JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Genetically Encoded 1,2-Aminothiols Facilitate Rapid and Site-Specific Protein Labeling via a Bio-orthogonal Cyanobenzothiazole Condensation

Duy P. Nguyen,[†] Thomas Elliott,[†] Matthew Holt,[‡] Tom W. Muir,[‡] and Jason W. Chin*[,†]

[†]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, U.K.

[‡]Laboratory of Synthetic Protein Chemistry, The Rockefeller University, New York, New York 10021, United States

Supporting Information

ABSTRACT: We report evolved orthogonal pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pairs that direct the efficient, site-specific incorporation of N^{ϵ}-L-thiaprolyl-L-lysine, N^{ϵ}-D-cysteinyl-L-lysine, and N^{ϵ}-L-cysteinyl-L-lysine into recombinant proteins in *Escherichia coli*. We demonstrate that the unique 1,2-aminothiol introduced by our approach can be efficiently, rapidly, and specifically labeled via a cyanobenzothiazole condensation to quantitatively introduce biophysical probes into proteins. Moreover, we show that, in combination with cysteine labeling, this approach allows the dual labeling of proteins with distinct probes at two distinct, genetically defined sites.

Functional groups that can participate in bioorthgonal reactions, including azides, alkynes, ketones, anilines, alkenes, and tetrazoles, have been genetically incorporated into proteins using amber suppressor aminoacyl-tRNA synthetase/$tRNA_{CUA}$ pairs.[1−11] This provides a route to the site-specific labeling of proteins at virtually any site. However, for established reactions that have been demonstrated on proteins the published rate constants for the corresponding model reactions[12] are in the range of 10^{-4} M^{-1} s^{-1} to 10^{-2} M^{-1} s^{-1}, although reactions with rate constants approaching 1 M^{-1} s^{-1} have recently been reported for emerging approaches.[13] The sluggishness of bio-orthogonal reactions for which a reaction partner can be encoded often means that large excesses of the labeling molecule and long reaction times are required to achieve protein labeling, and in practice, quantitative labeling may be challenging or impossible to achieve.

Other protein labeling methods are available including the use of fluorescent proteins, self-labeling tags (e.g., tetracysteine, tetraserine), self-labeling proteins (e.g., SNAPtag, HALOtag, CLIPtag), and ligases (e.g., biotin ligase, lipoic acid ligase, sortase, phosphopantetheinyl-transferase).[14] While some of these approaches allow rapid labeling and have had a substantial impact, they require either the use of protein fusions or the introduction of defined additional sequences of amino acids in the protein, which can perturb the structure and function of the protein and make it challenging to place probes at any position in a protein. In some cases the types of probes that can be incorporated with these approaches are limited, and some approaches are incompatible with the introduction of two independent probes into the same protein.

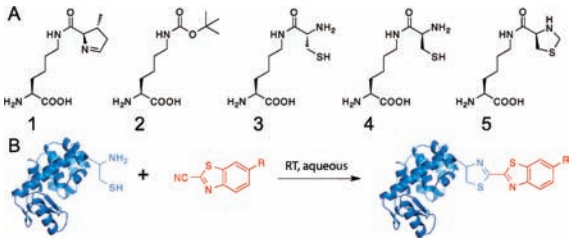
The condensation of 1,2-aminothiols with 2-cyanobenzothiazole (CBT) is the last step in the synthesis of luciferin.[15] The reaction proceeds under aqueous conditions at physiological temperatures and pressures. Since 1,2-aminothiols are not found in proteins and CBT does not react to form stable adducts with the functional groups found in natural proteins,[16] the 1,2-aminothiol/CBT combination is bio-orthogonal with respect to the functional groups found in natural proteins. Rao and co-workers have recently explored the use of this condensation to create polymer assemblies in mammalian cells[17] and to label proteins on the 1,2-aminothiol of an N-terminal cysteine, revealed by protease cleavage.[16] We were interested in developing methods for the site-specific labeling of proteins, at any user-defined site, via this reaction. Since the reaction between CBT and 1,2-aminothiols proceeds with a second-order rate constant of approximately 9 M^{-1} s^{-1}, it is orders of magnitude faster than reactions for which a reaction partner can currently be site specifically incorporated into proteins by genetic code expansion.[16]

Pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pairs (PylRS/$tRNA_{CUA}$) from *Methanosarcina* species, which naturally incorporate pyrrolysine (**1**) are orthogonal to endogenous tRNAs and aminoacyl-tRNA synthetases in *Escherichia coli*, yeast, and mammalian cells.[18−20] We and others have previously demonstrated that several unnatural amino acids, including post-translational modifications, chemical handles, and photocaged amino acids, can be site-specifically incorporated using this pair and its synthetically evolved derivatives.[9,10,18,19,21−24] Recently Chan and co-workers reported the site-specific incorporation of N^{ϵ}-D-cysteinyl-L-lysine **3** (Scheme 1) into a recombinant protein in response to an amber codon using the wild-type *M. mazei* PylRS/$tRNA_{CUA}$ pair.[22] Since **3** contains a 1,2-aminothiol we realized that it might be possible to site specifically label proteins containing **3** (or **4**) with CBT-based biophysical probes. In addition, since the 1,2-aminothiols of N-terminal cysteines commonly react with pyruvate in vivo[25,26] we investigated the incorporation of **5** in which the 1,2-aminothiol is protected from in vivo modification through the formation of a thiazolidine ring. This amino acid will not be alkylated by common reagents for labeling cysteine residues with biophysical probes but can easily be converted to the corresponding 1,2-aminothiol with methoxyamine,[25−27] which should permit labeling with CBT derivatives. We envisioned that this would be an advantage that would ultimately facilitate the dual labeling of proteins with two independent probes at distinct sites in a protein.

We began by synthesizing amino acids **3**, **4**, and **5**, as described in the Supporting Information (SI). We added **3** (2 mM) to *E. coli*

Received: April 5, 2011
Published: July 07, 2011

Scheme 1. (A) Structure of Pyrrolysine (pyl, 1), N^{ϵ}-*tert*-Butyloxycarbonyl-L-lysine (2), N^{ϵ}-D-Cysteinyl-L-lysine (3), N^{ϵ}-L-Cysteinyl-L-lysine (4), and N^{ϵ}-L-Thiaprolyl-L-lysine (5); (B) Condensation between 1,2-Aminothiol and CBT Derivatives



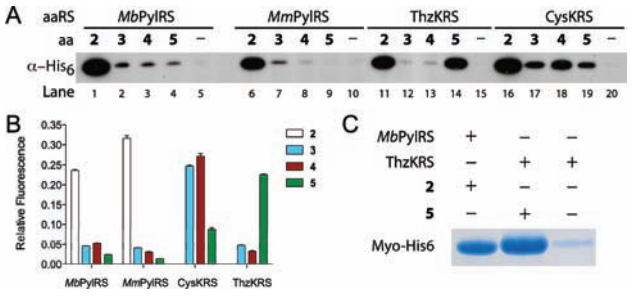


Figure 1. Analysis of incorporation efficiency of **3**, **4**, and **5** in myoglobin and GFP using different synthetases and amino acids. (A) Immunoblotting of sperm whale myoglobin-His6 on cell lysates derived from equal numbers of cells, as judged by OD_{600}. Amino acid [2 mM] (B) Fluorescence assay of incorporation efficiency by GFP expression. GFP is expressed with **2** (1 mM), **3**, **4**, **5** (10 mM). Data are shown for equal numbers of cells as judged by OD_{600}. (C) Expression of myoglobin-His6 incorporating **5**. Myoglobin-His6 is purified from cells containing pMyo4TAG-pylT-*his6* and pBK-*Mb*PylRS or pBK-ThzKRS in the presence of 1 mM **2** or 3 mM **5**.

containing a hexahistidine tagged sperm whale myoglobin gene with an amber codon at position 4 and the PylRS/$tRNA_{CUA}$ pairs from *Methanosarcina barkeri* (*Mb*) or *Methanosarcina mazei* (*Mm*).[1,19] In parallel, we performed identical experiments but replaced **3** with **4**, **5**, or **2**. N^{ϵ}-*tert*-butyloxycarbonyl-L-lysine (**2**) is a known substrate for the PylRS/$tRNA_{CUA}$ pair that has been previously incorporated by our laboratory and a number of other laboratories to give good yields of protein.[10,18,23,28,29] We used western blots on cell lysates (Figure 1 and Supplementary [SI] Figure 1) to evaluate the relative efficiency of incorporation with each amino acid/synthetase combination. This method of assessing amino acid incorporation does not introduce bias—that may result from purification steps—into the observed protein yield. As previously reported,[22] **3** is incorporated more efficiently than **4** using *Mm*PylRS/$tRNA_{CUA}$. In addition we find that **5** is not detectably incorporated using this pair. With the *Mb*PylRS/$tRNA_{CUA}$ all amino acids tested are detectably incorporated. However comparison of the myoglobin produced with the PylRS/$tRNA_{CUA}$ pairs in the presence of **3**, **4**, and **5** to the myoglobin produced in the presence of **2** demonstrates that the incorporation of amino acids **3**,**4**, and **5** using PylRS/$tRNA_{CUA}$ pairs is far from optimal. These results were further confirmed by adding each amino acid (**2**, **3**, **4**, and **5**) to *E. coli* containing the PylRS/$tRNA_{CUA}$ pairs and a green fluorescent protein (GFP) gene with an amber codon at position 150, a known permissive site. At each amino acid concentration tested (1, 2, 5, and 10 mM) the GFP resulting from incorporation of **3**, **4**, or **5** using the PylRS/$tRNA_{CUA}$ pairs was much lower than the GFP resulting from incorporating **2**, as judged by both quantitative fluorescence and western blot (Figure 1 and SI Figure 1). For example, the GFP fluorescence observed when **3** or **5** was added to cells at 10 mM was 8-fold or 24-fold lower, respectively, than when **2** was added to cells at 1 mM.

Since amino acids **3**, **4**, and **5** are very poorly incorporated by the PylRS/$tRNA_{CUA}$ pairs, we decided to evolve PylRS variants to efficiently direct the incorporation of each of these amino acids into proteins. To evolve the orthogonal *Mb*PylRS/$tRNA_{CUA}$ pair for the incorporation of **3** in response to the amber codon, we used a library in which six residues (L266, L270, Y271, L274, C313, and W383) in the active site of the synthetase[19] were randomized to all possible amino acids. These residues surround the pyrroline ring of bound pyrrolysine, and their mutation may allow the accommodation of the new unnatural amino acids. We performed three rounds of alternating positive and negative selections as previously described.[19] The surviving clones from the selection were transformed into cells containing pREP-pylT, which encodes *Mb*$tRNA_{CUA}$, and a chloramphenicol resistance gene with an amber codon at a permissive position. Fifteen clones out of the 96 we screened conferred chloramphenicol resistance up to 250 μg/mL chloramphenicol in the presence of **3**, but did not survive at 50 μg/mL chloramphenicol in the absence of the unnatural amino acid. Sequencing of the evolved synthetases revealed that they contain a C313 V mutation. We named the selected synthetase cysteinyl-lysyl-tRNA synthetase (CysKRS).

Efforts to directly evolve the *Mb*PylRS/$tRNA_{CUA}$ pair for the incorporation of **5** using the existing synthetase library were unsuccessful. Despite this observation the CysRS synthetase evolved from this library can direct the incorporation of **5**. To directly select for improved synthetases for **5** we created a new library of 10^8 *Mb*PylRS mutants in which five residues (A267, Y271, L274, C313, and M315) in the active site of the synthetase that also surround the bound pyrroline ring of pyrrolysine were randomized to all possible amino acids (SI). We performed five rounds of alternating positive and negative selections[19] on this library with **5**. Each of the 96 synthetases we tested from this selection survived on 300 μg/mL chloramphenicol in the presence of **5**, but did not survive at 50 μg/mL chloramphenicol in the absence of the unnatural amino acid. The 15 clones we sequenced had the same sequence, which contains three programmed mutations: A267S, C313V, M315F and one nonprogrammed mutation D344G. We named the resulting synthetase thiazolidinyl-lysyl-tRNA synthetase (ThzKRS).

To investigate the specificity and efficiency of the selected synthetases we compared the ability of each selected synthetase to incorporate each unnatural amino acid using both western blots and quantitative fluorescence assays at a range of substrate concentrations (Figure 1 and SI Figure 1). We found that the ThzKRS/$tRNA_{CUA}$ pair and the CysKRS/$tRNA_{CUA}$ pair can incorporate **5** (2 mM) and **3** or **4** (2 mM) respectively much more efficiently than the parent *Mb*PylRS/$tRNA_{CUA}$ pair, as judged by western blots. The increase in efficiency was further quantified by GFP assays. For example, when **5** (10 mM) is added to cells, 16-fold more GFP is produced using the ThzKRS/$tRNA_{CUA}$ pair than with the PylRS/$tRNA_{CUA}$ pair. Similarly the CysKRS/$tRNA_{CUA}$ pair incorporates **3** (10 mM) and **4** (10 mM) respectively 6-fold and 9-fold more efficiently

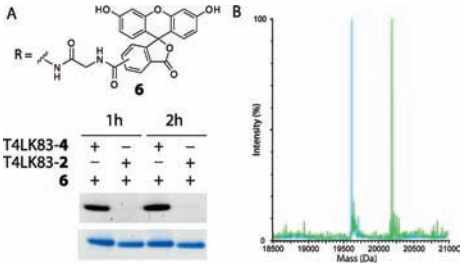


Figure 2. Efficient and specific labeling of a genetically encoded 1,2-aminothiol with CBT-fluorescein (**6**). (A) T4 lysozyme-His_6 containing **4** (T4L-K83-**4**) and **2** (T4L-K83-**2**) were labeled for 1 or 2 h with **6**. Top panel, in-gel fluorescence imaging; bottom panel, Coomassie-stained gel. (B) ESI-MS of the labeled protein confirms that T4L-K83-**4** (blue) is efficiently labeled with CBT-conjugated fluorescein probe **6** to produce fluorescein-labeled protein (green).

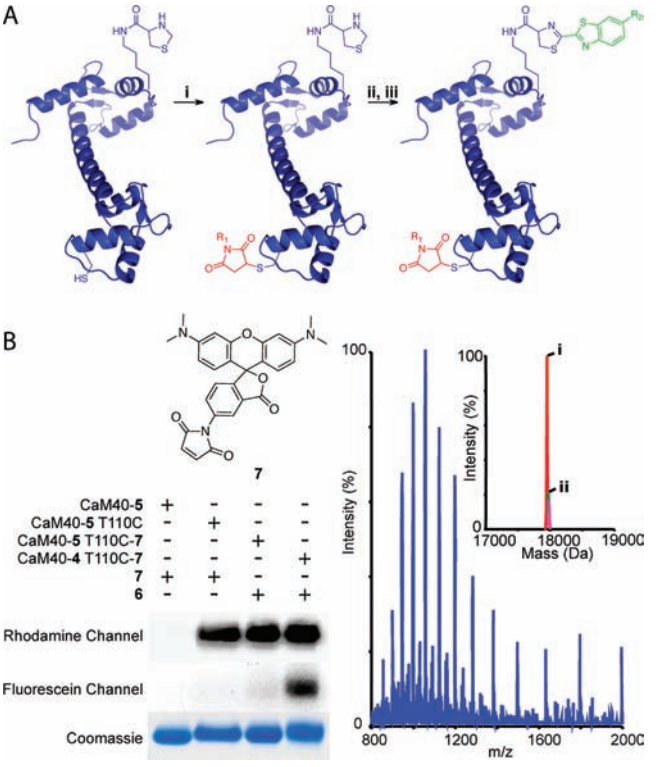


Figure 3. Site-specific double labeling of a genetically encoded cysteine and 1,2-aminothiol with TMR maleimide (**7**) and CBT-fluorescein (**6**) respectively. (A) **i.** Thiol specific cysteine labeling with **7**; **ii.** Thiazolidine cleavage with methoxyamine; **iii.** CBT condensation between 1,2-aminothiol and CBT derivative **6**. (B) Structure of TMR maleimide dye (**7**), in gel fluorescence of cysteine labeled and double labeled calmodulin and ESI-MS analysis of calmodulin double labeled with **6** and **7**. Calmodulin bearing **5** at position 40 and cysteine at position 110 (CaM40-**5** T110C) was labeled with **7** under standard conditions to quantitatively yield the cysteine labeling product (CaM40-**5** T110C-**7**). Control experiments used CaM40-**5**, which lacks the cysteine. The thiazolidine (**5**) in CaM40-**5** T110C-**7** was quantitatively deprotected with methoxyamine to yield (CaM40-**4** T110C-**7**), which was quantitatively labeled with **6**. The control for double labeling used CaM40-**5** T110C-**7**. The Rhodamine Channel detects red fluorescence. The Fluorescein Channel detects green fluorescence. ESI-MS analysis of the final double labeled calmodulin shows a mass **i.** 17954 ± 1.0 Da (expected mass 17954.6 Da) and minor peak **ii.** 17975 Da corresponds to sodium adduct. Mass spectra of each intermediate in labeling are in SI Figure 5.

than the unevolved synthetase incorporates these amino acids. Similar results were observed at a range of other amino acid concentrations (SI Figure 1). The levels of protein produced with **3** (10 mM) or **4** (10 mM) and the CysKRS/$tRNA_{CUA}$ pair and with **5** (10 mM) and the ThzKRS/$tRNA_{CUA}$ pair are comparable to those produced with 1 mM N^{ε}-*tert*-butyloxycarbonyl-L-lysine (**2**) and the wild-type synthetase/tRNA pairs. This demonstrates that it is now possible to incorporate these amino acids with an absolute protein yield comparable to that produced when incorporating a good substrate for the wild-type PylRS/$tRNA_{CUA}$ pair.

Myoglobin expressed in the presence of **5** (3 mM) was purified by Ni-NTA chromatography (Figure 1C) with a yield of 2 mg/L, which is comparable to yields obtained with other unnatural amino acids of demonstrated utility, and ESI-MS of the purified myoglobin confirmed the incorporation of **5** (SI Figure 2A). When we used CysKRS to incorporate **3** into T4 lysozyme, ESI-MS revealed peaks corresponding to the pyruvate adduct and its decarboxylation product attached to the purified protein (SI Figure 2B).[25,26] This is in good agreement with previous reports on the in vivo modification of N-terminal cysteines by cellular metabolites.[25,26,30]

To demonstrate that recombinant proteins can be efficiently and site-specifically labeled with CBT probes, we expressed and purified T4 lysozyme (T4L) incorporating **5** at position 83 in good yield (2 mg/L) from cells containing the ThzKRS/$tRNA_{CUA}$ pair and a T4L gene with an amber codon at position 83. ESI-MS demonstrated the incorporation of the thiazolidine **5** (SI Figure 3A). The purified protein incorporating **5** was treated with 200 mM *O*-methylhydroxyamine in 6 M guanidinium chloride at pH4 for 4 h[27,30] to quantitatively reveal **4** (SI Figure 3B). The protein was labeled with five equivalents of CBT-fluorescein **6** in PBS buffer (pH7.4) supplemented with 1 mM DTT at room temperature. SDS-PAGE analysis shows that labeling is complete after 2 h, whereas no nonspecific labeling was detected with the control protein containing amino acid **2** at the same site (Figure 2A). ESI-MS confirms the near-quantitative site-specific labeling of the protein with the fluorescein probe **6** (Figure 2B and SI Figure 3C). In contrast, when we incorporated an aliphatic alkyne N^{ε}-[(2-propynyloxy)carbonyl]-L-lysine into T4 lysozyme at position 83 using *Mb*PylRS/$tRNA_{CUA}$ pair and attempted to label the protein with 10 equiv of a fluorescein azide in a Cu(I)-catalyzed reaction,[10] the labeling was only 50% complete after 24 h (SI Figure 3E). This demonstrates that the rate acceleration for the CBT 1,2-aminothiol reaction translates into a practical advantage in protein labeling.

To further demonstrate the generality of this approach we expressed calmodulin containing **5** at position 40 in good yield (1 mg/L) from cells containing the ThzKRS/$tRNA_{CUA}$ pair and a calmodulin gene with an amber codon at position 40 and demonstrated that this could also be quantitatively deprotected and site specifically labeled with **6** (SI Figure 4). Our initial experiments on T4L used chaotropic denaturant and low pH to rapidly deprotect the thiazolidine, as is commonly used for peptides.[27] We have subsequently shown that the thiazolidine (**5**) in recombinant proteins can be deprotected using 200 mM *O*-methylhydroxyamine in PBS at pH 7 (SI Figure 4B).

To demonstrate that this approach can be used for the dual labeling of proteins at distinct sites with independent

probes we expressed calmodulin incorporating **5** at position 40 and bearing a threonine→cysteine mutation at position 110, from cells containing the ThzKRS/$tRNA_{CUA}$ pair and a calmodulin gene containing a threonine→cysteine codon change at position 110 and an amber codon at position 40.

We quantitatively labeled the protein at position 110 with a maleimide derivative of rhodamine (**7**), deprotected the thiazolidine, and labeled the resulting 1,2-aminothiol with CBT-fluorescein (**6**), as described above for the single labeling. SDS-PAGE analysis and ESI-MS demonstrate the quantitative site-specific dual labeling of the protein (Figure 3 and SI Figure 5).

In summary, we have demonstrated the efficient, genetically directed, site-specific introduction of aliphatic amino acids containing a 1,2-aminothiol chemical handle (**3**,**4**,**5**). We demonstrated that proteins containing **5** can be site-specifically and rapidly labeled via the CBT condensation. The reaction is compatible with cysteine labeling[16] for dual protein labeling, and we have demonstrated that this facilitates labeling with two independent probes at distinct genetically defined sites. Moreover, the CBT condensation should be compatible with, and orthogonal to, the functional groups for other bio-orthogonal reactions that can be encoded. This will facilitate the dual labeling of proteins, regardless of their cysteine content.[31]

■ ASSOCIATED CONTENT

Supporting Information. Experimental protocols and supplementary data. This material is available free of charge via the Internet. http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
chin@mrc-lmb.cam.ac.uk

■ ACKNOWLEDGMENT

This work was supported by Medical Research Council, UK and a grant from European Research Council to J.W.C. D.N. is grateful for a Laboratory of Molecular Biology (LMB) Cambridge scholarship.

■ REFERENCES

(1) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.
(2) Zhang, Z.; Wang, L.; Brock, A.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2840–2842.
(3) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. *Science* **2003**, *301*, 964–967.
(4) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 11782–11783.
(5) Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 935–939.
(6) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56–61.
(7) Deiters, A.; Cropp, T. A.; Summerer, D.; Mukherji, M.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5743–5745.
(8) Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. *Chem. Commun.* **2008**, 1205–1207.
(9) Fekner, T.; Li, X.; Lee, M. M.; Chan, M. K. *Angew. Chem., Int. Ed.* **2009**, *48*, 1633–1635.
(10) Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2009**, *131*, 8720–8721.
(11) Wang, J.; Zhang, W.; Song, W.; Wang, Y.; Yu, Z.; Li, J.; Wu, M.; Wang, L.; Zang, J.; Lin, Q. *J. Am. Chem. Soc.* **2010**, *132*, 14812–14818.
(12) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644–648.
(13) Wang, Y.; Song, W.; Hu, W. J.; Lin, Q. *Angew. Chem., Int. Ed.* **2009**, *48*, 5330–3.
(14) Hinner, M. J.; Johnsson, K. *Curr. Opin. Biotechnol.* **2010**, *21*, 766–776.
(15) White, E. H.; McCapra, F.; Field, G. F.; McElroy, W. D. *J. Am. Chem. Soc.* **1961**, *83*, 2402–2403.
(16) Ren, H.; Xiao, F.; Zhan, K.; Kim, Y. P.; Xie, H.; Xia, Z.; Rao, J. *Angew. Chem., Int. Ed.* **2009**, *48*, 9658–9662.
(17) Liang, G.; Ren, H.; Rao, J. *Nature Chem.* **2010**, *2*, 54–60.
(18) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818–822.
(19) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Chem. Biol.* **2008**, *4*, 232–234.
(20) Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 14819–14824.
(21) Polycarpo, C. R.; Herring, S.; Bérubé, A.; Wood, J. L.; Söll, D.; Ambrogelly, A. *FEBS Lett.* **2006**, *580*, 6695–6700.
(22) Li, X.; Fekner, T.; Ottesen, J. J.; Chan, M. K. *Angew. Chem., Int. Ed.* **2009**, *48*, 9184–9187.
(23) Nguyen, D. P.; Garcia Alai, M. M.; Kapadnis, P. B.; Neumann, H.; Chin, J. W. *J. Am. Chem. Soc.* **2009**, *131*, 14194–14195.
(24) Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 4086–4088.
(25) Chiang, K. P.; Jensen, M. S.; McGinty, R. K.; Muir, T. W. *ChemBioChem* **2009**, *10*, 2182–2187.
(26) Ottesen, J. J.; Bar-Dagan, M.; Giovani, B.; Muir, T. W. *Biopolymers* **2008**, *90*, 406–414.
(27) Villain, M.; Vizzavona, J.; Rose, K. *Chem. Biol.* **2001**, *8*, 673–679.
(28) Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W. *Nat. Chem. Biol.* **2010**, *6*, 750–757.
(29) Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S.; Chin, J. W. *Chem. Biol.* **2010**, *17*, 1072–1076.
(30) Gentle, I. E.; De Souza, D. P.; Baca, M. *Bioconjugate Chem.* **2004**, *15*, 658–663.
(31) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441–444.