Trt)-OH, Fmoc-Asn(Trt)-OH were purchased from Bachem (Torrance, CA) or EMD Chemicals (Philadelphia, PA). Benzotriazol-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP), Fmoc-Ala-OH, Fmoc-Leu-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Thr(tBu)-OH, Fmoc-Met-OH, Fmoc-Gly-OH, Fmoc-Val-OH, Fmoc-Phe-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Cys(Trt)-OH, 2-chlorotrityl chloride resin, and Rink amide resin were purchased from Novabiochem (San Diego, CA). Piperidine and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from American Bioanalytical (Natick, MA). Sigma-cote, *N,N*-diisopropyl ethylamine (DIPEA), thiophenol, tri-fluoroacetic acid (TFA), tris(2-carboxyethyl)phosphine hydrochloride (TCEP), were purchased from Sigma-Aldrich (St. Louis, MO). Ni-NTA resin was from Qiagen (Valencia, CA). *E. coli* BL21(DE3) cells were purchased from Stratagene (La Jolla, CA). Sequencing-grade trypsin was purchased from Promega (Madison, WI). Restriction Grade Factor Xa protease was purchased from Novagen (San Diego, CA). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA).

Milli-Q filtered (18 M $\Omega$ ) water was used for all solutions (Millipore; Billerica, MA). Matrix-assisted laser desorption ionization (MALDI) mass spectra were collected with a Bruker Ultraflex III MALDI-TOF-TOF mass spectrometer (Billerica, MA). UV/Vis absorbance spectra were obtained with a Hewlett-Packard 8452A diode array spectrophotometer (Agilent Technologies, Santa Clara, CA). Fluorescence spectra were collected with a Varian

Cary Eclipse fluorescence spectrophotometer fitted with a Peltier multicell holder (currently Agilent Technologies).

## Peptide Synthesis

All peptides were synthesized using a manual, Fmoc-based solid-phase procedure described previously.<sup>4,7</sup> Thioamide monomers were synthesized as previously described.<sup>4,83</sup>

## CaM Expression and Purification

Each plasmid containing a CaM mutant (pCaM-TAG<sub>13</sub>, pCaM-TAG<sub>17</sub>, pCaM-TAG<sub>93</sub>, or pCaM-TAG<sub>100</sub>, see Supporting information for plasmid construction) was transformed into pDULE2-Cnf-containing *E. coli* BL21(DE3) cells. These cells were cultured and made competent through the Hanahan method.<sup>84</sup> First, the pDULE2-Cnf plasmid was transformed into BL21(DE3) cells. Then, a colony was picked based on streptomycin (Strep) resistance to grow cultures of competent cells. Once pCaM was transformed into these cells, selection was based on both ampicillin (Amp) and Strep resistance. Single colonies were used to inoculate 5 mL of LB media containing Amp (100 µg/mL) and Strep (100 µg/mL). This primary 5 mL culture was incubated at 37 °C with shaking at 250 RPM overnight, and then added to 1 L of M9 minimal media containing Amp and Strep at the same concentrations. M9 minimal media was prepared by adding the following autoclaved solutions to 1 L of auto-claved water containing 6 g Na<sub>2</sub>HPO<sub>4</sub>, 3 g KH<sub>2</sub>PO<sub>4</sub>, 0.5 g NaCl and 1 g NH<sub>4</sub>Cl: 1 mL of 2 M MgSO<sub>4</sub>, 1 mL of 15 mg/mL FeCl<sub>2</sub> (in 1.0 M HCl), 1 mL of 15 mg/mL ZnCl<sub>2</sub> (in acidified H<sub>2</sub>O), 2 mL of 10% Bacto™ Yeast Extract, 20 mL of 10% glycerol (v/v) and 1 µL of 1 M CaCl<sub>2</sub>. The 1 L secondary culture was incubated at 37 °C with shaking at 250 RPM until the OD<sub>600</sub> reached 0.7 – 0.9 AU. Protein expression was induced with isopropyl D-galactoside (IPTG) and 95 mg (0.5 mM) Cnf, and the culture was incubated at 37 °C for an additional 12 h. The cells were recovered from media at 6000 × g for 10 minutes. The resulting pellet was suspended in 15 mL of CaM resuspension buffer (50 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS), 100 mM KCl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.5). Cell were lysed by sonication, alternating between one minute of sonication. CaCl<sub>2</sub> was added to the sonicated lysate to a final concentration of 20 mM prior to centrifugation for 10 minutes at 6,000 × g, 4 °C. Following centrifugation, the supernatant was collected for purification. CaM was purified from the cleared cell lysate using a phenyl-sepharose (PS) CL-4B column with EDTA as the eluent. A column with a total resin bed volume of approximately 8 mL was first equilibrated with four 10 mL washes of PS Buffer A (50 mM Tris base, 1 mM CaCl<sub>2</sub>). Then the cleared cell lysate was added to the column and collected as it ran through to save any protein that did not bind to the column. Next, the column was washed four times with 10 mL PS Buffer A. The column was then washed with four 10 mL portions of high-salt PS Buffer B (50 mM Tris base, 0.5 M NaCl, 0.1 mM CaCl<sub>2</sub>). Finally, an additional 10 mL wash of PS Buffer A was performed to restore the column to a low-salt condition. CaM was eluted with 18 mL of PS Buffer C (10 mM Tris base, 10 mM EDTA, pH 7.5). This elution solution was then re-saturated with CaCl<sub>2</sub> to a concentration of 20 mM. This elution solution was repurified on a second PS column using the same procedure to reduce truncation product. For the second PS column, rather than adding 18 mL of PS Buffer C directly to the column, 1 mL elutions were collected and analyzed by SDS-PAGE gel. Collection stopped when the protein was no longer visible by gel. The elution fractions were dialyzed against CaM Buffer (140 mM KCl, 15 mM HEPES, 6 mM CaCl<sub>2</sub>, pH 6.7) and stored at –20 °C. Proteins were analyzed by SPS PAGE gel and MALDI MS for purity.

## CaM/pOCNC Binding Assays

After synthesis, HPLC purification, and lyophilization, the various pOCNC peptides were redissolved in CaM buffer (140 mM KCl, 10 mM HEPES, 6 mM CaCl<sub>2</sub>, pH 6.7). The

concentrations of the peptides were determined by UV/Vis spectroscopy, using an extinction coefficient of  $10,270 \text{ M}^{-1} \text{ cm}^{-1}$  at 273 nm for the thioamide-containing peptides and  $600 \text{ M}^{-1} \text{ cm}^{-1}$  at 257 nm for the oxoamide-containing peptide. The concentration of the CaM proteins was determined through the bicinchoninic acid (BCA) assay. For each peptide/protein combination, six solutions were prepared that contained  $10 \mu\text{M}$  CaM protein and 0 to  $15 \mu\text{M}$  pOCNC peptide. In addition, a solution of  $10 \mu\text{M}$  peptide in the absence of protein was made to confirm that no background fluorescence resulted from the peptide. The solutions were prepared three times in order to obtain data in triplicate.  $E_Q$  was computed from the raw fluorescence data for CaM with thioamide-labeled pOCNC ( $F_{\text{Thio}}$ ) or pOCNC ( $F_{\text{Oxo}}$ ).

### C-Terminal Fragment Expression (Factor Xa)

The plasmid pET16b- $\text{H}_{\text{Tag}}\text{-}\alpha\text{S}_{\Delta 1-8}\text{C}_9\text{F}^*_{39}$  (see Supporting Information for plasmid construction) was transformed into *E. coli* BL21(DE3) cells containing pDULE2-Cnf (see above) and colonies were selected on the basis of Amp and Strep resistance. Following growth of a primary culture in LB media, 1L of M9 minimal media containing glycerol as a carbon source (0.25% v/v) was inoculated and allowed to grow at  $37^\circ\text{C}$  with shaking at 250 RPM. When the  $\text{OD}_{600}$  reached 0.8, Cnf was added (150 mg), and the culture was allowed to incubate overnight at  $37^\circ\text{C}$ . The boiled and cleared lysate was prepared as described and incubated with 2 mL Ni-NTA resin for 1 hour at room temperature.<sup>7</sup> The resin was washed with 20 mL of 50 mM Tris, 150 mM NaCl, pH 8. The protein was eluted in eight 1.5 mL portions of 50 mM Tris, 150 mM NaCl, 250 mM imidazole, pH 8. Elution fractions were analyzed by SDS-PAGE gel, combined, and dialyzed against Factor Xa cleavage buffer (50 mM Tris, 150 mM NaCl, 5 mM  $\text{CaCl}_2$ , pH 8.0) overnight. Following dialysis, His-tag cleavage was achieved by overnight incubation with Factor Xa protease (10 units per 1 mg of  $\text{H}_{\text{Tag}}\text{-}\alpha\text{S}_{\Delta 1-8}\text{C}_9\text{F}^*_{39}$ ). Factor Xa was inactivated by boiling the reaction for 20 minutes and the precipitated enzyme was removed by centrifugation for 20 minutes at  $30,000 \times g$ ,  $4^\circ\text{C}$ . The supernatant was then subjected to incubation with 1 mL of Ni-NTA resin for 1 hr at room temperature to capture any remaining  $\text{H}_{\text{Tag}}\text{-}\alpha\text{S}_{9-140}\text{C}_9\text{F}^*_{39}$ . The flow-through containing  $\alpha\text{S}_{\Delta 1-8}\text{C}_9\text{F}^*_{39}$  was collected, analyzed by SDS-PAGE gel and MALDI MS, dialyzed overnight against water, and stored at  $-80^\circ\text{C}$ .

### C-Terminal Fragment Expression (Met Aminopeptidase)

The plasmids pRK- $\alpha\text{S}^{\text{F}}_{\Delta 2-8}\text{C}_9\text{F}^*_{39}$  and pRK- $\alpha\text{S}^{\text{F}}_{\Delta 2-8}\text{C}_9\text{F}^*_{94}$  (see Supporting Information for plasmid construction) were transformed into competent *E. coli* BL21(DE3) cells with pDULE2-Cnf (see above). Transformed cells were selected on the basis of Amp and Strep resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp (100  $\mu\text{g}/\text{mL}$ ) and Strep (100  $\mu\text{g}/\text{mL}$ ). The primary 5 mL culture was incubated at  $37^\circ\text{C}$  with shaking at 250 RPM for 5 hours and then used to inoculate 1 L of LB media. When the  $\text{OD}_{600}$  of the secondary culture reached 0.8, the cells were isolated by centrifugation at  $5000 \times g$  for 15 minutes and gently re-suspended in 1 L of M9 minimal media containing Cnf (100 mg). M9 minimal media was prepared as described above. The cells were harvested at  $5000 \times g$  for 15 min and the resulting pellet was resuspended in 20 mM Tris pH 8. Following sonication, the cell lysate was boiled for 20 min prior to centrifugation for 20 min at  $30,000 \times g$ ,  $4^\circ\text{C}$ . The cleared lysate was loaded onto a Superdex 200 column (25 cm) connected to a BioCad Sprint (FPLC) system and eluted with 20 mM Tris, pH 8. FPLC fractions from size-exclusion chromatography were analyzed by SDS PAGE. The fractions containing  $\alpha\text{S}$  were then loaded onto a HighTrap<sup>TM</sup> Q HP column and eluted over a 100 min NaCl gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8). FPLC fractions from ion-exchange chromatography were dialyzed against 20 mM Tris, 100 mM NaCl, pH 7.4 and analyzed by SDS PAGE gel, MALDI MS, UV/Vis, and fluorescence spectroscopy. Following dialysis, guanidinium hydrochloride and methoxylamine hydrochloride were added to a final

concentration of 6 M and 400 mM, respectively. The pH was adjusted to 4.0 with NaOH and the solution was stirred overnight at 4 °C. Removal of N-terminal cysteine adducts was monitored by MALDI MS. Following deprotection, the solution was dialyzed against water and dried in a vacuum centrifuge.

### Thioester Synthesis (PyBOP)

Peptide synthesis and selective cleavage from Cl-Trt resin was performed as previously described.<sup>7</sup> Cleaved N-terminal peptide was dissolved in THF, and 2-mercaptopropionate (3 equiv) was added to the solution. After stirring for 5 min, DIPEA (6 equiv) and a solution of PyBOP pre-dissolved in warm THF (3 equiv) were added to the reaction mixture. The solution was allowed to stir for 30 min at room temperature, at which point the solvent was removed by rotary evaporation. Sidechain protecting groups were removed by incubating the peptide for 10 min with a cleavage cocktail containing TFA, TIPS, thioanisole, and 1,2-ethanedithiol (47:1:1:1). Following concentration of the cleavage solution by rotary evaporation, the peptide was precipitated with ether and isolated by centrifugation. The peptide thioester was purified by HPLC and dried in a vacuum centrifuge.

### Thioester Synthesis (C<sup>b</sup>PG<sub>0</sub>)

Ac-MDV'<sub>3</sub>FMKGLC<sup>b</sup>PG<sub>0</sub> (αS<sub>1-8</sub>V'<sub>3</sub>-C<sup>b</sup>PG<sub>0</sub>) was synthesized essentially as previously described, with the following modifications.<sup>7</sup> Bromoacetic acid (0.0695 g, 500 μmol, 5 equiv) was pre-activated with diisopropyl carbodiimide (78 μL, 500 μmol, 5.0 equiv) for 30 min in dry DMF (6 mL). Also, the peptide (αS<sub>1-8</sub>V'<sub>3</sub>-C<sup>b</sup>PG<sub>0</sub>) was cleaved from the resin with the treatment of CH<sub>2</sub>Cl<sub>2</sub>/TFA/thioanisole (38:1:1 v/v, 5 mL) on a rotisserie for 1 h. The resulting solution was concentrated by rotatory evaporation, precipitated in cold diethyl ether (20 mL), centri-fuged, and decanted ether. The precipitated peptide was dissolved in acetonitrile/water (50:50 v/v) for HPLC purification.

### Ligation to form αS<sup>F</sup>F'<sub>4</sub>C<sub>9</sub>F\*<sub>39</sub>, αS<sup>F</sup>V'<sub>3</sub>C<sub>9</sub>F\*<sub>39</sub>, and αS<sup>F</sup>F'<sub>4</sub>C<sub>9</sub>F\*<sub>94</sub>

The expressed C-terminal protein fragment (1 equiv, 0.1 μmol) was dissolved in 100 μL of freshly prepared ligation buffer (6 M GdnHCl, 200 mM Na<sub>2</sub>HPO<sub>4</sub>, 20 mM TCEP, 1% v/v thiophenol, pH 7.2 for PyBOP-based thioesters or pH 8.0 for C<sup>b</sup>PG<sub>0</sub> thioesters). The dissolved protein fragment was transferred to a microcentrifuge tube containing 1.2 equiv of the dried N-terminal peptide thioester, purged with argon, and allowed to incubate overnight at 37 °C with shaking at 1000 RPM. The ligation solution was diluted to 3 mL and dialyzed against water prior to HPLC purification (see Supporting Information). Following HPLC purification, the product fractions were combined, concentrated, and exchanged into 20 mM Tris, 100 mM NaCl, pH 7.4 using an Amicon (Millipore) Ultra 0.5 mL 3 kDa spin column.

### Expression and purification of αS<sub>1-122</sub>F\*<sub>39</sub>-Int

A plasmid encoding αS<sub>1-122</sub>F\*<sub>39</sub>-Int-H<sub>T</sub>ag was transformed into *E. coli* BL21(DE3) cells containing the pDULE2-Cnf gene (prepared and made competent using the Hanahan method).<sup>84</sup> A starter culture of 4 mL LB media was inoculated with a single colony and was grown at 37 °C in the presence of Amp (100 mg/L) and Strep (100 mg/L) for approximately 4 h. A secondary culture of 1 L M9 media was inoculated with 4 mL of the starter culture and grown at 37 °C with the same concentrations of Amp and Strep. M9 minimal media was prepared as described above. The culture was induced by the addition of Cnf (150 mg) followed by 0.5 mM of IPTG when it reached an OD<sub>600</sub> of 0.8 and allowed to grow at 37 °C overnight. Cells were harvested at 5,000 RPM using a GS3 rotor and Sorvall RC-5 centrifuge for 20 min at 4 °C. After discarding the supernatant, the pellet was resuspended in 30 mL of resuspension, pH 8.3, protease inhibitor cocktail, 1 mM phenylmethanesulfonyl fluoride (PMSF), and 10 units/mL DNase I–Grade II. The buffer: 40 mM Tris, 5 mM EDTA

cells were lysed by sonication, and the lysate was centrifuged at 13,900 RPM for 30 min at 4 °C. The supernatant was then incubated with Ni-NTA resin for 1 h on ice. This slurry was loaded into a column and the liquid allowed to flow through. The resin was washed extensively, first with 15 mL of Wash Buffer A (50 mM HEPES, pH 7.0), then twice with 10 mL of Wash Buffer B (50 mM HEPES, 10 mM imidazole, pH 7.0). The protein was eluted with five 2 mL portions of elution buffer (50 mM HEPES, 300 mM imidazole, pH 7.0). The eluted fractions were combined and incubated with 400 mM sodium 2-mercaptoethanesulfonate (MESNA) for thiolysis for 48 h, stirring at 4 °C. The protein thioester was dialyzed against 20 mM Tris, pH 8.0 overnight, and incubated with Ni-NTA resin for 1 h on ice. The flow-through was collected, dialyzed against 20 mM sodium citrate, pH 5.0, purified over HiTrap SP column using a 100 min NaCl gradient (0 to 0.5 M NaCl in sodium citrate, pH 5). The product fractions were pooled and dialyzed against Milli-Q water overnight, dried in a vacuum centrifuge, and stored at – 80 °C until further use.

### Ligation to Form $\alpha S^{F^*_{94}C123A'124}$

Protein thioester  $\alpha S_{1-122}F^*_{39}$  (1.5 equiv, 0.45  $\mu$ mol, 1.5 mM) and C-terminal protein fragment  $\alpha S_{123-140}C123A'_{124}$  (1 equiv, 0.3  $\mu$ mol, 1 mM) were dissolved in 300  $\mu$ L of a freshly made, degassed ligation buffer (6 M Gdn•HCl, 200 mM  $Na_2HPO_4$ , 20 mM TCEP, and 1% v/v thiophenol, pH 7.0). The reaction solution was placed in an incubator at 37 °C, shaking at 1000 RPM. The ligation progress was monitored by MALDI MS, and allowed to react until all the protein thioester was either consumed or hydro-lyzed. Upon completion, the ligation solution was brought up to 1.5 mL with water and dialyzed against 20 mM Tris, pH 8.0 for 4 h. Then, the ligated product,  $\alpha S^{F^*_{39}C123A'_{124}}$ , was purified over a HiTrap™ Q HP column using a 100 min NaCl gradient (0 to 0.5 M NaCl in 20 mM Tris, pH 8). The product fractions were pooled, concentrated, and the buffer was ex- changed for 20 mM Tris, 100 mM NaCl, pH 7.4 using an Amicon (Millipore) Ultra 0.5 mL 10 kDa spin column.

### Trypsin Digest

Proteins (10 – 50  $\mu$ g) were incubated with 5 – 10  $\mu$ L aliquots of sequencing-grade modified trypsin (0.1 mg/mL) in 20 mM Tris, 100 mM NaCl, pH 7.5 at 37 °C for 4 – 8 hours. An aliquot (1.0  $\mu$ L) of the digestion reaction was taken and analyzed by MALDI MS (Figure 6, Figure 8, and Supporting Information).

### Refolding Assays

Prior to performing the experiment, the concentrations of all double-labeled proteins were adjusted using the BCA assay to match their respective oxoamide (donor only) control proteins. Tris buffers (20 mM Tris, 100 mM NaCl, 1 mM  $\beta$ ME, pH 7.5) containing trimethylamine oxide (TMAO) were prepared such that upon addition of protein (to a final concentration of 1  $\mu$ M), the final TMAO concentrations were 0 M, 1 M, 2 M, 3 M, and 4 M. Buffers containing urea were similarly prepared such that upon addition of protein, the final urea concentration was 6 M. Fluorescence spectra were measured immediately following sample dilution in TMAO or urea buffer (to a final volume of 130  $\mu$ L). All samples were prepared and measured in triplicate.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

This work was supported by funding from the University of Pennsylvania, including a grant from the Institute on Aging, the Searle Scholars Program (10-SSP-214 to EJP), the National Science Foundation (NSF CHE-1150351)

and the National Institutes of Health (NIH NS081033 to EJP). Instruments supported by the NSF and NIH include: HRMS (NIH RR-023444) and MALDI-MS (NSF MRI-0820996). We thank Prof. Ryan Mehl for the gift of the pDULE2-Cnf plasmid. We also thank Dr. Christopher Lanci of the UPenn Biological Chemistry Resource Center for assistance with circular dichroism spectroscopy (supported by UPenn Laboratory for Research on the Structure of Matter with NSF MRSEC DMR 0520020, 1120901). RFW thanks the NIH for funding through the Chemistry-Biology Interface Training Program (T32 GM07133). CMF thanks the Roy and Diana Vagelos Scholars Program in the Molecular Life Sciences for support.

## ABBREVIATIONS

<b>FRET</b>	Förster resonant energy transfer Cnf or F*, <i>p</i> -cyanophenylalanine
<b>PET</b>	photoinduced electron transfer
<b>SPPS</b>	solid phase peptide synthesis
<b>NCL</b>	native chemical ligation
<b>Uaa</b>	unnatural amino acid
<b>RS</b>	aminoacyl tRNA synthetase
<b>CaM</b>	calmodulin
<b><math>\alpha</math>S</b>	$\alpha$ -synuclein
<b>OCNC</b>	olfactory cyclic nucleotide-gated channel
<b>IPTG</b>	isopropyl $\beta$ -D-thiogalactopyranoside
<b>PAGE</b>	polyacrylamide gel electrophoresis
<b>MPAA</b>	4-mercaptophenylacetic acid
<b>Fmoc</b>	9-fluorenylmethoxycarbonyl
<b>Cl-Trt</b>	2-chlorotrityl
<b>TFA</b>	trifluoroacetic acid
<b>PyBOP</b>	benzotri-azol-1-yl-oxytri-pyrrolidinophosphonium hexafluoro-phosphate
<b>C<sup>b</sup></b>	<i>S</i> - <i>tert</i> -Butyl-L-cysteine
<b>G<sub>o</sub></b>	glycolic acid
<b>MESNA</b>	mercap-toethane sulfonate
<b>TMAO</b>	trimethylamine oxide
<b>BME</b>	$\beta$ -mercaptoethanol
<b>Acđ</b>	acridonylalanine thioamide residues are represented by the one or three letter code of the equivalent oxo-amide amino acids with a prime symbol (e.g. L' or Leu' represent thioleucine)

## REFERENCES

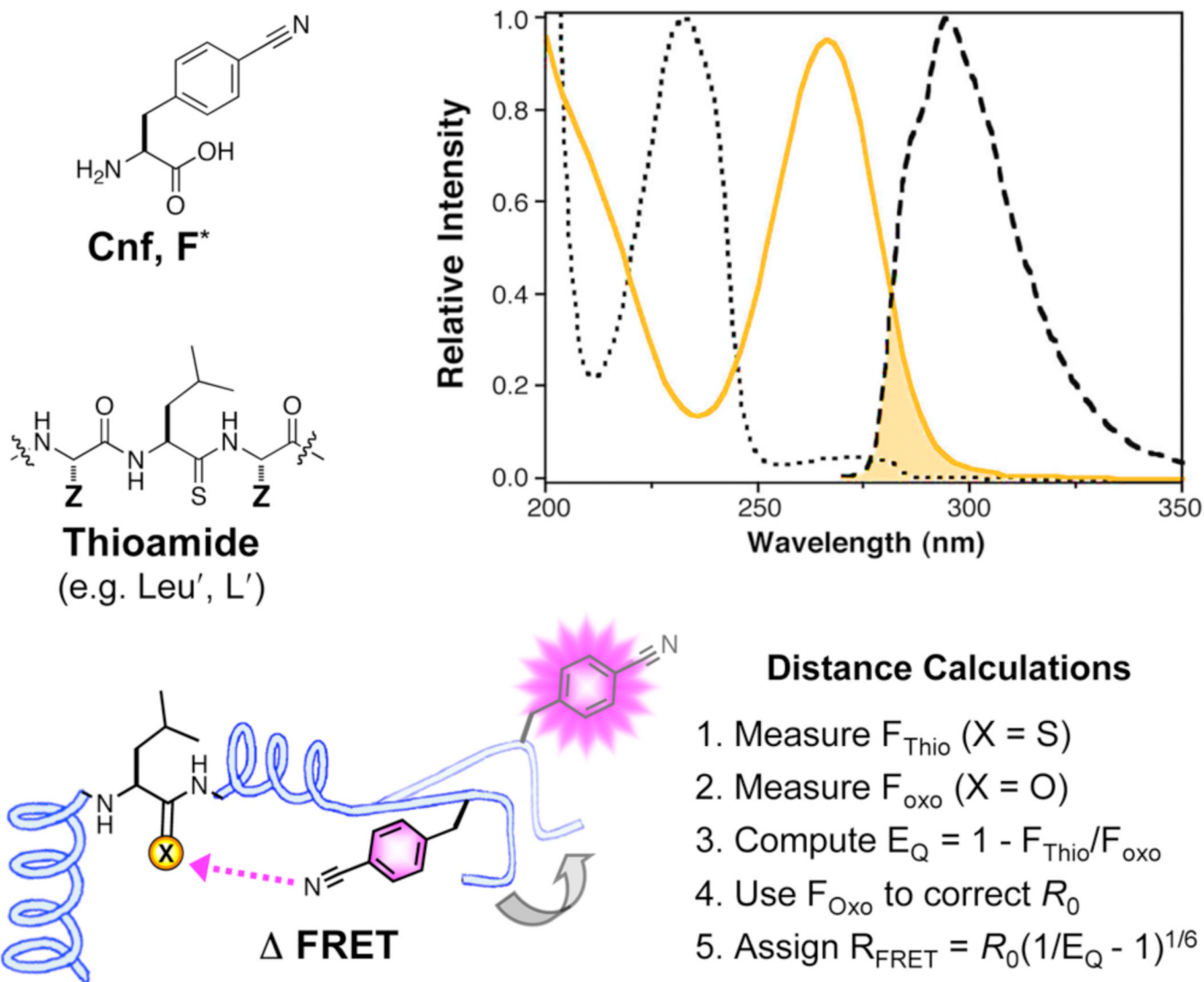
1. Forster T. Discuss. Faraday Soc. 1959; No. 27:7–17.
2. Royer CA. Chem. Rev. 2006; 106:1769–1784. [PubMed: 16683754]
3. Piston DW, Kremers G-J. Trends Biochem. Sci. 2007; 32:407–414. [PubMed: 17764955]
4. Goldberg JM, Batjargal S, Petersson EJ. J. Am. Chem. Soc. 2010; 132:14718–14720. [PubMed: 20886849]
5. Goldberg JM, Wissner RF, Klein AM, Petersson EJ. Chem. Commun. 2012; 48:1550–1552.
6. Goldberg JM, Speight LC, Fegley MW, Petersson EJ. J. Am. Chem. Soc. 2012; 134:6088–6091. [PubMed: 22471784]

7. Batjargal S, Wang YJ, Goldberg JM, Wissner RF, Petersson EJ. *J. Am. Chem. Soc.* 2012; 134:9172–9182. [PubMed: 22468862]
8. Alewood P, Engelhard M, Kent SBH. *J. Peptide Sci.* 2010; 16:513–513. [PubMed: 20862718]
9. Muralidharan V, Muir TW. *Nature Methods.* 2006; 3:429–438. [PubMed: 16721376]
10. Wang L, Schultz PG. *Angew. Chem. Int. Ed.* 2005; 44:34–66.
11. Wang L, Brock A, Herberich B, Schultz PG. *Science.* 2001; 292:498–500. [PubMed: 11313494]
12. Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R, Goedert M. *Nature.* 1997; 388:839–840. [PubMed: 9278044]
13. Wood SJ, Wypych J, Steavenson S, Louis JC, Citron M, Biere AL. *J. Biol. Chem.* 1999; 274:19509–19512. [PubMed: 10391881]
14. Auluck PK, Caraveo G, Lindquist S. *Ann. Rev. Cell Dev. Biol.* 2010; 26:211–233. [PubMed: 20500090]
15. Nguyen DP, Elliott T, Holt M, Muir TW, Chin JW. *J. Am. Chem. Soc.* 2011; 133:11418–11421. [PubMed: 21736333]
16. Li X, Fekner T, Ottesen JJ, Chan MK. *Angew. Chem. Int. Ed.* 2009; 48:9184–9187.
17. Choudhary A, Raines RT. *ChemBioChem.* 2012; 12:1801–1807. [PubMed: 21751326]
18. Valiyaveetil FI, Sekedat M, MacKinnon R, Muir TW. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:17045–17049. [PubMed: 15563591]
19. Fu YW, Bieschke J, Kelly JW. *J. Am. Chem. Soc.* 2005; 127:15366–15367. [PubMed: 16262389]
20. Fu YW, Gao JM, Bieschke J, Dendle MA, Kelly JW. *J. Am. Chem. Soc.* 2006; 128:15948–15949. [PubMed: 17165703]
21. Autric F, Ferraz C, Kilhoffer MC, Cavadore JC, Demaille JG. *Biochim. Biophys. Acta.* 1980; 631:139–147. [PubMed: 7397242]
22. Brzeska H, Venyaminov SV, Grabarek Z, Drabikowski W. *FEBS Lett.* 1983; 153:169–173. [PubMed: 6825857]
23. Chabbert M, Kilhoffer MC, Watterson DM, Haiech J, Lami H. *Biochemistry.* 1989; 28:6093–6098. [PubMed: 2775754]
24. Chapman ER, Alexander K, Vorherr T, Carafoli E, Storm DR. *Biochemistry.* 1992; 31:12819–12825. [PubMed: 1463753]
25. Gilli R, Lafitte D, Lopez C, Kilhoffer MC, Makarov A, Briand C, Haiech J. *Biochemistry.* 1998; 37:5450–5456. [PubMed: 9548926]
26. Kilhoffer MC, Demaille JG, Gerard D. *Biochemistry.* 1981; 20:4407–4414. [PubMed: 7284330]
27. Moorthy AK, Gopal B, Satish PR, Bhattacharya S, Bhattacharya A, Murthy MRN, Surolia A. *FEBS Lett.* 1999; 461:19–24. [PubMed: 10561489]
28. Marlow MS, Dogan J, Frederick KK, Valentine KG, Wand AJ. *Nat. Chem. Biol.* 2010; 6:352–358. [PubMed: 20383153]
29. Vetter SW, Leclerc E. *Eur. J. Biochem.* 2003; 270:404–414. [PubMed: 12542690]
30. Liu MY, Chen TY, Ahamed B, Li J, Yau KW. *Science.* 1994; 266:1348–1354. [PubMed: 7526466]
31. Contessa GM, Orsale M, Melino S, Torre V, Paci M, Desideri A, Cicero DO. *J. Biomol. NMR.* 2005; 31:185–199. [PubMed: 15803393]
32. Serrano AL, Troxler T, Tucker MJ, Gai F. *Chem. Phys. Lett.* 2010; 487:303–306. [PubMed: 20419080]
33. Du H, Fuh RCA, Li JZ, Corkan LA, Lindsey JS. *Photochem. Photobiol.* 1998; 68:141–142.
34. Taskent-Sezgin H, Chung J, Patsalo V, Miyake-Stoner SJ, Miller AM, Brewer SH, Mehl RA, Green DF, Raleigh DP, Carrico I. *Biochemistry.* 2009; 48:9040–9046. [PubMed: 19658436]
35. Chen Y, Liu B, Yu HT, Barkley MD. *J. Am. Chem. Soc.* 1996; 118:9271–9278.
36. Lakowicz, JR. *Principles of fluorescence spectroscopy.* Third ed.. New York, NY: Springer; 2006.
37. Chen H, Ahsan SS, Santiago-Berrios MEB, Abrunifa HD, Webb WW. *J. Am. Chem. Soc.* 2010; 132:7244–7245. [PubMed: 20446733]
38. Muir TW. *Ann. Rev. Biochem.* 2003; 72:249–289. [PubMed: 12626339]

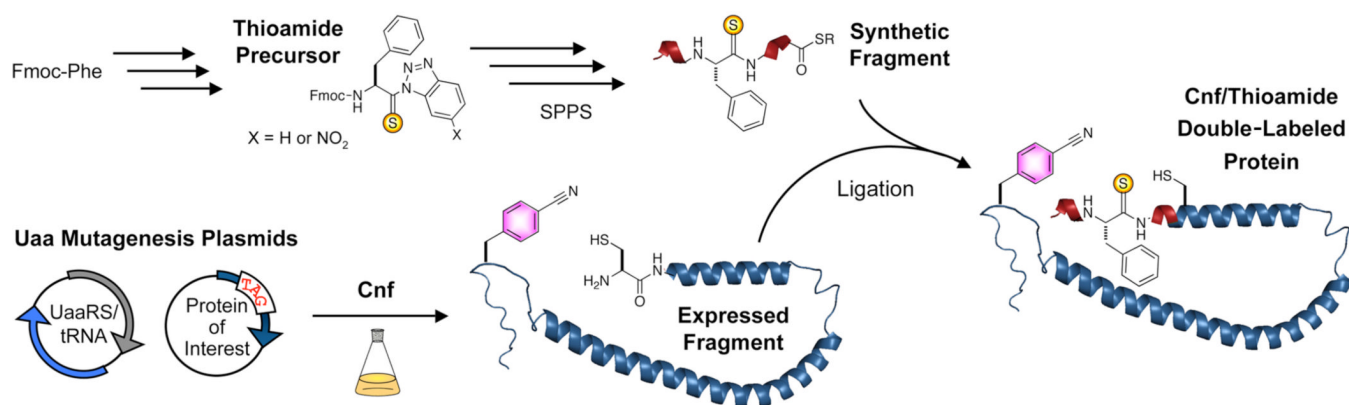
39. Hejjaoui M, Haj-Yahya M, Kumar KSA, Brik A, Lashuel HA. *Angew. Chem. Int. Ed.* 2011; 50:405–409.
40. Gentle IE, De Souza DP, Baca M. *Bioconj. Chem.* 2004; 15:658–663.
41. Flemer S. *J. Peptide Sci.* 2009; 15:693–696. [PubMed: 19790043]
42. Camarero JA, Mitchell AR. *Protein Peptide Lett.* 2005; 12:723–728.
43. Nagalingam AC, Radford SE, Warriner SL. *Synlett.* 2007:2517–2520.
44. Kang J, Macmillan D. *Organic & Biomolecular Chemistry.* 2010; 8:1993–2002. [PubMed: 20401371]
45. Kawakami T, Aimoto S. *Tetrahedron.* 2009; 65:3871–3877.
46. Young TS, Ahmad I, Yin JA, Schultz PG. *J. Mol. Biol.* 2010; 395:361–374. [PubMed: 19852970]
47. Uversky VN, Li J, Fink AL. *J. Biol. Chem.* 2001; 276:10737–10744. [PubMed: 11152691]
48. Vilar M, Chou HT, Luhrs T, Maji SK, Riek-Loher D, Verel R, Manning G, Stahlberg H, Riek R. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105:8637–8642. [PubMed: 18550842]
49. Desplats P, Lee HJ, Bae EJ, Patrick C, Rockenstein E, Crews L, Spencer B, Masliah E, Lee SJ. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:13010–13015. [PubMed: 19651612]
50. Volpicelli-Daley LA, Luk KC, Patel TP, Tanik SA, Riddle DM, Stieber A, Meaney DF, Trojanowski JQ, Lee VMY. *Neuron.* 2011; 72:57–71. [PubMed: 21982369]
51. Winner B, Jappelli R, Maji SK, Desplats PA, Boyer L, Aigner S, Hetzer C, Loher T, Vilar M, Campion S, Tzitzilonis C, Soragni A, Jessberger S, Mira H, Consiglio A, Pham E, Masliah E, Gage FH, Riek R. *Proc. Natl. Acad. Sci. U.S.A.* 2011; 108:4194–4199. [PubMed: 21325059]
52. Conway KA, Harper JD, Lansbury PT. *Biochemistry.* 2000; 39:2552–2563. [PubMed: 10704204]
53. Hejjaoui M, Butterfield S, Fauvet B, Vercruyse F, Cui J, Dikiy I, Prudent M, Olschewski D, Zhang Y, Eliezer D, Lashuel HA. *J. Am. Chem. Soc.* 2012; 134:5196–5210. [PubMed: 22339654]
54. Meier F, Abeywardana T, Dhall A, Marotta NP, Varkey J, Langen R, Chatterjee C, Pratt MR. *J. Am. Chem. Soc.* 2012; 134:5468–5471. [PubMed: 22404520]
55. Fauvet B, Fares MB, Samuel F, Dikiy I, Tandon A, Eliezer D, Lashuel HA. *J. Biol. Chem.* 2012
56. Allison JR, Varnai P, Dobson CM, Vendruscolo M. *J. Am. Chem. Soc.* 2009; 131:18314–18326. [PubMed: 20028147]
57. Celej MS, Sarroukh R, Goormaghtigh E, Fidelio GD, Ruyschaert JM, Raussens V. *Biochem. J.* 2012; 443:719–726. [PubMed: 22316405]
58. Karyagina I, Becker S, Giller K, Riedel D, Jovin TM, Griesinger C, Bennati M. *Biophys. J.* 2011; 101:L1–L3. [PubMed: 21723808]
59. Ramakrishnan M, Jensen PH, Marsh D. *Biochemistry.* 2006; 45:3386–3395. [PubMed: 16519533]
60. Wu KP, Baum J. *J. Am. Chem. Soc.* 2010; 132:5546–5547. [PubMed: 20359221]
61. Taskent-Sezgin H, Marek P, Thomas R, Goldberg D, Chung J, Carrico I, Raleigh DP. *Biochemistry.* 2010; 49:6290–6295. [PubMed: 20565125]
62. Nath A, Sammalkorpi M, DeWitt DC, Trexler AJ, Elbaum-Garfinkle S, O'Hern CS, Rhoades E. *Biophys. J.* 2012; 103:1940–1949. [PubMed: 23199922]
63. Ferreon ACM, Gambin Y, Lemke EA, Deniz AA. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:5645–5650. [PubMed: 19293380]
64. Ferreon ACM, Moosa MM, Gambin Y, Deniz AA. *Proc. Natl. Acad. Sci. U.S.A.* 2012; 109:17826–17831. [PubMed: 22826265]
65. Lee JC, Gray HB, Winkler JR. *J. Am. Chem. Soc.* 2005; 127:16388–16389. [PubMed: 16305213]
66. Lee JC, Langen R, Hummel PA, Gray HB, Winkler JR. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:16466–16471. [PubMed: 15536128]
67. Deniz AA, Dahan M, Grunwell JR, Ha TJ, Faulhaber AE, Chemla DS, Weiss S, Schultz PG. *Proc. Natl. Acad. Sci. U.S.A.* 1999; 96:3670–3675. [PubMed: 10097095]
68. Miyake-Stoner SJ, Miller AM, Hammill JT, Peeler JC, Hess KR, Mehl RA, Brewer SH. *Biochemistry.* 2009; 48:5953–5962. [PubMed: 19492814]
69. Tucker MJ, Oyola R, Gai F. *J. Phys. Chem. B.* 2005; 109:4788–4795. [PubMed: 16851563]
70. Makarov DE, Plaxco KW. *J. Chem. Phys.* 2009:131.
71. O'Brien EP, Morrison G, Brooks BR, Thirumalai D. *J. Chem. Phys.* 2009:130.



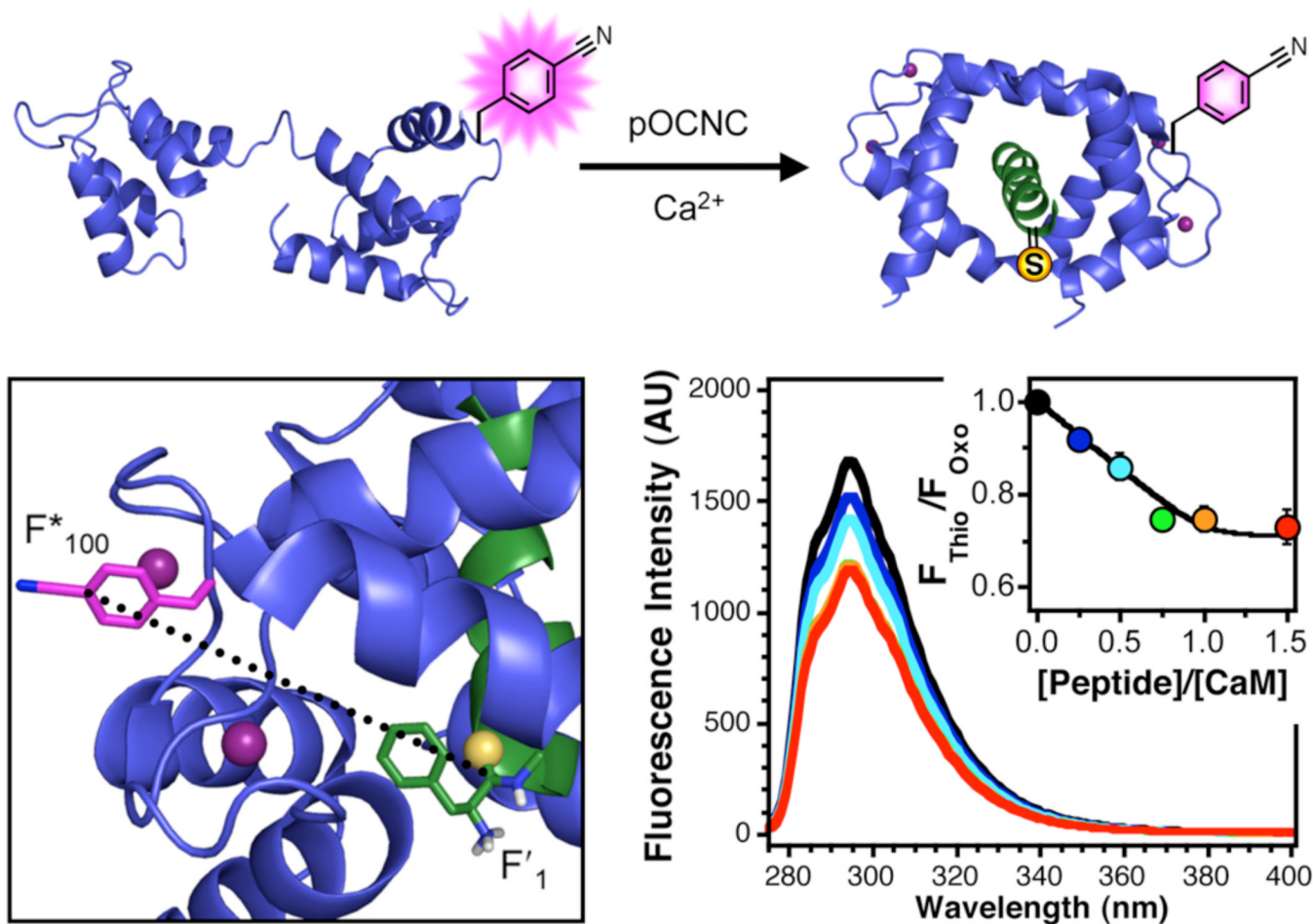
72. McCarney ER, Werner JH, Bernstein SL, Ruczinski I, Makarov DE, Goodwin PM, Plaxco KW. *J. Mol. Biol.* 2005; 352:672–682. [PubMed: 16095607]
73. Trexler AJ, Rhoades E. *Biophys. J.* 2010; 99:3048–3055. [PubMed: 21044603]
74. Wang JY, Xie JM, Schultz PG. *J. Am. Chem. Soc.* 2006; 128:8738–8739. [PubMed: 16819861]
75. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. *J. Am. Chem. Soc.* 2002; 124:9026–9027. [PubMed: 12148987]
76. Speight LC, Goldberg JM, Warner JB, Mehl RA, Petersson EJ. Unpublished results.
77. Speiser S. *Chem. Rev.* 1996; 96:1953–1976. [PubMed: 11848817]
78. Uversky VN, Li J, Fink AL. *FEBS Lett.* 2001; 509:31–35. [PubMed: 11734201]
79. Brustad EM, Lemke EA, Schultz PG, Deniz AA. *J. Am. Chem. Soc.* 2008; 130:17664–+. [PubMed: 19108697]
80. Huang Y, Russell WK, Wan W, Pai PJ, Russell DH, Liu WS. *Mol. Biosys.* 2010; 6:683–686.
81. Wan W, Huang Y, Wang ZY, Russell WK, Pai PJ, Russell DH, Liu WR. *Angew. Chem. Int. Ed.* 2010; 49:3211–3214.
82. Neumann H, Wang KH, Davis L, Garcia-Alai M, Chin JW. *Nature.* 2010; 464:441–444. [PubMed: 20154731]
83. Shalaby MA, Grote CW, Rapoport H. *J. Org. Chem.* 1996; 61:9045–9048. [PubMed: 11667896]
84. Hanahan D, Jessee J, Bloom FR. *Methods Enzymol.* 1991; 204:63–113. [PubMed: 1943786]



**Figure 1.** Cyanophenylalanine/Thioamide FRET Interactions. Top Left: Cyanophenylalanine (Cnf, F\*) and leucyl thioamide (Leu', L'). Top Right: Normalized absorption (dotted black line) and fluorescence emission (dashed black line) spectrum of Cnf and absorption spectrum of Leu' Ala (yellow line). The shaded area indicates the spectral overlap that contributes to FRET. Bottom: Distance ( $R_{\text{FRET}}$ ) calculations are made using Förster theory as described in text.

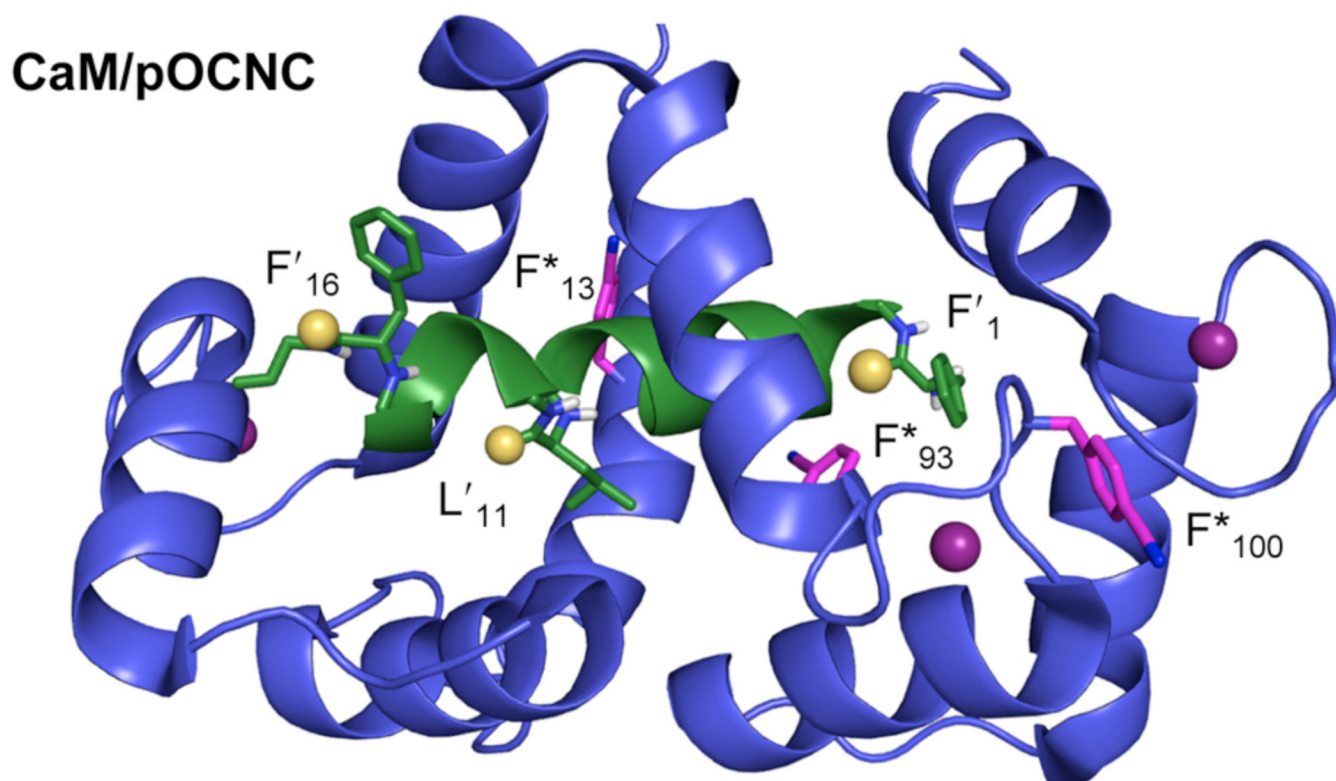


**Figure 2.** Combining Unnatural Amino Acid Mutagenesis and Native Chemical Ligation. Thioamide-bearing peptides are synthesized on solid phase starting from activated benzotriazole precursors and standard Fmoc-protected amino acids. For N-terminal thioamides, a thioester is installed as a masked group or formed by post-synthetic activation. To generate the Cnf-labeled protein fragment, *E. coli* cells are transformed with two DNA plasmids; one encoding the protein with a UAG stop codon at the site of interest, and one encoding an aminoacyl tRNA synthetase that is selective for an unnatural amino acid (UaaRS) with a tRNA that recognizes the UAG stop codon. Protein is expressed under standard conditions and the N-terminal Cys generated by proteolysis. Ligation of the two fragments gives the full length protein, labeled with both Cnf and the thioamide.

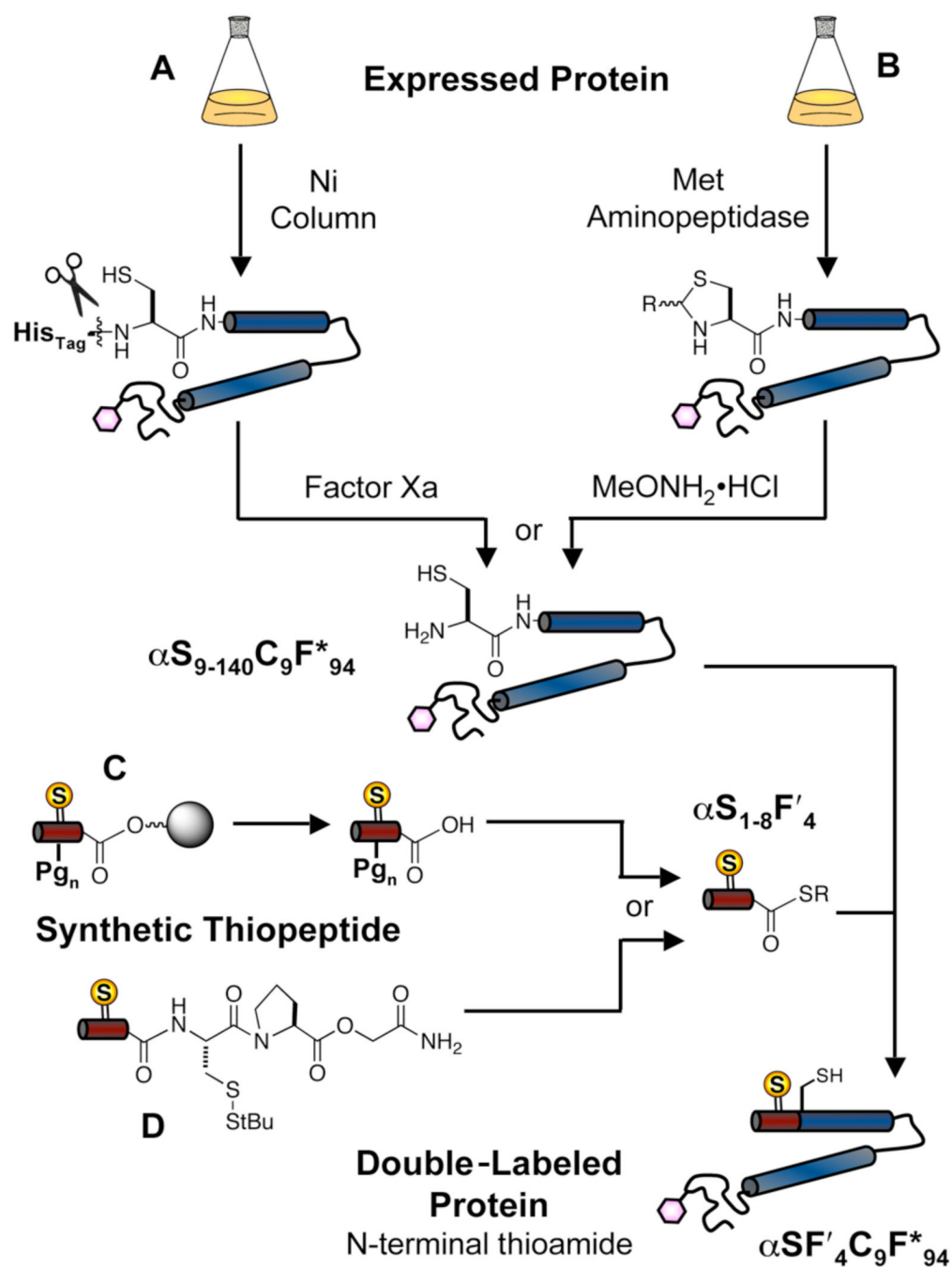


**Figure 3.**

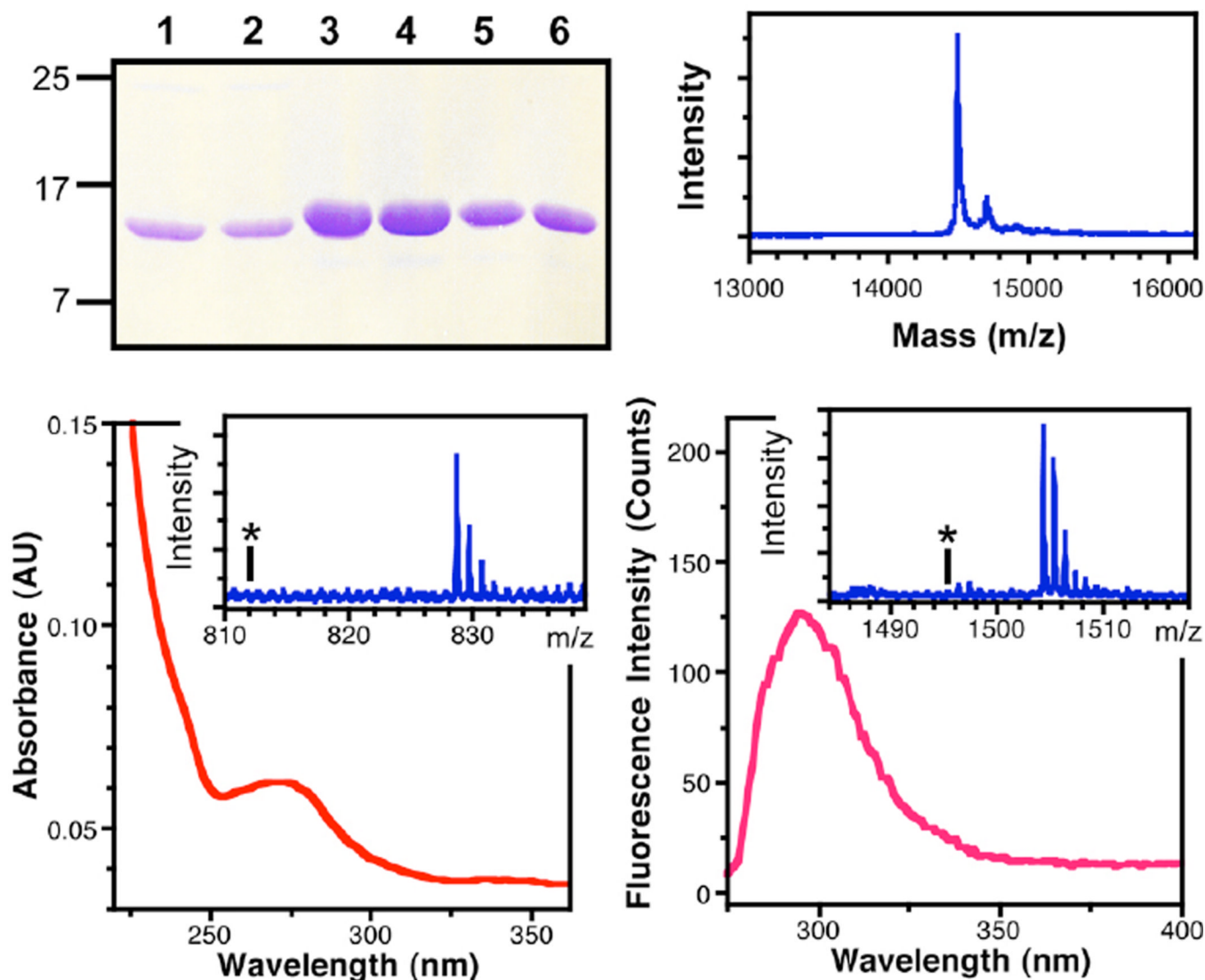
FRET in CaM/Peptide Complexes. Top: Binding of thioamide-labeled peptides (green) to Cnf-labeled CaM (blue) in the presence of Ca<sup>2+</sup>. Bottom Left: An image showing Cnf (pink) at position 100 in calmodulin (CaM<sup>F\*F\*100</sup>) and thiophenylalanine (yellow sulfur atom) at position 1 in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



**Figure 4.** CaM/pOCNC Label Sites. Cnf shown at three positions in CaM with three possible pOCNC thiopeptides (coloring as in Fig. 3). Images based on one of the 20 low energy structures in PDB ID 1SYD, modified using PyMol.

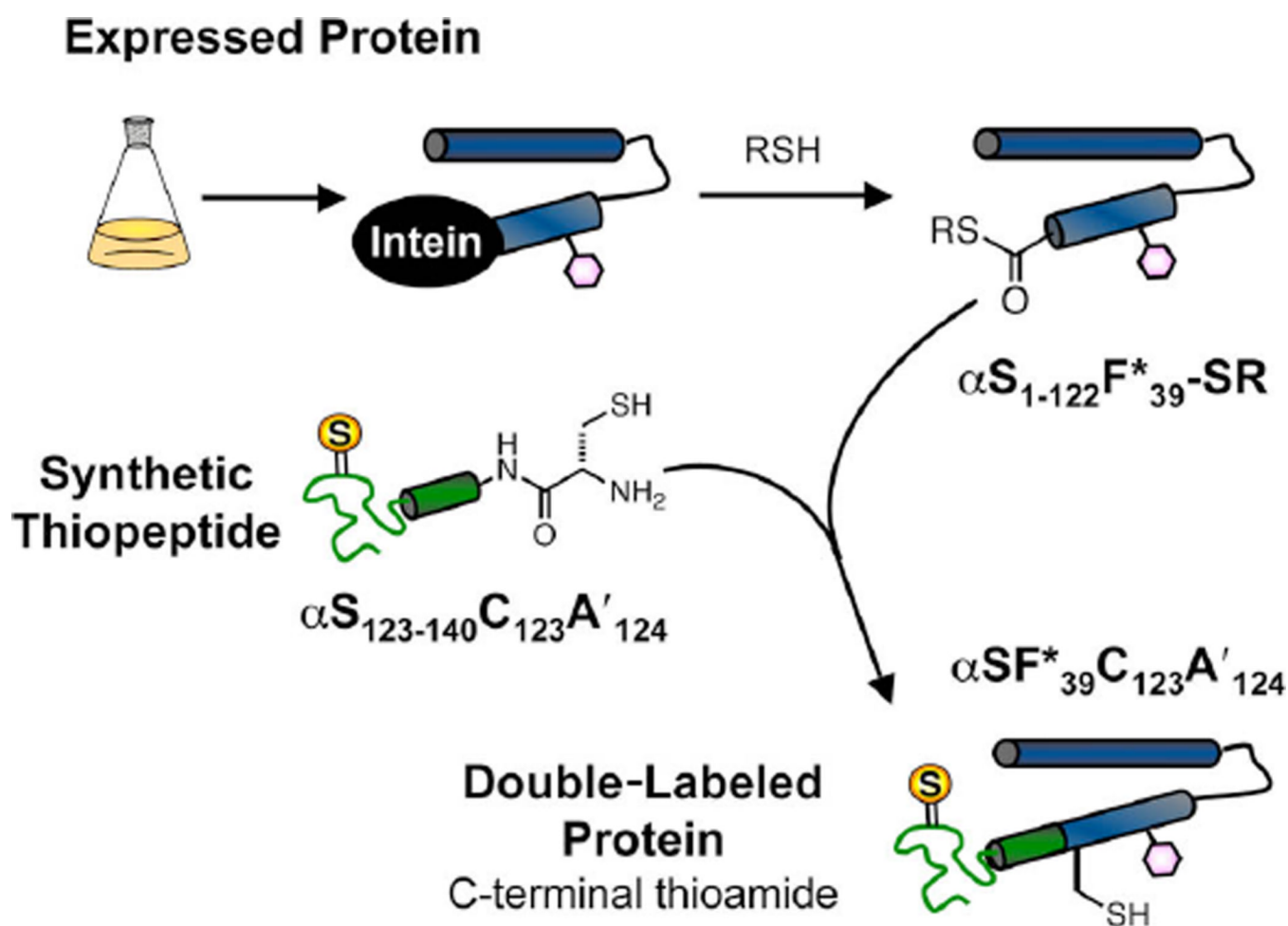


**Figure 5.** Semi-synthesis of Double-Labeled Proteins with N-terminal Thioamides. Expressed C-terminal protein fragment: A) Using a plasmid encoding an N-terminal His<sub>Tag</sub> with a Factor Xa site; B) Using a plasmid encoding the truncated protein where N-terminal Met is removed by Met aminopeptidase. MeONH<sub>2</sub>•HCl used to cleave the thiazolidine which forms with intracellular aldehydes. (R = Me or CH<sub>2</sub>CO<sub>2</sub>H) Synthetic N-terminal thiopeptide thioester: C) Synthesis on Cl-Trt resin and activation to form thioester with PyBOP, followed by deprotection with TFA; D) Synthesis on Rink amide resin with C<sup>b</sup>PG<sub>0</sub> masked thioester. The expressed protein fragment and synthetic thiopeptide are combined in the presence of thiophenol and the ligated, full-length protein isolated.



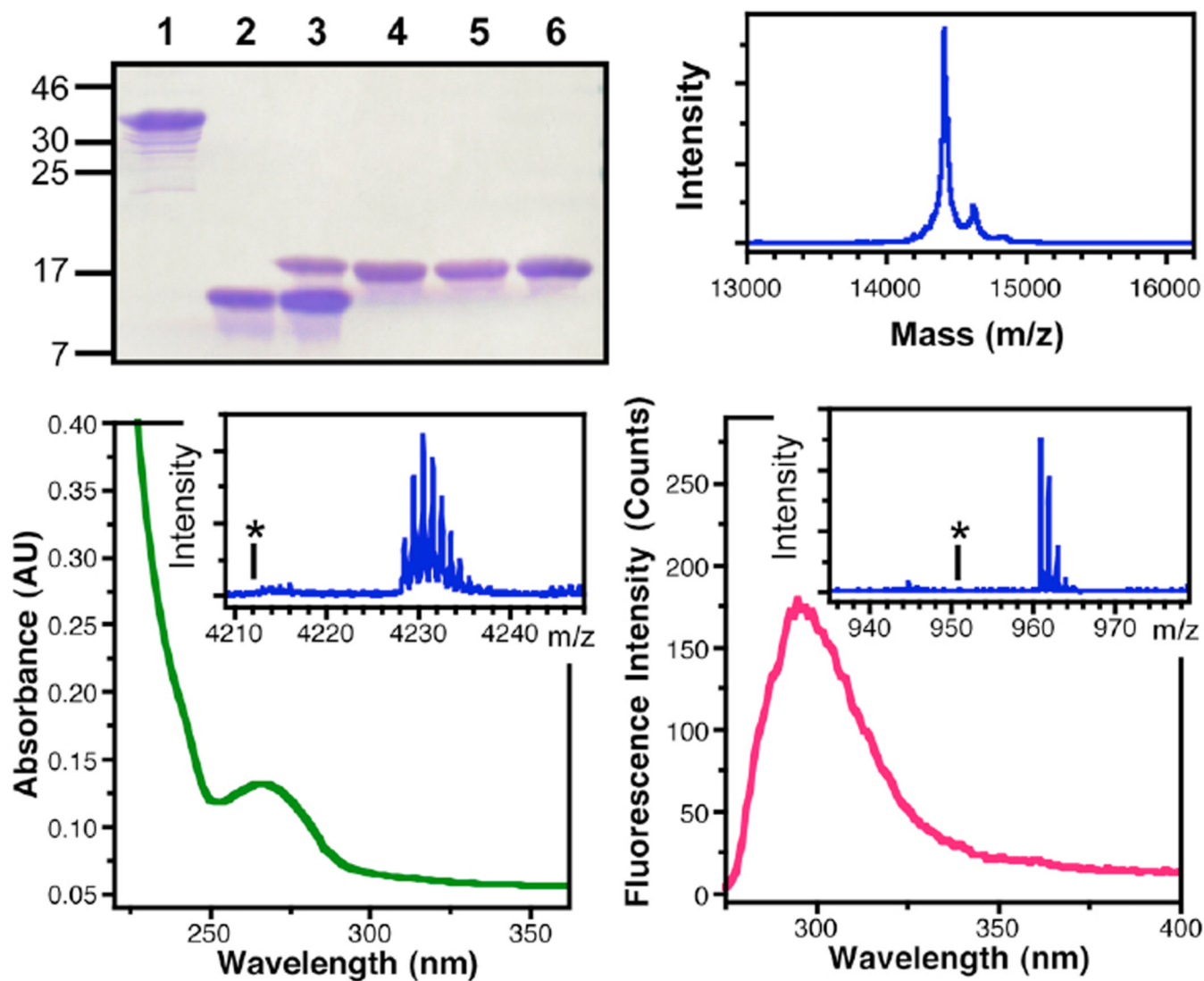
**Figure 6.**

Characterization of  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$  Ligation Product. Top Left: PAGE gel showing:  $\alpha\text{S}^{\text{F}'_{9-140}\text{C}_9\text{F}^*_{94}}$  before (1) and after (2)  $\text{MeONH}_2\cdot\text{HCl}$  treatment, ligation of  $\alpha\text{S}^{\text{F}'_{9-140}\text{C}_9\text{F}^*_{94}}$  to  $\alpha\text{S}^{\text{F}'_{1-8}\text{F}'_4\text{-SR}}$  thioester (3), purified  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$  (4),  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$  (5), and  $\alpha\text{S}^{\text{WT}}$  (6). See HPLC chromatogram in Supporting Information for separation of ligated product. Top Right: MALDI MS analysis of full-length  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$ . Bottom Left: UV/Vis absorption spectrum of  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$  showing thioamide absorption at 266 nm. Inset: MALDI MS analysis of trypsinized 1–6 fragment, confirming the presence of the thioamide at  $\text{F}'_4$ ; Calcd  $m/z$  (M+H): 828.35, Obsvd: 828.65. The asterisk indicates the absence of peaks corresponding an oxoamide at  $\text{Phe}_4$  in the 1–6 fragment; Calcd  $m/z$  (M+H): 812.38. Bottom Right: Fluorescence emission spectra of  $\alpha\text{S}^{\text{F}'_4\text{C}_9\text{F}^*_{94}}$  showing Cnf emission at 295 nm. Inset: MALDI MS analysis of trypsinized 81–96 fragment, confirming the presence of the Cnf at position 94; Calcd  $m/z$  (M+H): 1503.78, Obsvd: 1504.30. The asterisk indicates the absence of peaks corresponding Tyr at position 94 in the 81–96 fragment; Calcd  $m/z$  (M+H): 1494.78.



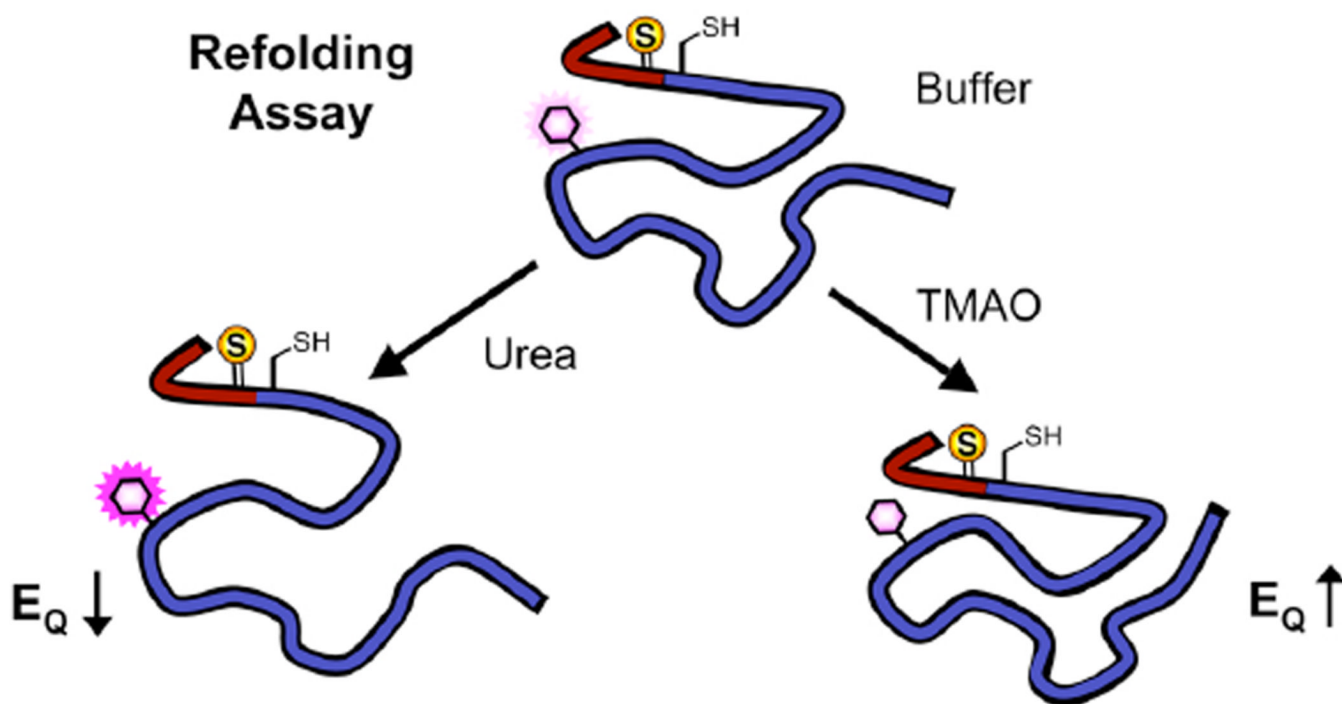
**Figure 7.** Semi-synthesis of Double-Labeled Proteins with C-terminal Thioamides. An expressed N-terminal protein fragment is generated using a plasmid encoding a truncated version of the protein with a C-terminal intein. Thiolysis of the stalled intein generates a thioester suitable for NCL. The synthetic thiopeptide with an N-terminal Cys is made by standard SPPS. The expressed protein fragment and synthetic thiopeptide are combined in the presence of thiophenol and ligated, full-length protein isolated.



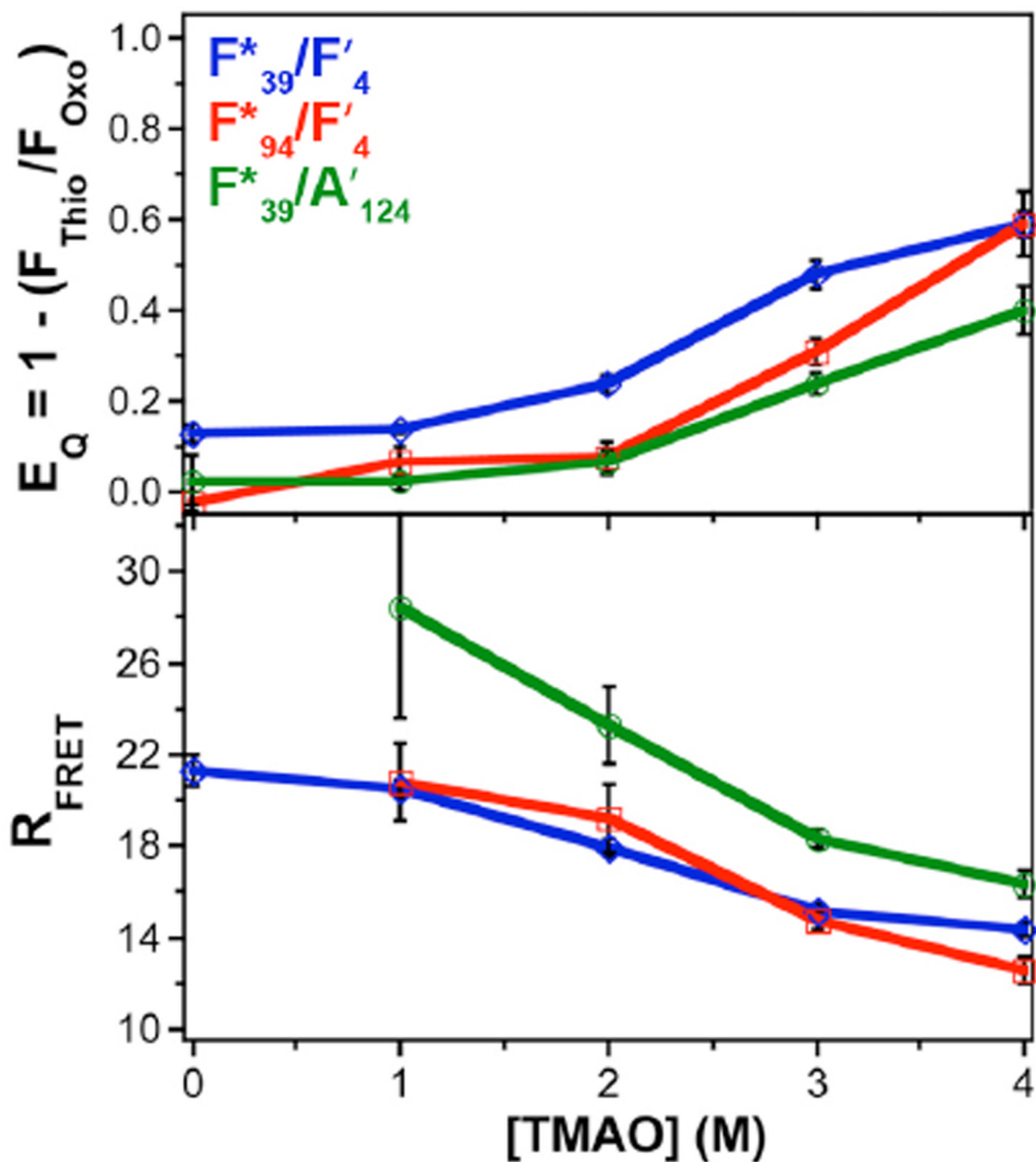


**Figure 8.**

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



**Figure 9.** Refolding Assay. Monomeric double-labeled  $\alpha$ S (e.g.  $\alpha$ S<sup>F'F\*</sup><sub>4C9F\*<sub>94</sub></sub>) is diluted to 1  $\mu$ M in either Tris buffer with 1 mM BME, or the same buffer with urea or TMAO added. Cnf fluorescence is measured at 295 nm ( $F_{\text{Thio}}$ ). Identical experiments are carried out with donor-only labeled  $\alpha$ S (e.g.  $\alpha$ S<sup>F</sup><sub>C9F\*<sub>94</sub></sub>) to determine  $F_{\text{Oxo}}$ .  $E_Q = 1 - F_{\text{Thio}}/F_{\text{Oxo}}$ .



**Figure 10.**

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

**Table 1**

## Distances in CaM/pOCNC Complexes

CaM/pOCNC Mutants ( $\Phi$ ) <sup>a</sup>	% E <sub>Q</sub> Obsvd <sup>b</sup>	R <sub>FRET</sub> (Å) <sup>c</sup>	R <sub>NMR</sub> (Å) <sup>d</sup>
All Phe control	0	—	—
F* <sub>13</sub> /F'₁ (0.003)	11 ± 1	13	14
F* <sub>13</sub> /L'₁₁ (0.003)	37 ± 2	11	14
F* <sub>17</sub>	NB <sup>e</sup>	—	—
F* <sub>93</sub> /F'₁ (0.014)	79 ± 2	12	9
F* <sub>93</sub> /L'₁₁ (0.014)	43 ± 4	13	12
F* <sub>100</sub> /F'₁ (0.110)	25 ± 3	20	19
F* <sub>100</sub> /L'₁₁ (0.110)	17 ± 2	21	24
F* <sub>100</sub> /F'₁₆ (0.110)	10 ± 3 <sup>f</sup>	23	31

<sup>a</sup> $\Phi$  determined by comparison to fluorescence emission of 10  $\mu$ M Cnf,  $\Phi = 0.110$ .

<sup>b</sup>E<sub>Q</sub> determined for 1:1 CaM/pOCNC complex.

<sup>c</sup>R<sub>FRET</sub> calculated from E<sub>Q</sub> as described in Supporting Information.

<sup>d</sup>R<sub>NMR</sub> is an average value calculated from the twenty lowest energy structures in PDB ID 1SYD.

<sup>e</sup>NB indicates no binding by native PAGE gel analysis.

<sup>f</sup>Mixture of R17 D- and L epimers of pOCNC-F'₁₆ peptide. Both bind with high affinity by native PAGE gel analysis.