JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Concerted, Rapid, Quantitative, and Site-Specific Dual Labeling of Proteins

Amit Sachdeva,† Kaihang Wang,† Thomas Elliott, and Jason W. Chin*

Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, United Kingdom

Supporting Information

ABSTRACT: Rapid, one-pot, concerted, site-specific labeling of proteins at genetically encoded unnatural amino acids with distinct small molecules at physiological pH, temperature, and pressure is an important challenge. Current approaches require sequential labeling, low pH, and typically days to reach completion, limiting their utility. We report the efficient, genetically encoded incorporation of alkyne- and cyclopropene-containing amino acids at distinct sites in a protein using an optimized orthogonal translation system in *E. coli*. and quantitative, site-specific, one-pot, concerted protein labeling with fluorophores bearing azide and tetrazine groups, respectively. Protein double labeling in aqueous buffer at physiological pH, temperature, and pressure is quantitative in 30 min.

The ability to attach two distinct molecules to programmed sites in proteins will facilitate a variety of applications including FRET[1,2] to study protein structure, conformation, and dynamics. Several approaches for the double labeling of proteins have been reported. One approach relies on the installation of one unnatural amino acid that is specifically labeled in combination with cysteine thiol labeling, but this approach is generally limited to proteins that do not contain more than one free thiol.[3,4] Chemical ligation approaches can be combined with the genetic encoding of a single unnatural amino acid for protein labeling,[5] but this may limit the size and/or sites that may be labeled. Perhaps the most generally applicable approach for protein double labeling is based on the genetic incorporation of two distinct amino acids in response to two distinct codons introduced at user defined sites in the gene of interest.

An ideal strategy for dual labeling requires (i) the efficient, cellular incorporation of two distinct unnatural amino acids bearing bioorthogonal functional groups that do not react together, into a protein and (ii) the quantitative, rapid, site-specific labeling of each encoded functional group at physiological temperature, pressure, and pH upon the simultaneous addition of both labeling reagents.

The cellular, genetically directed incorporation of two distinct unnatural amino acids into proteins has been demonstrated in response to an amber codon and a quadruplet codon,[6] two distinct stop codons,[7,8] or two distinct quadruplet codons.[9] We previously demonstrated the evolution of an orthogonal ribosome (ribo-Q1) that efficiently reads quadruplet codons and amber codons on an orthogonal mRNA using cognate extended anticodon tRNAs or amber suppressors, respectively.[6] We demonstrated that the PyrrolysyltRNA synthetase (PylRS)/tRNA pair and synthetically evolved derivatives of the *Methanococcus janaschii* Tyrosyl-tRNA synthetase (*Mj*TyrRS)/*Mj*tRNA pair are mutually orthogonal in their aminoacylation specificity and can be used to direct the incorporation of pairs of unnatural amino acids in response to amber and quadruplet codons.[6] We recently described several major advances in this system, including the evolution of a series of quadruplet decoding tRNAs based on the PylRS/tRNA pair that efficiently direct the incorporation of unnatural amino acids in response to quadruplet codons using the evolved orthogonal translation machinery.[9] We demonstrated efficient incorporation of numerous pairs of unnatural amino acids using the evolved PylRS/$tRNA_{UACU}$ pair and derivatives of the *Mj*TyrRS/$tRNA_{CUA}$ pair with orthogonal messages bearing AGTA and TAG codons and ribo-Q1, as well as the incorporation of unnatural amino acids in response to two distinct quadruplet codons.[9]

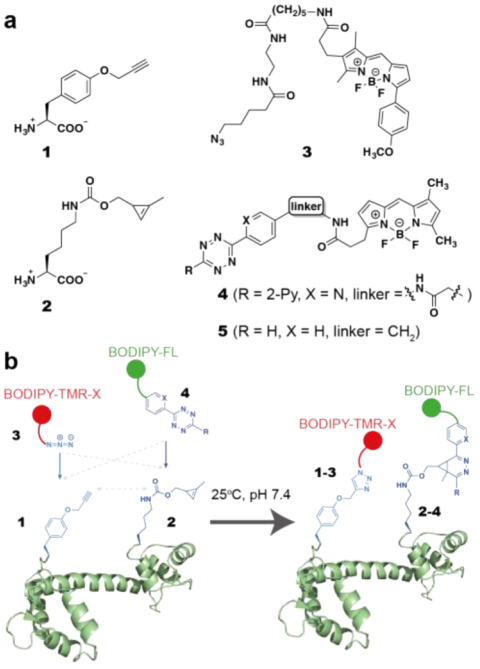
A variety of approaches have been reported for labeling two distinct bioorthogonal groups in proteins. These approaches are slow, typically taking tens of hours to days to reach completion. Azides and alkynes have been encoded in the same protein,[6,7] but these react together when placed in proximity[6] and the alkynes and azide probes used to label them will react together if added simultaneously. Azides and ketones have been encoded,[8,10] but unnatural amino acids bearing azides are prone to reduction,[8,11] and ketone mediated reactions commonly require a low pH and have very slow rates (rate constant approximately 10^{-4} M^{-1} s^{-1}).[12] We recently genetically installed a deactivated tetrazine containing amino acid[13] and a norbornene containing amino acid[14−16] at distinct sites in a single protein[9] and selectively labeled the encoded amino acids with fluorophores for FRET studies.[9] However, the labeling reactions, while rapid and proceeding at physiological temperature and pH, did need to be implemented sequentially to avoid reactions between the two labeling reagents.

A promising pair of mutually orthogonal reactions for one-pot labeling under aqueous conditions at physiological pH are the Cu(I)-catalyzed (3 + 2) cycloaddition between azides and terminal alkynes,[17] and the inverse electron demand Diels−Alder reaction of strained alkenes and tetrazines[18−23] (Scheme 1). The reaction of strained alkynes and azides can also be orthogonal to strained alkene−tetrazine reactions, but since tetrazines react with strained alkynes, this approach requires

Received: December 20, 2013
Published: May 23, 2014

   dx.doi.org/10.1021/ja4129789 | J. Am. Chem. Soc. 2014, 136, 7785−7788

Scheme 1. Concerted, Rapid, One-Pot Quantitative Dual Labeling of Proteins in Aqueous Medium at Physiological pH and Temperature[a]



[a](a) Unnatural amino acids and fluorophores used in this study.(b) Concerted labeling at an encoded terminal alkyne and an encoded cyclopropene via mutually orthogonal cycloadditions.

careful tuning of the rate constants for each reaction.[24] No combination of (3 + 2) cycloaddition and inverse electron demand Diels–Alder reaction has been demonstrated for labeling a single protein.

We recently demonstrated that a 1,3 disubstituted cyclopropene containing amino acid, **2**, can be efficiently and site-specifically incorporated into proteins using the PylRS/$tRNA_{CUA}$ pair.[25] This amino acid is smaller than most bioorthogonal dienophiles and reacts with tetrazines[19,26] with an on-protein rate constant of 27 M^{-1} s^{-1}.[25] Here we demonstrate the efficient genetic encoding of a terminal alkyne containing amino acid **1** and a cyclopropene containing amino acid **2** into a single protein and their rapid, quantitative, one-pot labeling with azide and tetrazine probes (Scheme 1). This work provides the first approach to the concerted double labeling of proteins in a one-pot process under aqueous conditions, at physiological pH, and provides a step change in the speed of double labeling, from days in previous work to 30 min in the approach reported here.

Proteins containing either **1** or **2** were overexpressed to examine the specificity of the proposed labeling reactions. A fusion protein of glutathione-S-transferase and calmodulin (GST-CaM) with amino acid **1** at position 1 in calmodulin was expressed from cells containing ribo-Q1 (an evolved orthogonal ribosome[6,27,28]), O-*gst-cam*$_{1TAG}$ (a fusion gene between *glutathione-S-transferase* (*gst*) and *calmodulin* (*cam*) on an orthogonal message[29] in which the first codon of *cam* is replaced with a TAG codon), and *Mj*PrpRS/$tRNA_{CUA}$ (a synthetase/tRNA pair developed for incorporating **1** in response to the TAG codon)[30] grown in the presence of **1** (2 mM). The yield of GST-CaM$\mathbf{1}_1$ was 4 to 5 mg per L of culture. The GST tag was subsequently removed by cleavage using thrombin at an engineered thrombin-cleavage site between GST and CaM. CaM$\mathbf{1}_1$ (CaM containing **1** at position 1, ~100 pmol) was labeled with the azide containing fluorophore **3** (2 nmole), in a Cu(I)-catalyzed click reaction. The reaction was quantitative as judged by both the quantitative shift of the fluorescently labeled protein by SDS-PAGE and electrospray ionization mass spectrometry (ESI-MS) (Figure 1a).

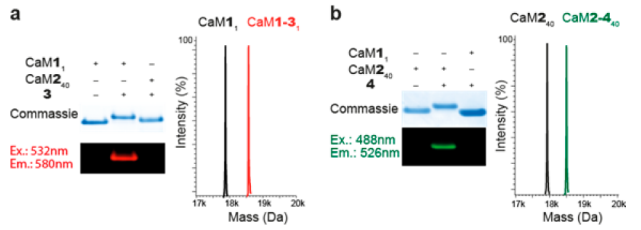


Figure 1. Specific protein labeling at genetically encoded unnatural amino acids **1** and **2**. (a) Genetically encoded **1**, but not **2**, in calmodulin is specifically labeled with probe **3**. Coomassie and fluorescence images demonstrate the specificity of labeling, and ESI MS before labeling (black, expected mass: 17875, found mass: 17874) and after labeling (red, expected mass: 18553, found mass: 18552) demonstrate the reaction is quantitative. (b) Genetically encoded **2**, but not **1**, in calmodulin is specifically labeled with probe **4**. Coomassie and fluorescence images demonstrate the specificity of labeling, and ESI MS before labeling (black, expected mass: 17930, found mass: 17930) and after labeling (green, expected mass: 18484, found mass: 18485) demonstrate the reaction is quantitative. Raw (before deconvolution) ESI-MS spectra in Supplementary Figure 2.

The cyclopropene containing amino acid, **2**, was site specifically incorporated at position 40 of calmodulin. The modified protein was expressed in cells bearing the PylRS/$tRNA_{CUA}$ (that efficiently directs the site-specific incorporation of **2**),[25] ribo-Q1, and O-*gst-cam*$_{40TAG}$ grown in the presence of **2** (1 mM). The yield of GST-CaM$\mathbf{2}_{40}$ was 4 to 5 mg per L of culture. CaM$\mathbf{2}_{40}$ (~100 pmol) (obtained after thrombin cleavage of the GST tag) was labeled with the tetrazine containing fluorophore **4** (2 nmol). The reaction was quantitative as judged by both the quantitative shift of the fluorescently labeled protein by SDS-PAGE and electrospray ionization mass spectrometry (ESI-MS) (Figure 1b). CaM$\mathbf{2}_{40}$was not labeled with **3** under the conditions that led to quantitative labeling of CaM$\mathbf{1}_1$ with **3** (Figure 1a). Similarly, CaM$\mathbf{1}_1$ was not labeled with **4** under conditions where CaM$\mathbf{2}_{40}$ was quantitatively labeled with **4**. These experiments demonstrate that the two labeling reagents react quantitatively with their target amino acid, but do not react with nontargeted unnatural or natural amino acids in proteins.

Next, we investigated labeling **1** and **2** within the same protein. We site-specifically incorporated **1** and **2** at positions 1 and 40 of calmodulin to produce CaM$\mathbf{1}_1\mathbf{2}_{40}$ (Figure 2). We directed the incorporation of amino acid **1** with an *Mj*PrpRS/$tRNA_{CUA}$ pair and the incorporation of amino acid **2** with the evolved PylRS/$tRNA_{UACU}$ pair, which efficiently decodes the quadruplet AGTA codon on orthogonal messages using ribo-Q1.[9] Unnatural amino acids were incorporated in response to UAG and AGTA codons at positions 1 and 40 in calmodulin within a GST-calmodulin gene on an orthogonal message (O-*gst-cam*$_{1TAG\text{-}40AGTA}$). Expression of full-length GST-CaM$\mathbf{1}_1\mathbf{2}_{40}$ was dependent on the addition of amino acids **1** and **2** to *E. coli*, and ESI-MS demonstrated the genetically directed incorpo-

dx.doi.org/10.1021/ja4129789 | J. Am. Chem. Soc. 2014, 136, 7785–7788

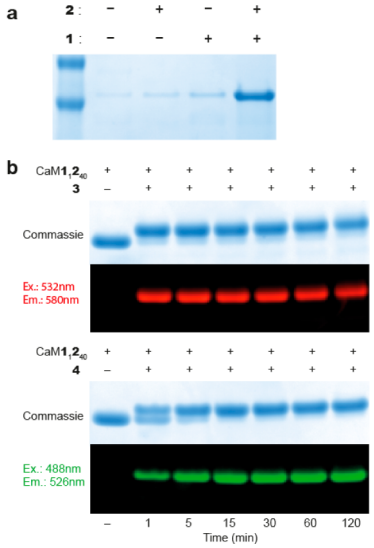


Figure 2. Incorporating **1** and **2** at positions 1 and 40 of calmodulin and the kinetics of specific labeling. (a) Expression was performed in *E. coli* bearing ribo-Q1, O-*gst-cam*$_{1TAG\text{-}40AGTA}$, the PylRS/tRNA$_{UACU}$ pair, and the *Mj*PrpRS/tRNA$_{CUA}$ pair. Amino acids **1** and **2** were used at 2 and 1 mM, respectively. (b) Labeling time course for reaction of CaM**1**$_1$**2**$_{40}$ with **3** and **4**. Each reaction was followed for 2 h by in-gel fluorescence and mobility shift.

ration of amino acids **1** and **2** (Figures 2 and 3, and Supplementary Figure 1). The yield of full length GST-CaM**1**$_1$**2**$_{40}$ was ∼2 mg per L of culture.

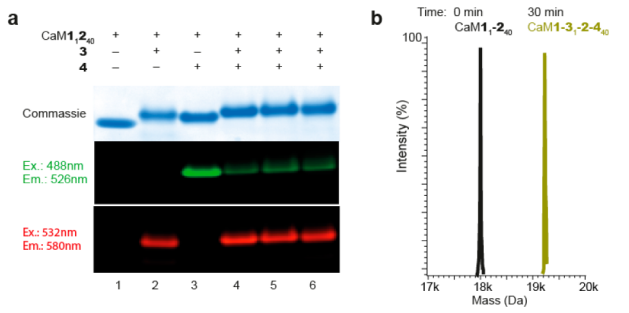


Figure 3. Concerted, quantitative, one-pot, dual labeling of calmodulin in 30 min. (a) Fluorophore-dependent labeling of CaM**1**$_1$**2**$_{40}$; sequential labeling with purification after first labeling in lane 4, sequential labeling without purification in lane 5, one-pot dual labeling in lane 6. (b) ESI-MS of one-pot protein labeling, before labeling (black, expected mass: 18000 found mass: 18000), after labeling (gold, expected mass: 19233 found mass: 19234). Raw (before deconvolution) ESI-MS spectra in Supplementary Figure 2.

To determine the time required to quantitatively label CaM**1**$_1$**2**$_{40}$ with azide **3** or tetrazine **4** we incubated 100 pmol of CaM**1**$_1$**2**$_{40}$ with 2 nmol of either **3** or **4** and followed each reaction by both mobility shift on SDS-PAGE and fluorescent imaging upon labeling (Figure 2b). These experiments demonstrate that fluorophore labeling is complete in 30 min.

Next we investigated the labeling of CaM**1**$_1$**2**$_{40}$ with both **3** and **4** (Figure 3). We first tested the addition of **4** (2 nmol) to CaM**1**$_1$**2**$_{40}$ (100 pmol) followed by purification to remove free **4**, and subsequent labeling with **3** (2 nmol) (Figure 3a lane 4). This led to efficient double labeling as judged by the SDS-PAGE mobility shift and fluorescence imaging. Next we performed sequential one-pot labeling without purification by incubating CaM**1**$_1$**2**$_{40}$ with **4** for 30 min and then adding **3** and click reagents and incubating further for 30 min (Figure 3a lane 5). This also led to efficient double labeling as judged by the SDS-PAGE mobility shift and fluorescence imaging. Finally, we simultaneously added **4** (2 nmol), **3** (2 nmol), and click reagents to CaM**1**$_1$**2**$_{40}$ (100 pmol) and incubated for 30 min (Figure 3a lane 6). This again led to efficient double labeling as judged by the SDS-PAGE mobility shift and fluorescence imaging. In all doubly labeled proteins we observe a decrease in the BODIPY-FL fluorescence relative to the singly labeled control upon excitation at 488 nm (compare lanes 4, 5 and 6 to lane 3 in Figure 3a), consistent with in-gel FRET. ESI-MS further demonstrates that this concerted, one-pot protocol leads to genetically directed efficient, rapid, and quantitative double labeling of proteins. Additional control experiments demonstrate that wild-type calmodulin is not labeled by **3** or **4** (Supplementary Figure 3), further confirming the specificity of the labeling reactions. We repeated the labeling and characterization of CaM**1**$_1$**2**$_{40}$ with **3** and **5** (Supplementary Figure 4). To further demonstrate the generality, we expressed and purified CaM**1**$_1$**2**$_{149}$ and quantitatively labeled these with **3** and **5** in 30 min, as judged by SDS PAGE and ESI-MS. Fluorescence spectra (Supplementary Figures 4 and 5) demonstrate FRET when calmodulin is labeled with donor and acceptor fluorophores at positions 1 and 40 and at positions 1 and 149, as expected.[9]

In summary, we report an efficient and rapid protocol for expressing recombinant proteins bearing a site-specifically incorporated alkyne and a site-specifically incorporated cyclopropene. We demonstrate that the inverse electron demand Diels−Alder reaction of an encoded 1,3 disubstituted cyclopropene and tetrazine probes, and the (3 + 2) cycloaddition reaction of the encoded alkyne and azide probes are mutually orthogonal to each other and to the functional groups in proteins. By combining the genetic encoding of an alkyne and a cyclopropene in a single protein and labeling with the mutually orthogonal reactions, we demonstrate the concerted, one-pot rapid double labeling of a protein in aqueous media at physiological pH and temperature. While the rate of protein labeling at specific sites in proteins may depend on local structure, sterics, and electrostatics, we anticipate that this strategy will prove useful for the double labeling of diverse proteins at diverse sites for a variety of studies. The strategy we have reported here may be extended to the double labeling of diverse molecules in cells and organisms, as well as by the use of nontoxic copper catalysts[31−33] or the development of additional bioorthogonal reactions.[34−36]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supplementary methods and details, including additional figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

chin@mrc-lmb.cam.ac.uk

Author Contributions

†A.S. and K.W. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Medical Research Council (U105181009, UD99999908) and European Research Council (MC-A024-5PG0A) for funding this work.

■ REFERENCES

(1) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906−918.
(2) Kajihara, D.; Abe, R.; Iijima, I.; Komiyama, C.; Sisido, M.; Hohsaka, T. *Nat. Methods* **2006**, *3*, 923−929.
(3) Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. *J. Am. Chem. Soc.* **2008**, *130*, 17664−17665.
(4) Nguyen, D. P.; Elliott, T.; Holt, M.; Muir, T. W.; Chin, J. W. *J. Am. Chem. Soc.* **2011**, *133*, 11418−11421.
(5) Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 6529−6540.
(6) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441−444.
(7) Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R. *Angew. Chem., Int. Ed.* **2010**, *49*, 3211−3214.
(8) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. *Biochemistry* **2013**, *52*, 1828−1837.
(9) Wang, K.; Sachdeva, A.; Cox, D. J.; Wilf, N. W.; Wallace, S.; Mehl, R. A.; Chin, J. W. *Nat. Chem.* **2014**, *6*, 393−403.
(10) Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R. *ChemBioChem.* **2012**, *13*, 1405−1408.
(11) Sasmal, P. K.; Carregal-Romero, S.; Han, A. A.; Streu, C. N.; Lin, Z.; Namikawa, K.; Elliott, S. L.; Koster, R. W.; Parak, W. J.; Meggers, E. *ChemBioChem.* **2012**, *13*, 1116−1120.
(12) Rotenberg, S. A.; Calogeropoulou, T.; Jaworski, J. S.; Weinstein, I. B.; Rideout, D. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2490−2494.
(13) Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A. *J. Am. Chem. Soc.* **2012**, *134*, 2898−2901.
(14) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. *Nat. Chem.* **2012**, *4*, 298−304.
(15) Plass, T.; Milles, S.; Koehler, C.; Szymański, J.; Mueller, R.; Wießler, M.; Schultz, C.; Lemke, E. A. *Angew. Chem., Int. Ed.* **2012**, *51*, 4166−4170.
(16) Kaya, E.; Vrabel, M.; Deiml, C.; Prill, S.; Fluxa, V. S.; Carell, T. *Angew. Chem., Int. Ed.* **2012**, *51*, 4466−4469.
(17) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192−3193.
(18) Devaraj, N. K.; Weissleder, R. *Acc. Chem. Res.* **2011**, *44*, 816−827.
(19) Yang, J.; Šečkutė, J.; Cole, C. M.; Devaraj, N. K. *Angew. Chem., Int. Ed.* **2012**, *51*, 7476−7479.
(20) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518−13519.
(21) Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 10317−10320.
(22) Borrmann, A.; Milles, S.; Plass, T.; Dommerholt, J.; Verkade, J. M. M.; Wießler, M.; Schultz, C.; van Hest, J. C. M.; van Delft, F. L.; Lemke, E. A. *ChemBioChem* **2012**, *13*, 2094−2099.
(23) Schoch, J.; Staudt, M.; Samanta, A.; Wiessler, M.; Jaschke, A. *Bioconjugate Chem.* **2012**, *23*, 1382−1386.
(24) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. *Angew. Chem., Int. Ed.* **2012**, *51*, 920−922.
(25) Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsasser, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W. *Nat. Biotechnol.* **2014**, DOI: 10.1038/nbt.2860.
(26) Kamber, D. N.; Nazarova, L. A.; Liang, Y.; Lopez, S. A.; Patterson, D. M.; Shih, H. W.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2013**, *135*, 13680−13683.
(27) Wang, K.; Schmied, W. H.; Chin, J. W. *Angew. Chem., Int. Ed. Engl.* **2012**, *51*, 2288−2297.
(28) Wang, K.; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Biotechnol.* **2007**, *25*, 770−777.
(29) Rackham, O.; Chin, J. W. *Nat. Chem. Biol.* **2005**, *1*, 159−166.
(30) Deiters, A.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1521−1524.
(31) Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2010**, *21*, 1912−1916.
(32) Uttamapinant, C.; Sanchez, M. I.; Liu, D. S.; Yao, J. Z.; Ting, A. Y. *Nat. Protoc.* **2013**, *8*, 1620−1634.
(33) Bevilacqua, V.; King, M.; Chaumontet, M.; Nothisen, M.; Gabillet, S.; Buisson, D.; Puente, C.; Wagner, A.; Taran, F. *Angew. Chem., Int. Ed. Engl.* **2014**, *32*, 465−472.
(34) Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16−20.
(35) Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764−4806.
(36) Wallace, S.; Chin, J. W. *Chem. Sci.* **2014**, *5*, 1742−1744.