In Vivo Protein Labeling

DOI: 10.1002/anie.201309847

Minimal Tags for Rapid Dual-Color Live-Cell Labeling and Super-Resolution Microscopy**

*Ivana Nikić, Tilman Plass, Oliver Schraidt, Jędrzej Szymański, John A. G. Briggs, Carsten Schultz, and Edward A. Lemke**

Abstract: *The growing demands of advanced fluorescence and super-resolution microscopy benefit from the development of small and highly photostable fluorescent probes. Techniques developed to expand the genetic code permit the residue-specific encoding of unnatural amino acids (UAAs) armed with novel clickable chemical handles into proteins in living cells. Here we present the design of new UAAs bearing strained alkene side chains that have improved biocompatibility and stability for the attachment of tetrazine-functionalized organic dyes by the inverse-electron-demand Diels–Alder cycloaddition (SPIEDAC). Furthermore, we fine-tuned the SPIEDAC click reaction to obtain an orthogonal variant for rapid protein labeling which we termed selectivity enhanced (se) SPIEDAC. seSPIEDAC and SPIEDAC were combined for the rapid labeling of live mammalian cells with two different fluorescent probes. We demonstrate the strength of our method by visualizing insulin receptors (IRs) and virus-like particles (VLPs) with dual-color super-resolution microscopy.*

The power of super-resolution microscopy (SRM) techniques heavily depends on the characteristics of the fluorophores. Most organic dyes have better photophysical properties and are typically more than 20-fold smaller than widely used fluorescent proteins.[2] With recent advances in Amber suppression technology, it is now possible to direct small, popular, and commercially available fluorophores into specific protein residues. By means of an orthogonal tRNA/tRNA synthetase pair (tRNA/RS) from *Methanosarcina mazei*, unnatural amino acids (UAAs) carrying strained alkyne and alkene side chains are genetically incorporated at positions encoded by an Amber (TAG) STOP codon.[3] These modifications add only a few atoms to the amino acid side chain and can be placed freely within the protein, lowering the risk of functional impact. Subsequently, strained alkyne and alkene UAAs can undergo catalyst-free strain-promoted alkyne–azide cycloaddition (SPAAC) and strain-promoted inverse-electron-demand Diels–Alder cycloaddition (SPIEDAC) reactions with organic fluorophores carrying azide or tetrazine (Tet) functionalities, respectively (Figure 1).[3c] The two reactions are fully biocompatible. They are additionally orthogonal to each other, since azides only react with alkynes but not with alkenes.[3c,4]

While encoding a single UAA has become relatively straightforward, there is still a demand for robust multicolor-labeling strategies in mammalian systems. At least two distinct strategies for UAA-based dual-color labeling and SRM are conceivable, which serve different experimental designs: 1) Simultaneous incorporation of two different UAAs, harboring two orthogonal functionalities (e.g. SPAAC and SPIEDAC), each recognizing a different codon in a single protein (e.g. for Förster resonance energy transfer

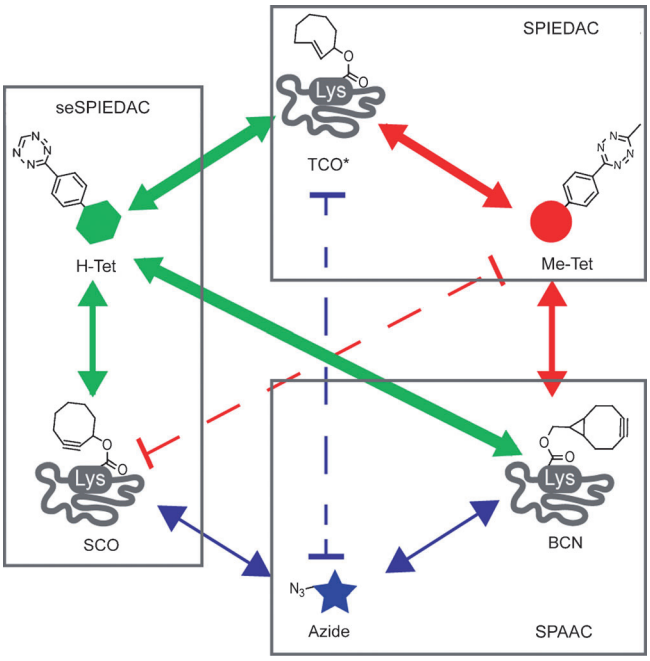

Figure 1. Reactivities of TCO*, SCO, and BCN with azide (blue), H-Tet (green), and Me-Tet (red) in SPAAC, SPIEDAC, and seSPIEDAC reactions. The thickness of arrows correlates with reaction rate (see also Figure 1 (SI) and Note 2 (SI)). Dashed lines highlight reactants that do not react with each other under the tested conditions. The three click-type reactions described in the text are in gray boxes.

[*] Dr. I. Nikić, Dr. T. Plass, Dr. O. Schraidt, Dr. J. Szymański, Dr. J. A. G. Briggs, Priv.-Doz. Dr. C. Schultz, Dr. E. A. Lemke
Structural and Computational Biology Unit
and Cell Biology and Biophysics Unit, EMBL
Meyerhofstrasse 1, 69117 Heidelberg (Germany)
E-mail: lemke@embl.de

[**] We thank S. Rizzoli for helpful discussions, N. Banterle, N. Davey, G. Estrada Girona, C. Koehler, S. Milles, S. Prinz, and S. Welsch for expert technical help, and H. G. Kraeusslich for the influenza plasmids. We thank V. VanDelinder and members of the Lemke group for proofreading this manuscript and for helpful discussions. The authors declare a conflict of interest, as some UAAs are covered by a patent application. This study was technically supported by the EMBL ALMF and PCF facilities. I.N. is an EMBO long-term fellow and T.P. a VCI fellow. O.S. is supported by a cofinanced Marie Curie/EIPOD fellowship. E.A.L. acknowledges funding by the Emmy Noether program and SPP1623 and CS from SPP1623 of the DFG.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201309847.

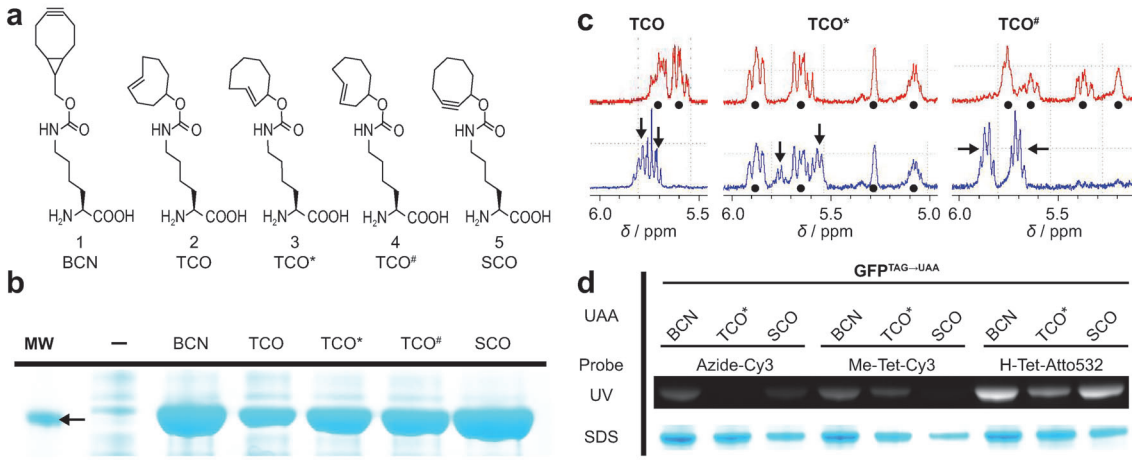

Figure 2. a) Structures of UAAs and b) the corresponding Coomassie-stained SDS-PAGE gel of purified $GFP^{TAG\rightarrow UAA}$ expressed in the absence (–) or presence of the UAAs. The arrow points to 35 kDa molecular weight marker. c) The plot shows relevant chemical shift data of TCO isomers (in the presence of cysteamine hydrochloride) measured for $t = 0$ h, room temperature in red, and for $t = 24$ h, 60 °C in blue. Black dots indicate signals of the double bond and *CH*O protons of the *trans* forms. Black arrows indicate corresponding *cis* isomer signals (for details see Figure 2 (SI)). TCO* shows the highest chemical stability (≈80% of *trans* isomer left after 24 h of thiol and heat treatment). d) Purified $GFP^{TAG\rightarrow UAA}$ (200 nM) was reacted with two tetrazines (15 μM, 20 min, 37 °C) and azide (45 μM, 10 h, 37 °C). The UV visualization and the Coomassie-stained SDS-PAGE gel are shown.

(FRET) studies) or in two different proteins (e.g. for colocalization microscopy of two different molecules). Despite progress in dual-codon suppression, in particular in prokaryotes,[5] this would require a tRNA/RS system that permits encoding a strained UAA (e.g. an alkene for SPIEDAC) simultaneously with another orthogonal tRNA/RS pair that recognizes a different distinct, but also strained UAA (e.g. an alkyne for SPAAC). This has not yet been achieved in mammalian cells. 2) Sequential encoding of two different UAAs, harboring two orthogonal functionalities, in response to the same codon using a single tRNA/RS system. This can be done in a pulse–chase manner, where the first UAA supplied to the growth medium is then chased by the second UAA. This method can be used, for example, to visualize protein sorting and transport. However, the typically employed screening of RS mutant libraries aims to find the best mutant tRNA/RS pair to incorporate one specific UAA with the highest efficiency rather than a dual-specificity tRNA/RS system.[6] Since a promiscuous tRNA/RS mutant that accepts two orthogonal UAAs with similarly high yields would make it possible to label two protein populations expressed at different times, this strategy was pursued in this work.

To develop a promiscuous system that accepts two suitable UAAs, we started from our previous observation that the double-mutant Y306A, Y384F of the *M. mazei* pyrrolysyl-tRNA-synthetase ($PylRS^{AF}$) accepts bicyclononyne-lysine (BCN) with high efficiency and *trans*-cyclooctene-lysine (TCO) with unsatisfactory yield.[3a,c] Based on the crystal structure of the $PylRS^{AF}$ mutant and the previously successfully encoded UAAs,[3,7] we hypothesized that screening the various ring-substitution isomers could yield a TCO variant that would be accepted as well as BCN. As shown in Figure 2a we synthesized two new ring isomers (Note 1 in the Supporting Information (SI)): the *trans*-cyclooct-2-ene TCO* and the *trans*-cyclooct-3-ene $TCO^{\#}$. Test expressions with a GFP^{TAG} reporter construct, which only gives full-length expression and hence fluorescence when the Amber mutation at position Y39 is suppressed, show that TCO* and $TCO^{\#}$ are accepted by the $tRNA^{Pyl}/PylRS^{AF}$ pair approximately three times better than the original cyclooct-4-ene TCO (Figure 2b, see Table S1 (SI) for mass spectrometry data), yielding roughly 10 mg of the $GFP^{TAG\rightarrow UAS}$ protein from 1 L of *E. coli* expression culture. Figure 2 (SI) shows that the three TCOs maintain similar reactivities in SPIEDAC reactions. However, *trans*-cyclooctenes as in TCO tends to isomerize to the nonreactive *cis* form especially in the presence of thiols.[8] Since thiols are abundant in the cytosol of mammalian cells, this can lead to interference with UAA biostability during long-term expressions. NMR measurements shown in Figure 2c and Figure 2 (SI) showed that TCO* is at least ten times more stable than TCO and $TCO^{\#}$ in the presence of thiols, which indicated efficient shielding of the *trans* double bond towards nucleophilic attack through the proximity of the carbamate bond. Thus only TCO* was used in the following experiments.

The $tRNA^{Pyl}/PylRS^{AF}$ mutant pair permits encoding TCO* and BCN, which can undergo SPIEDAC and SPAAC reactions, respectively (Figure 2d, Figure 1). To explore the potential of this UAA pair for dual-color labeling of live cells, we used it for the pulse–chase labeling of the insulin receptor (IR). The function of the IR and its recycling are topics of high contemporary relevance due to its central role in diabetes and recently discovered roles in gene regulation.[9] We selected a position located on the extracellular side of the protein (K676) for an Amber (TAG) mutation and expressed the IR^{TAG} in the presence of a plasmid coding for the $tRNA^{Pyl}/PylRS^{AF}$ in HEK293T cells.

We then performed the pulse–chase experiment as outlined in Figure 3a, where the growth medium was first pulsed

for 4 h with the first UAA (TCO*) which was chased with 4 h of incubation with the second UAA (BCN). Living cells expressing IR^{TAG} were labeled by incubation with first azide-Cy5 (10 min) and then with Methyl-Tet-Cy3 (Me-Tet-Cy3) (10 min). As can be seen in Figure 3b (first panel), confocal

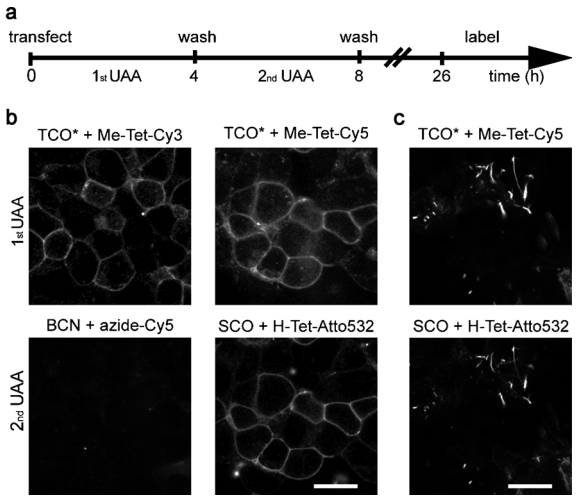

Figure 3. a) Outline of the expression and labeling scheme employed for dual-color labeling of IR^{TAG}. b) Confocal images of dual-color labeling of IR with different combinations of UAAs and dyes. The left panels show a combination of SPIEDAC (TCO*/Me-Tet-Cy3 and SPAAC labeling (BCN/azide-Cy5). The right panels show a combination of SPIEDAC (TCO*/Me-Tet-Cy5) and seSPIEDAC (SCO/H-Tet-Atto532). c) Virus-like particles (VLPs) visualized by dual-color labeling achieved by SPIEDAC (TCO*/Me-Tet-Cy5; top) and seSPIEDAC (SCO/H-Tet-Atto532; bottom). Scale bars are 20 μm.

imaging was used to visualize the membrane staining of $IR^{TAG\rightarrow TCO^*}$ with Me-Tet-Cy3 from the SPIEDAC reaction. The short incubation of $IR^{TAG\rightarrow BCN}$ with azide-Cy5 gave no results. As detailed in Figure 3 (SI), this could be explained by the speed of the SPAAC reaction, which is three to four orders of magnitudes slower than the SPIEDAC reaction. We therefore set out to find an alternative fast orthogonal click reaction.

We observed previously that tetrazines can react with strained alkene and alkyne UAAs (Figure 1).[3a,c] Due to the markedly different reaction properties of strained alkynes and alkenes,[4a] we reasoned that it might be possible to identify tetrazine-UAA combinations that permit orthogonal labeling. The previously developed cyclooctynyl-lysine derivative (SCO) is accepted by the same $tRNA^{Pyl}/PylRS^{AF}$ pair and in similar yields as TCO* (Figure 2a,b).[3b] While TCO* reacts with H-Tet and Me-Tet with reactions rates exceeding $1000\,M^{-1}s^{-1}$ in in vitro kinetic assays and labeling experiments, SCO shows no substantial reactivity with Me-Tet under the tested conditions (see Figure 1 (SI) for reaction kinetics; Figure 2d and Figure 1). However, the SPIEDAC reaction of SCO with H-Tet is still about two orders of magnitude faster than the SPAAC reaction of BCN with azide.[3a,10] We repeated the pulse–chase experiment with TCO* and SCO (Figure 3a), followed by labeling with Me-Tet-Cy5, and then H-Tet-Atto532 for 10 min. As shown in Figure 3b, this combination resulted in the bright labeling of the IR in the plasma membrane for both channels. SCO selectively reacts with H-Tet but not Me-Tet on the timescale of our experiments and thus this reaction is orthogonal to the SPIEDAC between TCO* and Me-Tet. We dub this subreaction type selectivity-enhanced SPIEDAC (seSPIEDAC). Note that as TCO* is highly reactive with both Me-Tet and H-Tet, experimental conditions must be chosen to ensure that all TCO* is consumed before proceeding to the second labeling step (see Figure 4 (SI) for details). Furthermore, we show in Figure 2d and Figure 4 (SI) that increasing the speed of seSPIEDAC by using the highly ring-strained BCN instead of SCO is not possible due to the reactivity of BCN towards Me-Tet (Figure 1 (SI) and Note 2 (SI); see also Figure 1).

Since Cy5 and Atto532 are commonly used for localization-based SRM, we performed dual-color SRM measurements.[11] The confocal (Figure 3) and widefield images showed overlapping plasma membrane staining of IR in both colors after dual-color labeling the TCO* and SCO. However, SRM revealed a heterogeneous distribution of IR clusters at the membrane (Figure 4a). Notably, clustering of other growth factor receptors has also been observed in previous SRM studies.[12]

To demonstrate the generality of our approach, we labeled virus-like particles (VLPs) assembled by the co-expression of influenza virus proteins hemagglutinin (HA) and matrix protein 1 (M1) (for a review see Ref. [13]). Viral genomes are compact and often contain overlapping genes, which makes inserting genetically encoded tags into viral proteins a particular challenge. We generated a TAG mutant of HA and expressed it together with M1 and the $tRNA^{Pyl}/PylRS^{AF}$ in HEK293T cells. We repeated the pulse–chase protocol using TCO*, SCO as UAAs and Me-Tet-Cy5 and H-Tet-Atto532 as fluorescence labels. As shown in Figure 3c, Atto532- and Cy5-stained filamentous protrusions, corresponding to assembled VLPs, became visible. The enhanced resolution of SRM makes it possible to visualize individual filaments (Figure 4b). Notably, there is significant spatial overlap between the two colors, suggesting that proteins translated at different times are incorporated into the same assembling VLPs.

In summary, we have tuned the genetically encoded SPIEDAC reaction into two mutually orthogonal SPIEDAC reactions which can be used to perform rapid labeling of proteins in living cells. In this sense, this expands the existing repertoire of biocompatible click labeling methods using an expanded genetic code from SPIEDAC and SPAAC to SPIEDAC, seSPIEDAC, and SPAAC. The exact origin of the orthogonality achieved between seSPIEDAC and SPIEDAC is still subject to further investigations, such as theoretical calculations.[4a] The two rapid SPIEDAC reactions allowed SRM-compatible dual-color-labeling experiments in mammalian cells, while the slow reactivity of SPAAC seemed insufficient for rapid high-contrast labeling of live cells.

TCO* offers higher biostability and incorporation efficiency than the original TCO. TCO* reacts rapidly with both tested tetrazines (Me-Tet, H-Tet), while BCN is much less reactive with Me-Tet. Under the experimental conditions, SCO reacted with H-Tet only and not with Me-Tet. Dual-color labeling was achieved using a promiscuous tRNA/RS

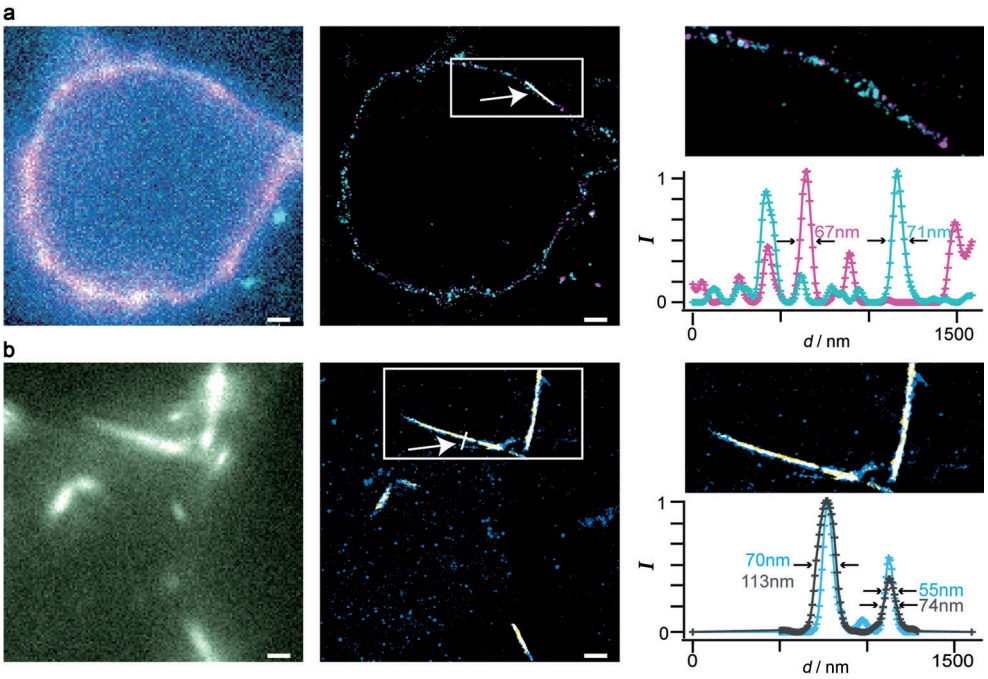

Figure 4. SRM images of cells expressing IRs and VLPs after SPIEDAC and seSPIEDAC labeling. a) Widefield (left) and SRM (middle) images of IR^{TAG} labeled according to Figure 3 a,b (Atto532 in magenta, Cy5 in cyan). Right: inset from the middle panel and a line plot (across the line shown in middle panel, which is highlighted by an arrow). The width of the marked peaks is given as full width at half maximum (FWHM). b) Widefield (left) and SRM (middle) images of VLPs analoguous to (a). In the images, Atto532 is in blue and Cy5 in yellow; overlaid regions appear white. In the line profile Atto532 is shown in blue and Cy5 in dark gray. SRM images are displayed at a resolution of 45 nm as determined by Fourier Ring correlation[1] (FRC, see the Supporting Information). Scale bars are 1 μm.

pair and a pulse–chase approach. The labeling step is done in living cells, creating new possibilities for studying protein functions with very high resolution. By combining our dual-color labeling with genetic switches, such as temperature-sensitive mutants and promoter control, it could be possible to label distinct proteins. In addition, the labeling methods developed here are general and can also be directly applied if specific encoding by two distinct codons becomes available in the future. For this, however, two orthogonal RS must be developed permitting specific encoding of the strained side chains necessary for seSPIEDAC and SPIEDAC.

The small size of the UAA tag, in comparison with other genetically encoded fluorescent tags, is a major advantage especially for studies of complex protein assemblies such as IR and HA, where multiple functional interactions with other proteins and lipids might be influenced by larger tags in unpredictable ways. In particular, viral genomes are frequently extremely compact and do not tolerate large modifications. The need for changing only a single codon, thus dramatically increases the chance of finding permissive sites that do not alter protein function.

Protein labeling through the expansion of the genetic code is in general a two-step process. Step 1 is the successful encoding of the reactive UAA in vivo, and thus proteins can be targeted anywhere in a cell. Step 2 is the coupling of the fluorescent dye to the protein. Most fluorescent dyes that are able to cross the plasma membrane are also hydrophobic and tend to stick nonspecifically to cellular structures. Here we bypassed this issue by limiting ourselves to proteins that are accessible to cell-surface labeling with charged hydrophilic dyes, such as sulfonated compounds, which cannot penetrate the plasma membrane in vivo. Presently, we are witnessing the burgeoning development of new dyes,[14] and improved fluorogenic dyes that give little to no background signals in living cells will soon become widely available. Since the described technique relies on the generic ligation mechanism of two tuned SPIEDAC reactions, it will thus be compatible with any new dyes suitable for live intracellular labeling and also applicable to a broad range of other disciplines relying on tags, such as MRI and PET studies.

Received: November 12, 2013
Published online: January 28, 2014

Keywords: amino acids · click chemistry · cycloaddition · protein engineering · super-resolution microscopy

[1] a) N. Banterle, K. H. Bui, E. A. Lemke, M. Beck, *J. Struct. Biol.* **2013**, *0*, 0; b) R. P. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grunwald, S. Stallinga, B. Rieger, *Nat. Methods* **2013**, *10*, 557–562.

[2] S. van de Linde, M. Heilemann, M. Sauer, *Annu. Rev. Phys. Chem.* **2012**, *63*, 519–540.

[3] a) A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. Verkade, M. Wiessler, C. Schultz, J. C. van Hest, F. L. van Delft, E. A. Lemke, *ChemBioChem* **2012**, *13*, 2094–2099; b) T. Plass, S. Milles, C. Koehler, C. Schultz, E. A. Lemke, *Angew. Chem.* **2011**, *123*, 3964–3967; *Angew. Chem. Int. Ed.* **2011**, *50*, 3878–3881; c) T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz, E. A. Lemke, *Angew. Chem.* **2012**, *124*, 4242–4246; *Angew. Chem. Int. Ed.* **2012**, *51*, 4166–4170; d) K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox, J. W. Chin, *J. Am. Chem. Soc.* **2012**, *134*, 10317–10320; e) S. Schneider, M. J. Gattner, M. Vrabel, V. Flugel, V. Lopez-Carrillo, S. Prill, T. Carell, *ChemBioChem* **2013**, *14*, 2114–2118.

[4] a) Y. Liang, J. L. Mackey, S. A. Lopez, F. Liu, K. N. Houk, *J. Am. Chem. Soc.* **2012**, *134*, 17904–17907; b) M. R. Karver, R. Weissleder, S. A. Hilderbrand, *Angew. Chem.* **2012**, *124*, 944–946; *Angew. Chem. Int. Ed.* **2012**, *51*, 920–922.

[5] a) W. Wan, Y. Huang, Z. Y. Wang, W. K. Russell, P. J. Pai, D. H. Russell, W. R. Liu, *Angew. Chem.* **2010**, *122*, 3279–3282; *Angew. Chem. Int. Ed.* **2010**, *49*, 3211–3214; b) H. Neumann, K. H. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, *464*,

441 – 444; c) W. Niu, P. G. Schultz, J. Guo, *ACS Chem. Biol.* **2013**, *8*, 1640 – 1645.
[6] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413 – 444.
[7] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187 – 1197.
[8] J. Yang, J. Seckute, C. M. Cole, N. K. Devaraj, *Angew. Chem.* **2012**, *124*, 7594 – 7597; *Angew. Chem. Int. Ed.* **2012**, *51*, 7476 – 7479.
[9] a) K. Siddle, *J. Mol. Endocrinol.* **2011**, *47*, R1 – 10; b) R. Sarfstein, H. Werner, *Endocrinology* **2013**, *154*, 1672 – 1679.
[10] M. F. Debets, S. S. van Berkel, J. Dommerholt, A. T. Dirks, F. P. Rutjes, F. L. van Delft, *Acc. Chem. Res.* **2011**, *44*, 805 – 815.
[11] a) M. Heilemann, S. van de Linde, M. Schuttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, M. Sauer, *Angew. Chem.* **2008**, *120*, 6266 – 6271; *Angew. Chem. Int. Ed.* **2008**, *47*, 6172 – 6176; b) M. Bates, B. Huang, G. T. Dempsey, X. Zhuang, *Science* **2007**, *317*, 1749 – 1753; c) J. Fölling, M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling, S. W. Hell, *Nat. Methods* **2008**, *5*, 943 – 945.
[12] S. Wilmes, M. Staufenbiel, D. Lisse, C. P. Richter, O. Beutel, K. B. Busch, S. T. Hess, J. Piehler, *Angew. Chem.* **2012**, *124*, 4952 – 4955; *Angew. Chem. Int. Ed.* **2012**, *51*, 4868 – 4871.
[13] J. S. Rossman, R. A. Lamb, *Virology* **2011**, *411*, 229 – 236.
[14] a) G. Lukinavičius, K. Umezawa, N. Olivier, A. Honigmann, G. Yang, T. Plass, V. Mueller, L. Reymond, I. R. Corrêa, Jr., Z. G. Luo, C. Schultz, E. A. Lemke, P. Heppenstall, C. Eggeling, S. Manley, K. Johnsson, *Nat. Chem.* **2013**, *5*, 132 – 139; b) J. C. Carlson, L. G. Meimetis, S. A. Hilderbrand, R. Weissleder, *Angew. Chem.* **2013**, *125*, 7055 – 7058; *Angew. Chem. Int. Ed.* **2013**, *52*, 6917 – 6920.

