

## Genetic Code Expansion Enables Live-Cell and Super-Resolution Imaging of Site-Specifically Labeled Cellular Proteins

Chayasith Uttamapinant, Jonathan D. Howe, Kathrin Lang,<sup>†</sup> Václav Beránek, Lloyd Davis, Mohan Mahesh, Nicholas P. Barry,<sup>\*</sup> and Jason W. Chin<sup>\*</sup>

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

### Supporting Information

**ABSTRACT:** Methods to site-specifically and densely label proteins in cellular ultrastructures with small, bright, and photostable fluorophores would substantially advance super-resolution imaging. Recent advances in genetic code expansion and bioorthogonal chemistry have enabled the site-specific labeling of proteins. However, the efficient incorporation of unnatural amino acids into proteins and the specific, fluorescent labeling of the intracellular ultrastructures they form for subdiffraction imaging has not been accomplished. Two challenges have limited progress in this area: (i) the low efficiency of unnatural amino acid incorporation that limits labeling density and therefore spatial resolution and (ii) the uncharacterized specificity of intracellular labeling that will define signal-to-noise, and ultimately resolution, in imaging. Here we demonstrate the efficient production of cytoskeletal proteins ( $\beta$ -actin and vimentin) containing bicyclo[6.1.0]nonyne-lysine at genetically defined sites. We demonstrate their selective fluorescent labeling with respect to the proteome of living cells using tetrazine-fluorophore conjugates, creating densely labeled cytoskeletal ultrastructures. STORM imaging of these densely labeled ultrastructures reveals subdiffraction features, including nuclear actin filaments. This work enables the site-specific, live-cell, fluorescent labeling of intracellular proteins at high density for super-resolution imaging of ultrastructural features within cells.

Super-resolution microscopic techniques such as STORM/PALM,<sup>1,2</sup> STED,<sup>3</sup> RESOLFT,<sup>4</sup> and SIM<sup>5</sup> have transformed biological imaging by allowing subdiffraction visualization of biomolecules in living cells. However, the performance of these techniques is limited by the photophysical properties and size of commonly used fluorescent labels. Fluorescent proteins (FPs) give off low photon numbers before bleaching,<sup>6</sup> thereby limiting the localization precision of individual fluorophores and lowering the image resolution. The low photon output from FPs is especially detrimental to live-cell superresolution imaging, where there is an intrinsic trade-off between spatial and temporal resolution, and hence a need to collect as many photons/localizations as possible in a given imaging period. While organic fluorophores have better photophysical properties<sup>7</sup> than FPs, methods to attach fluorophores to proteins in cells, including SNAP/CLIP,<sup>8</sup> DHFR,<sup>9</sup> HaloTag,<sup>10</sup> FAsH,<sup>11</sup> and PRIME,<sup>12</sup> involve the addition of a protein or peptide appendage (size

ranging from 12 to 297 amino acids) which can disrupt the structure, function, and localization of the protein of interest, may limit the sites that can be labeled to protein termini and, may ultimately limit the extent to which the dimensions of native protein assemblies can be defined by imaging.

Genetic code expansion via the use of an orthogonal aminoacyl-tRNA synthetase/tRNA pair, which allows substitution of a single amino acid on a protein with a designer unnatural amino acid, is the only site-specific protein labeling method that does not require the addition of protein or peptide bulk. The unnatural amino acid is directed to a desired incorporation site in response to the amber stop codon (UAG). Among different orthogonal tRNA synthetase/tRNA pairs used for incorporation of unnatural amino acids, the pyrrolysyl-tRNA synthetase (PylRS)/Pyl tRNA<sub>CUA</sub> pair from *Methanosarcina* species is preferred for two main reasons: (i) the active site of PylRS and its engineered variants efficiently accept structurally diverse unnatural amino acids as substrates, but not the common 20 amino acids;<sup>13</sup> and (ii) the PylRS/Pyl tRNA<sub>CUA</sub> pair is orthogonal in a range of hosts including *E. coli*,<sup>14</sup> yeast,<sup>15</sup> mammalian cells,<sup>16–18</sup> *C. elegans*,<sup>19</sup> and *D. melanogaster*,<sup>20</sup> facilitating the transposition of PylRS variants evolved in *E. coli* for use in other hosts.

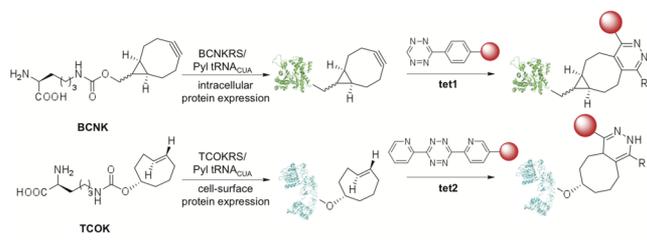
The installation of probes for super-resolution imaging via genetic code expansion is best achieved using a two-step strategy. In a first step an amino acid bearing a bioorthogonal group is co-translationally installed in the protein, and in a second step the probe is site-specifically attached to the protein. This strategy has several advantages: (i) it provides a modular approach for installing diverse probes with a single genetic system and, (ii) it removes limitations that the translational machinery may place on the size of probes that may be used. Bicyclo[6.1.0]nonyne-lysine and *trans*-cyclooctene-lysine (BCNK and TCOK respectively, structures in Scheme 1) can be genetically encoded using variants of the PylRS/Pyl tRNA<sub>CUA</sub> pair and are the best amino acid candidates for protein labeling in cells. These amino acids react chemoselectively with tetrazine conjugates via an inverse-electron-demand Diels–Alder cycloaddition.<sup>21</sup> Such cycloadditions are exceptionally fast (second-order rate constants in the range of 10<sup>2</sup>–10<sup>4</sup> M<sup>-1</sup> s<sup>-1</sup> have been reported *in vitro*),<sup>22</sup> are compatible with living cells,<sup>21,23,24</sup> and can be fluorogenic if the tetrazine is linked to a suitable fluorophore.

Despite the excellent properties of the inverse-electron-demand Diels–Alder reaction, previous uses of genetic code

Received: December 17, 2014

Published: April 1, 2015

### Scheme 1. Fluorophore Labeling of Cellular Proteins via Genetic Encoding of Unnatural Amino Acids and Diels–Alder Cycloaddition



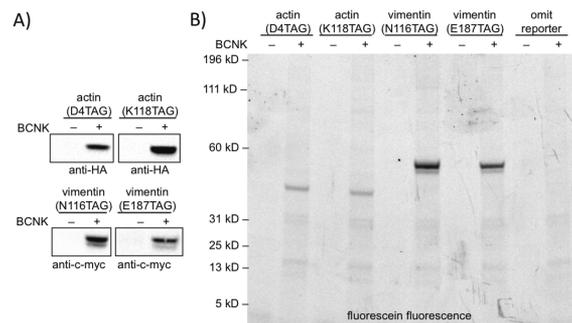
expansion in conjunction with Diels–Alder reactions for protein labeling have been limited to cell-surface proteins,<sup>25</sup> with one example of live-cell intracellular labeling on a nuclear transcription factor, jun (reported by our laboratory).<sup>21</sup> One study has attempted super-resolution imaging,<sup>26</sup> using sparsely labeled cell-surface proteins. However, one difficulty in applying genetic code expansion to protein imaging stems from the inefficiency of unnatural amino acid incorporation in mammalian cells.<sup>27</sup> This presents a substantial challenge for sub-diffraction imaging, where the distance between fluorophores contributes to the image resolution (Nyquist criterion) and densely labeled ultrastructures are required.

We recently optimized the PylRS/Pyl tRNA<sub>CUA</sub> pair for unnatural amino acid incorporation in mammalian cells by increasing Pyl tRNA<sub>CUA</sub> levels.<sup>27</sup> Here, we show that this enables efficient unnatural amino acid incorporation and subsequent dense fluorescent labeling of intracellular proteins. To demonstrate the utility of genetic code expansion for live-cell protein imaging, we incorporate TCOK and BCNK into epidermal growth factor receptor (EGFR) (Scheme 1). We tackle the more difficult, and previously unaddressed, challenge of live-cell, site-specific labeling of intracellular proteins for subdiffraction imaging by maximizing specific labeling signal and dissecting and minimizing sources of background. We incorporate BCNK into actin microfilaments and vimentin intermediate filaments and label the proteins in living cells with tetrazine-fluorophores (Scheme 1). Super-resolution imaging of these membrane and cytoskeletal proteins reveals fine structural details, including those of nuclear actin filaments, previously unresolved by traditional fluorescence microscopy.

PylRS mutants specific for TCOK- and BCNK-amino acids (TCOKRS and BCNKRS, respectively) with good incorporation efficiencies have been discovered.<sup>21</sup> While TCOK is an order of magnitude more reactive than BCNK in inverse-electron-demand Diels–Alder reactions, it is challenging to remove excess TCOK from cells by rinsing cells with fresh media (Figure S1). This leads to a high labeling background upon addition of a tetrazine-fluorophore. In contrast, BCNK can be more efficiently removed from cells. These experiments suggest BCNK will be more useful for intracellular labeling, though we note that TCOK is still useful for cell-surface protein labeling when a membrane-impermeable tetrazine-fluorophore is employed.

To maximize the expression of the BCNKRS/Pyl tRNA<sub>CUA</sub> pair in mammalian cells for intracellular labeling, we created a plasmid construct bearing a single copy of BCNKRS on an EF1 $\alpha$  constitutive promoter, and four copies of PylT (the gene encoding Pyl tRNA<sub>CUA</sub>) each on a U6 promoter.<sup>27</sup> To further ensure high levels of Pyl tRNA<sub>CUA</sub> in our expression system, we included four copies of U6-PylT in the expression plasmid for the protein of interest (POI)— $\beta$ -actin or vimentin—bearing the

amber codon. As there are no previous reports of site-specific unnatural amino acid incorporation in actin or vimentin, we selected several amino acid positions that are surface-exposed and not involved in interstrand contacts based on available structural information on actin<sup>28,29</sup> and vimentin,<sup>30</sup> to be replaced with BCNK. Upon cotransfection of BCNKRS/PylT and POI/PylT plasmids and subsequent addition of BCNK, we observed efficient, BCNK-dependent expression of two amber variants of actin, D4TAG and K118TAG, as well as two amber variants of vimentin, N116TAG and E187TAG, by western blot (Figures 1A and S2). The expression of these recombinant



**Figure 1.** Efficient BCNK-dependent expression of cytoskeletal proteins and specific labeling with a tetrazine-fluorophore. (A) HEK293T cells were transfected with plasmids encoding (Pyl tRNA<sub>CUA</sub>)<sub>4</sub>/BCNKRS and (Pyl tRNA<sub>CUA</sub>)<sub>4</sub>/POI (POI = protein of interest = actin<sup>D4TAG</sup> or K118TAG, vimentin<sup>N116TAG</sup> or E187TAG) and grown in the presence or absence of 1 mM BCNK for 48 h. Actin was HA-tagged, while vimentin was c-myc-tagged. (B) Analyses of labeling specificity. HEK cells were transfected and grown with/without BCNK as in (A). After removal of excess BCNK, cells were treated with 400 nM tet1-CFDA for 20 min, then washed for 2 h before lysis. Lysates were resolved and their fluorescence visualized. Cells were transfected with plasmids encoding (Pyl tRNA<sub>CUA</sub>)<sub>4</sub>/BCNKRS and (Pyl tRNA<sub>CUA</sub>)<sub>4</sub> in the omit reporter controls. Loading controls for both (A) and (B) are shown in Figure S5.

proteins is greatly improved in the new expression cassette, with respect to a previously reported system with much lower Pyl tRNA<sub>CUA</sub> levels (Figure S3).<sup>21</sup>

We next tested whether actin and vimentin bearing BCNK can be specifically derivatized inside living cells with a tetrazine-fluorophore conjugate via Diels–Alder cycloaddition. Among tetrazine variants, we chose a 3-phenyl-1,2,4,5-tetrazine (tet1, Scheme 1 and Figure S4) as it has excellent cell entry and conjugated it to a membrane-permeable diacetylfluorescein (CFDA) fluorophore. To assess labeling specificity, we expressed amber variants of actin and vimentin in the presence of BCNK and, after excess BCNK was removed, added tet1-CFDA to live cells to effect the cycloaddition. Analyses of cell lysates by in-gel fluorescence imaging show strong fluorescent bands corresponding to the molecular weights of actin (42 kDa) or vimentin (58 kDa, Figures 1B, S5 and S6). There was minimal off-target labeling of the proteome in the presence of BCNK, suggesting that BCNK incorporation in response to the amber termination codons of endogenous genes is minimal in comparison to specific BCNK incorporation on the desired protein target—even though the target proteins are expressed at levels at or below those of other proteins present in the cell. This observation may reflect i) a distinction between the decoding of an amber codon presented in the middle of a coding sequence and the decoding of an amber codon presented in the context of a

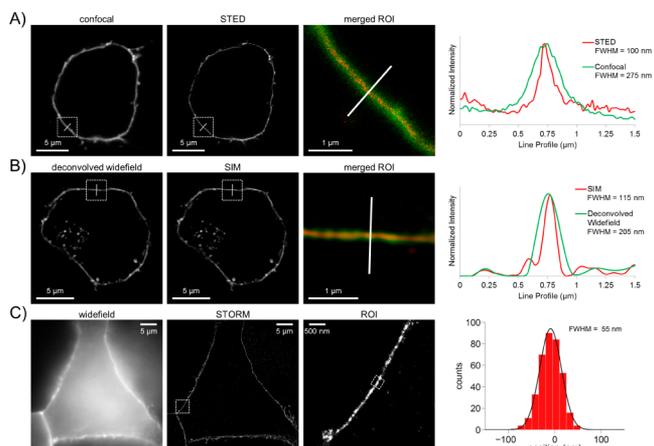
natural termination site at the 3' end of an mRNA and/or ii) the degradation of aberrantly extended proteins that result from the read-through of endogenous stop codons. No off-target proteome labeling was observed in the absence of BCNK, indicating that the chemoselectivity of tet1-CFDA toward BCNK is excellent. Variants of actin and vimentin were fluorescently tagged to a similar extent regardless of the amber codon position, suggesting that numerous amino acid sites within a protein that are surface-exposed and not functionally important may be targeted for replacement by BCNK.

While we did not observe significant non-specific labeling on the proteome, we saw evidence of off-target fluorophores in the nuclei of a subset of cells (Figure S7). This nuclear background is specific to transfected cells and is most evident when the target protein is exclusively cytosolic (e.g., vimentin) and could not contribute to labeling signal in the nuclei. Fluorescence *in situ* hybridization demonstrates that the nuclear labeling colocalizes with Pyl tRNA<sub>CUA</sub> (Figure S7), consistent with the background resulting from a population of nuclear Pyl tRNA<sub>CUA</sub> acylated with BCNK being derivatized with tetrazine probes. This is consistent with reports that the aminoacylation of nuclear tRNAs acts as a quality control mechanism for nuclear export.<sup>31</sup> The nuclear background is reduced if HEK cells are extensively washed before the tetrazine-fluorophore is added.

We next synthesized the tetrazine-fluorophore conjugates for various super-resolution imaging applications (Figure S4). For cell-surface labeling, we conjugated 3,6-dipyridyl-1,2,4,5-tetrazine based compound (tet2), a kinetically efficient tetrazine, to ATTO488 and Alexa Fluor 647. For intracellular labeling, we coupled tet1 to a membrane-permeable near-infrared silicon-rhodamine (SiR) fluorophore.<sup>32</sup>

We proceeded to perform cellular labeling and super-resolution imaging using unnatural amino acids and tetrazine-fluorophores. At the cell surface, we expressed EGFR bearing an amber codon at position 128 (EGFR128TAG) in HEK cells containing the BCNKRS/Pyl tRNA<sub>CUA</sub> and grown in the presence of BCNK. Treatment of cells with tet2-ATTO488 resulted in green fluorescent rings at the cell membrane, consistent with specific labeling on EGFR (Figure S8). After fixing the labeled cells, STED and SIM imaging produced superior image resolution (100 and 115 nm, respectively, based on line profile analyses) compared to conventional diffraction-limited imaging (205–275 nm, Figure 2A,B). Likewise, EGFR128TAG-GFP fusion expressed in HEK cells containing the TCOKRS/Pyl tRNA<sub>CUA</sub> pair grown in the presence of TCOK was labeled with tet2-Alexa Fluor 647. Alexa Fluor 647 labeling was observed only in GFP-positive cells, confirming labeling specificity (Figure S8). After fixing the labeled cells, STORM imaging of EGFR showed a significant improvement in the resolution compared to conventional widefield fluorescence imaging (Figure 2C). We performed a Gaussian fit on the cross-sectional distributions of localizations on a thin section of the cell membrane and found a full width at half-maximum (FWHM) of 55 nm, a value that is well below the diffraction limit.

We also labeled the intracellular protein targets actin and vimentin in live COS-7 cells with tet1-SiR and after fixation, confirmed labeling specificity (Figure S8), and performed STORM imaging (Figure 3). In the presence of BCNK, we observed production of actin and vimentin filaments, indicating that our tagged proteins are polymerization competent, and subsequent tet1-SiR labeling revealed the expected localizations for both proteins. Upon STORM imaging, different populations of vimentin intermediate filaments can be seen, including thick

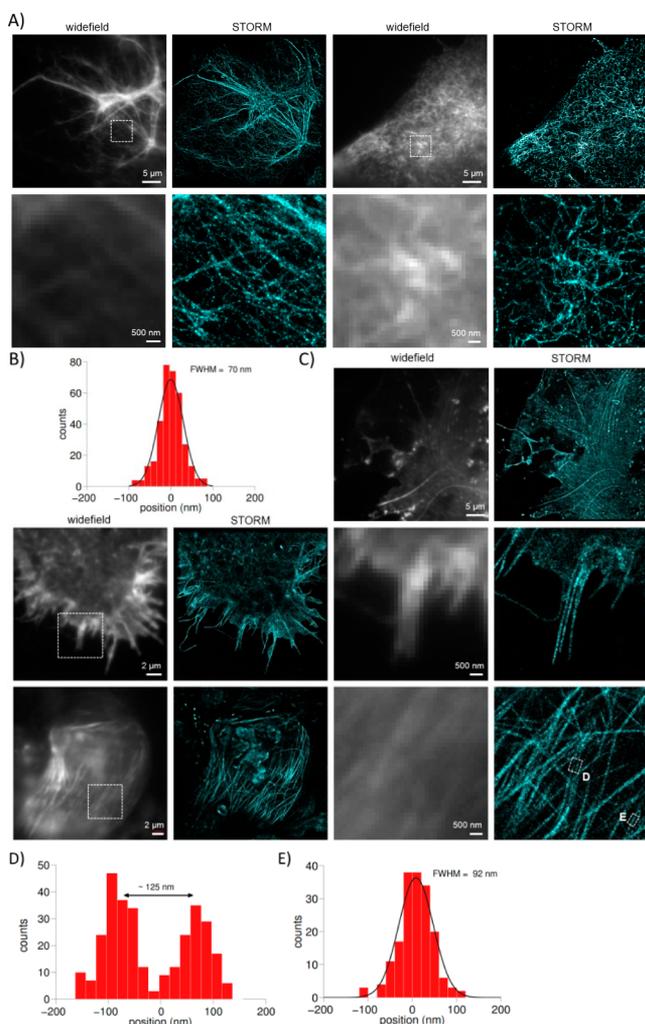


**Figure 2.** STED, SIM, and STORM imaging of EGFR in HEK293T cells. (A) Confocal image of EGFR labeled with BCNK and tet2-ATTO488 and corresponding STED image of the same cell. Line-profile analysis of the indicated membrane segment shows resolution improvement of STED over confocal microscopy. (B) Deconvolved widefield fluorescence image of EGFR labeled as in (A), corresponding SIM image of the same cell and line-profile analysis of the indicated membrane segment. (C) Widefield fluorescence image of EGFR-GFP labeled with TCOK and tet2-Alexa Fluor 647, corresponding STORM image of the same cell and cross-sectional analysis of the indicated membrane segment. ROI = region of interest.

vimentin bundles and thinner structures (Figure 3A). We measured the apparent width of an individual thinner vimentin structure in the STORM image (Figure S9) by performing a Gaussian fit on its cross-sectional distribution of localizations and found the FWHM to be 70 nm (Figure 3B; our mean localization precision is  $\sim 7$ –13 nm, Figure S11). The measured width of the thinner vimentin structure may be attributed to tightly packed vimentin filament bundles, which are observed by electron microscopy of vimentin in cells.<sup>33</sup> The resolution of the images, as judged by Fourier ring correlation (FRC) analysis,<sup>34</sup> was comparable to those of endogenous filamentous proteins labeled by standard methods in previous reports (Figure S12).

For actin, we observed diffuse globular actin, along with distinct classes of filamentous actin: stress fibers and actin meshworks throughout the cytosol (top row, Figure 3C), cortical actin at the plasma membrane (middle row), and actin filaments in the nucleus (bottom row). Upon STORM imaging, these ultrastructural features are further refined. For example, actin filaments that are 125 nm apart can now be clearly separated (Figure 3D, see Figure S12 for FRC analysis). Interestingly, we observed uniform,  $\sim 92$  nm-wide filaments (Figure 3E) within the cell nuclei of a small population of cells (co-staining of labeled actin with a nuclear marker shown in Figure S13). Actin is known to exist as monomers or short oligomers within the nucleus,<sup>35</sup> so the long filamentous actin we observe may be a result of ectopic expression conditions, or linked to known functions of filamentous actin in the cell nuclei, such as control of cell<sup>36</sup> and nucleus morphology<sup>37</sup> or mediation of gene expression.<sup>36</sup>

In summary, we have developed a protein labeling strategy based on the genetic encoding of BCNK and TCOK and inverse-electron-demand Diels–Alder cycloaddition for live-cell and super-resolution imaging. With the optimized PylRS/Pyl tRNA<sub>CUA</sub> system and labeling protocol, we were able to densely express and fluorescently label cell-surface and cytoskeletal proteins in living cells. In conjunction with advances in fluorophore design, we have demonstrated, to our knowledge,



**Figure 3.** STORM imaging of vimentin (N116TAG) and actin (K118TAG) in COS-7 cells. (A) Widefield fluorescence images of vimentin labeled with BCNK and tet1-SiR and corresponding STORM images. Images in the second row are zoom-ins of the boxed regions in the first row. (B) Cross-sectional distribution of localizations along a vimentin filament segment from a STORM image (Figure S9) and its Gaussian fit, which gives FWHM of 70 nm. (C) Widefield images of actin labeled with BCNK and tet1-SiR and corresponding STORM images. Third and fourth columns are zoom-in images of the boxed regions in the first and second columns, respectively. (D) Cross-sectional distribution of localizations along two nearby actin filaments in box D. The two filaments are separated by  $\sim 125$  nm. (E) Cross-sectional distribution of localizations along a nuclear actin filament segment from box E and its corresponding Gaussian fit, which gives FWHM of 92 nm. Larger versions of all images are shown in Figure S10.

the first application of genetic code expansion to super-resolution imaging of mammalian intracellular proteins. The flexibility in labeling site placement upon using genetic code expansion in combination with super-resolution imaging should allow for creative applications of the method, such as delineation of macromolecular organization and architecture<sup>38</sup> in living cells.

## ■ ASSOCIATED CONTENT

### Supporting Information

Complete references, supplementary figures, and methods. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*chin@mrc-lmb.cam.ac.uk

\*nbarry@mrc-lmb.cam.ac.uk

### Present Address

†Department of Chemistry, Technische Universität München, Institute for Advanced Study, Lichtenbergstrasse 4, 85748 Garching, Germany

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Alice Ting (MIT), Simon Elsässer, and Yu-Hsuan Tsai for plasmids, Mathias Pasche for assistance with microscopy and Kai Johnsson (EPFL) for SiR. Funding: MRC grants U105181009 and UD99999908 (J.W.C.), Medical Research Council grant U105178788 (J.D.H. and N.P.B.), MRC/BBSRC/EPSC MR/K015680/1 (N.P.B. and J.W.C.), Marie Curie International Incoming Fellowship (C.U.), MRC Centenary Early Career Award (K.L.), MRC CASE studentship (Nikon) (V.B.).

## ■ REFERENCES

- (1) Bates, M.; et al. *Science* **2007**, *317*, 1749.
- (2) Betzig, E.; et al. *Science* **2006**, *313*, 1642.
- (3) Willig, K. I.; et al. *Nature* **2006**, *440*, 935.
- (4) Hofmann, M.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17565.
- (5) Gustafsson, M. G. J. *Microsc.* **2000**, *198*, 82.
- (6) Wang, S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 8452.
- (7) Dempsey, G. T.; et al. *Nat. Methods* **2011**, *8*, 1027.
- (8) Gautier, A.; et al. *Chem. Biol.* **2008**, *15*, 128.
- (9) Calloway, N. T.; et al. *ChemBioChem* **2007**, *8*, 767.
- (10) Los, G. V.; et al. *ACS Chem. Biol.* **2008**, *3*, 373.
- (11) Martin, B. R.; et al. *Nat. Biotechnol.* **2005**, *23*, 1308.
- (12) Uttamapinant, C.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10914.
- (13) Chin, J. W. *Ann. Rev. Biochem.* **2014**, *83*, 379.
- (14) Neumann, H.; et al. *Nat. Chem. Biol.* **2008**, *4*, 232.
- (15) Hancock, S. M.; et al. *J. Am. Chem. Soc.* **2010**, *132*, 14819.
- (16) Gautier, A.; et al. *J. Am. Chem. Soc.* **2010**, *132*, 4086.
- (17) Chen, P. R.; et al. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 4052.
- (18) Mukai, T.; et al. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818.
- (19) Greiss, S.; et al. *J. Am. Chem. Soc.* **2011**, *133*, 14196.
- (20) Bianco, A.; et al. *Nat. Chem. Biol.* **2012**, *8*, 748.
- (21) Lang, K.; et al. *J. Am. Chem. Soc.* **2012**, *134*, 10317.
- (22) Lang, K.; et al. *ACS Chem. Biol.* **2014**, *9*, 16.
- (23) Plass, T.; et al. *Angew. Chem. Intl. Ed. Engl.* **2012**, *51*, 4166.
- (24) Borrmann, A.; et al. *Chembiochem.* **2012**, *13*, 2094.
- (25) Lang, K.; et al. *Chem. Rev.* **2014**, *114*, 4764.
- (26) Nikic, I.; et al. *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 2245.
- (27) Schmied, W. *J. Am. Chem. Soc.* **2014**, *136*, 15577.
- (28) Galkin, V. E.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 20568.
- (29) Oda, T.; et al. *Nature* **2009**, *457*, 441.
- (30) Chernyatina, A. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 13620.
- (31) Sarkar, S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14366.
- (32) Lukinavicius, G.; et al. *Nat. Chem.* **2013**, *5*, 132.
- (33) Liu, D. S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, E4551.
- (34) Nieuwenhuizen, R. P.; et al. *Nat. Methods* **2013**, *10*, 557.
- (35) Belin, B. J.; et al. *Mol. Biol. Cell* **2013**, *24*, 982.
- (36) Kokai, E.; et al. *Histochem. Cell Biol.* **2014**, *141*, 123.
- (37) Bohnsack, M. T.; et al. *Nat. Cell Biol.* **2006**, *8*, 257.
- (38) Szymborska, A.; et al. *Science* **2013**, *341*, 655.