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

Genetically Encoded Cleavable Protein Photo-Cross-Linker

Shixian Lin,[†] Dan He,[†] Teng Long,[‡] Shuai Zhang,[†] Rong Meng,[†] and Peng R. Chen*[,†,‡]

[†]Beijing National Laboratory for Molecular Sciences, Synthetic and functional Biomolecules Center, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, College of Chemistry, and Molecular Engineering, Peking University, Beijing 100871, China

[‡]Peking-Tsinghua Center for Life Sciences, Beijing, China

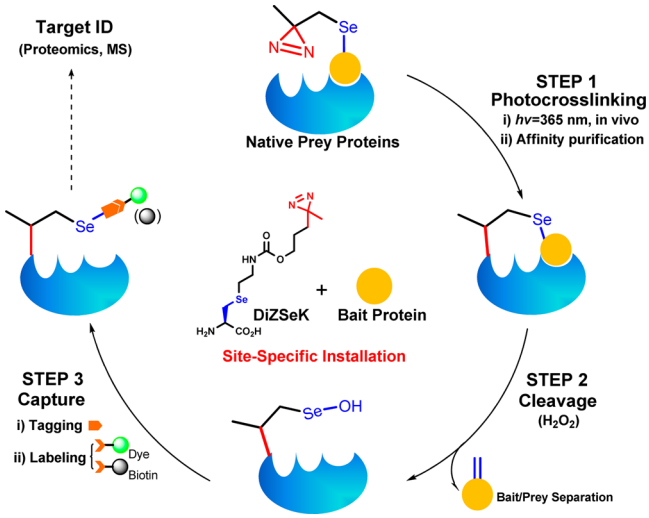
Supporting Information

ABSTRACT: We have developed a genetically encoded, selenium-based cleavable photo-cross-linker that allows for the separation of bait and prey proteins after protein photo-cross-linking. We have further demonstrated the efficient capture of the *in situ* generated selenenic acid on the cleaved prey proteins. Our strategy involves tagging the selenenic acid with an alkyne-containing dimethoxyaniline molecule and subsequently labeling with an azide-bearing fluorophore or biotin probe. This cleavage-and-capture after protein photo-cross-linking strategy allows for the efficient capture of prey proteins that are readily accessible by two-dimensional gel-based proteomics and mass spectrometry analysis.

Protein photo-cross-linking has become a highly valuable technique for the study of protein−protein interactions, particularly for weak or transient interactions that are difficult to capture with traditional methodology.[1] For example, a panel of photo affinity moieties (e.g., benzophenones, aryl azide, and diazirine groups) has been site-specifically incorporated into "bait" proteins via the genetic code expansion strategy. This incorporation allows for covalent linkage with its partner proteins ("prey") in close proximity under native cellular conditions.[1] Such site-specific protein photo-cross-linking methods significantly enhanced cross-linking efficiency and fidelity, thus enabling subsequent proteomics and mass spectrometry (MS) analysis for target identification. Despite these advantages, these methods still suffer from high false positive and false negative rates.[2] Nonspecific interactions between the bait and various cellular proteins have the potential to generate a large amount of false positive data; what is more, low enrichment efficiency for prey proteins causes missed hits as well.[3] Therefore, a strategy to separate the bait protein, and thus its nonspecific interacting partners, from the cross-linked prey proteins is desirable as a strategy to eliminate false-positive backgrounds. In addition, an easy capture and label strategy for use on the separated prey protein pool would further boost proteomic analysis efficiency. Herein, we have developed a genetically encoded, cleavable protein photo-cross-linker that is capable of separating bait and prey proteins after photo-cross-linking. Furthermore, we demonstrated that the cleavable handle allowed for efficiently capturing and labeling the prey protein pool that may be later accessed through proteomics and MS studies (Scheme 1).

Mild cleavage reactions, the reduction of disulfide bonds in particular, have been successfully applied to small-molecule-mediated protein photo-cross-linking or chemical cross-linking.[4] These methods easily remove the bait protein and facilitate the subsequent labeling or enrichment of the prey protein pool. Inspired by these examples, we reasoned that the introduction of a functional group into our previously developed, alkyl diazirine-containing photo-cross-linking unnatural amino acid (UAA)-DiZPK (Figure S1) could produce a genetically encoded cleavable photo-cross-linker.[5] Most current cleavable moieties are either prohibitively large in size or not chemically inert inside cells, thus rendering them useless. Indeed, for successful photo-cross-linking, the protein translational machinery-the amino acid tRNA-synthetase/tRNA pair must first recognize these UAAs.[6] Notably, amino acids that contain a selenium atom at the γ

Scheme 1. A General Procedure for Protein Photo-Cross-Linking Using a Cleavable Photo-Cross-Linker[a]



[a]Step 1, Photo-cross-linking of bait and prey proteins via a Se-containing DiZSeK (N^{ε}-3-((3-methyl-3H-diazirine-3-yl))-propamino-carbonyl-γ-seleno-L-lysine) amino acid before affinity purification of the protein complex. Step 2, Separation of bait and prey protein complexes via a H_2O_2-mediated oxidative cleavage reaction. Step 3, Capturing the prey proteins harboring an SA moiety generated *in situ* by (i) transferring a bioorthogonal handle and (ii) the subsequent bioorthogonal labeling with a fluorophore or biotin probe that are suitable for proteomics and MS analysis.

Received: May 1, 2014
Published: August 1, 2014

   dx.doi.org/10.1021/ja504371w | J. Am. Chem. Soc. 2014, 136, 11860−11863

position could undergo the intramolecular β-elimination reaction in the presence of mild oxidizing agents, generating dehydroanaline (Dha) and selenenic acid (SA).[7] This chemistry has been successfully employed for the *in situ* generation of Dha on a protein of interest (e.g., histone), which can be further modified to resemble various biologically active lysine or serine posttranslational modifications.[8] In an attempt to create a cleavable photo-cross-linker, we turned our attention to a Se atom that could replace the γ position carbon atom of DiZPK. We showed that our newly designed photoaffinity UAA, termed DiZSeK (Scheme 1), was successfully encoded by the pyrrolysine tRNA-synthetase (PylRS)-$tRNA^{Pyl}_{CUA}$ pair. Moreover, we were then able to use this newly designed UAA, in conjunction with our cleavage-and-capture after protein photo-cross-linking (CAPP) strategy, to efficiently separate and capture prey proteins for target identification.

We began by synthesizing DiZSeK through a tristep synthetic route shown in Figure S1 with an overall yield of 58% (Figures S2−4). As expected, the Se-containing linker on DiZSeK was cleaved by H_2O_2, generating Dha and DiZ-SA (Figure 1A).

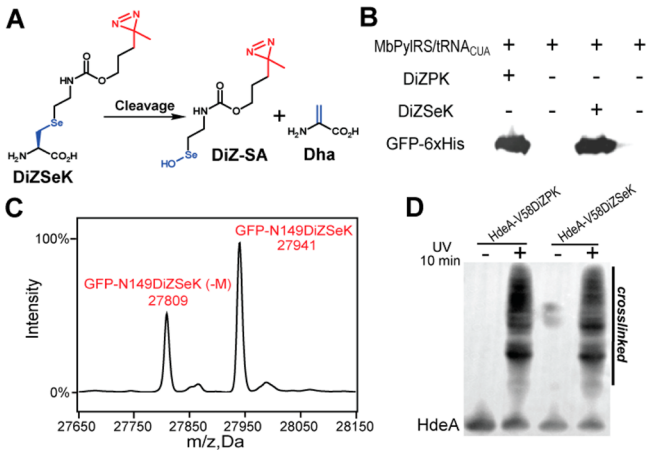


Figure 1. Development of a genetically encoded cleavable protein photo-cross-linker-DiZSeK. (A) The H_2O_2-mediated cleavage reaction on DiZSeK to generate DiZ-SA and Dha. The photoaffinity group and cleavable linker are shown in red and blue, respectively. (B) Western blot analysis of the amber suppression efficiency of the PylRS mutant-tRNA pair in recognizing DiZPK or DiZSeK. Full-length GFP carrying an in-frame amber mutation and a C-terminal Histag was produced only in the presence but not absence of 1 mM DiZPK or DiZSeK. (C) MS analysis of GFP-N149DiZSeK (calculated: 27943 Da, found: 27941 Da.) (D) *E. coli* cells expressing the periplasm-residing HdeA-V58DiZPK or HdeA-V58DiZSeK protein (carrying a C-terminal Histag) were incubated at pH 2.3 followed by 365 nm photolysis. Cell extracts were separated by SDS-PAGE and analyzed by immunobloting. An anti-His_6 antibody was used for both (B) and (D).

DiZPK and DiZSeK are highly similar structurally; we therefore tested if our previously obtained DiZPK-recognizing PylRS mutant could incorporate DiZSeK into proteins with its cognitive $tRNA^{Pyl}_{CUA}$. Indeed, this PylRS mutant showed similar amber suppression efficiency in the presence of DiZPK or DiZSeK, which allowed us to site-specifically incorporate DiZSeK at an in-frame amber mutation site N149 on green fluorescent protein (Figures 1B and S5). The yield of this DiZSeK-incorporated GFP, named GFP-N149DiZSeK, produced by *E. coli* cells was estimated to be ~8 mg·l^{-1} (Supporting Information). We used ESI-MS analysis on the produced protein in order to verify the specificity and fidelity of DiZSeK incorporation into proteins produced in *E. coli* (Figures 1C and S6). Notably, the GFP-N149DiZSeK can be converted to GFP-N149Dha upon H_2O_2-mediated oxidative cleavage; these results confirm that the DiZSeK probe can be oxidative cleaved when incorporated into an intact protein (Figure S7). Next, we examined the photo-cross-linking efficiency of DiZSeK by capturing the binding proteins of an *E. coli* acid chaperone HdeA inside living bacteria. *E. coli* cells expressing the HdeA protein carrying DiZSeK at residue S27 (HdeA-S27DiZSeK) were treated at pH = 2.3 for 30 min to allow HdeA's binding with its substrate proteins. Photoirradiation (365 nm) was next performed for 10 min on live *E. coli* cells, and the results were analyzed by SDS-PAGE gel and immunoblotting assay. High cross-linking efficiency comparable to that of DiZPK-mediated protein photo-cross-linking was obtained (Figures 1D and S8). In addition, the ratio of photo-cross-linked HdeA dimer complex to the monomer at pH 7 was obtained at a similar yield when DiZSeK or DiZPK was introduced at the protein dimer interface (residue Phe35). This result further suggests that use of Se to replace the γ-carbon on DiZPK has no effect on the function of the diazirine moiety (Figure S9). Taken together, this new DiZSeK probe showed amber suppression and photo-cross-linking efficiency similar to that of DiZPK photo-cross-linker.

To determine if the Se-containing handle can be cleaved upon oxidative treatment at the bait and prey protein interface, we first produced a covalent-linked HdeA dimer by photo-cross-linking HdeA at pH 7 with DiZSeK incorporated at residue Phe 35 within its dimer interface. The addition of H_2O_2 up to 8 mM[8b] still did not result in a decrease in the HdeA dimeric band on SDS-PAGE gel (Figure S10A). We reasoned that since Phe 35 is located at the center of the highly hydrophobic HdeA dimer, the H_2O_2 molecule may not be able to access it (Figure S11). We then used SDS to denature the cross-linked HdeA-dimer complex and performed the same oxidative-cleavage reaction afterward. The HdeA dimer was progressively disappearing on SDS-PAGE gel upon the treatment with an increasing amount of H_2O_2 (Figure S10B). As a control, no change was detected when DiZPK was used as the photo-cross-linking probe (Figure S10B). Completion of this cleavage reaction can be achieved in either Tris buffer (150 mM, pH = 8.0) or PBS buffer (pH 8) in the presence of 8 mM H_2O_2. In this composition we observed no detectable damage to GFP (carrying surface exposed cysteine and methionine residues) (Figure S7). Therefore, the concentration is compatible with proteomics and MS analysis. We also found that 0.5% SDS served as a better denaturation reagent than 8 M urea for this purpose (Figure S12). With an optimized reaction condition in hand, we investigated the separation of HdeA and its previously identified *in vivo* binding proteins (e.g., DegP and SurA) after *in vitro* photo-cross-linking under an acidic condition (pH 2.3).[5] The cross-linked low-pH protein complexes, HdeA/DegP and HdeA/SurA, were clearly detectable on the SDS-PAGE gel but disappeared after the addition of 8 mM H_2O_2 for 1 h (Figure S13). Together, these results confirm that the H_2O_2-mediated cleavage reaction may be used on the Se handle of DiZSeK in order to yield an efficient separation of bait and prey proteins after photo-cross-linking.

Next, we developed a CAPP strategy in order to capture the liberated prey proteins after oxidative-cleavage and separation of the prey and bait proteins. We designed a covalently linked GFP-Biotin system as the "prey-bait" model for proof-of-concept (Figures 2A,B). GFP was chosen as the prey protein because we have previously gained intensive experience on LC-MS/MS analysis of this protein containing different unnatural mod-

dx.doi.org/10.1021/ja504371w | *J. Am. Chem. Soc.* 2014, 136, 11860−11863

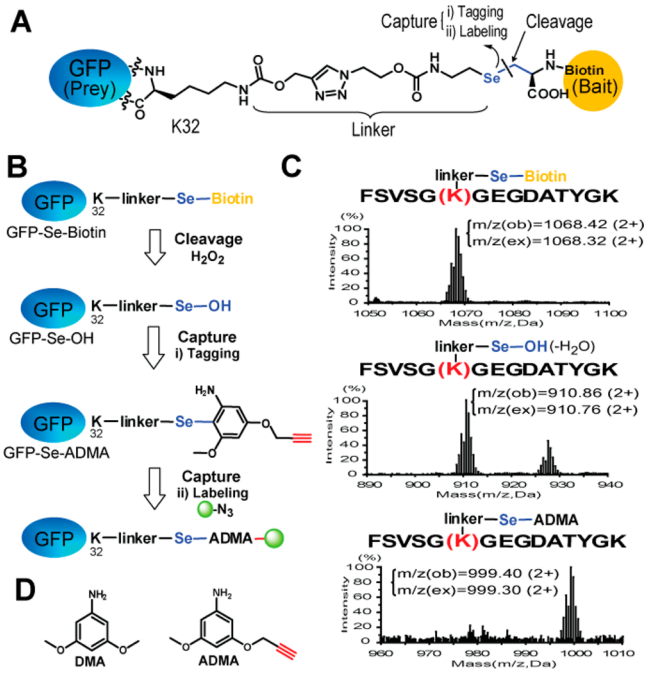


Figure 2. Demonstration of the CAPP strategy on a "prey-bait" model system. (A) A covalent-linked GFP-Biotin complex containing a Se-linker was created as a "prey-bait" model system: the Se-linker can be first oxidative-cleaved to generate SA moiety, which can be subsequently captured by tagging with an alkynyl-SA scavenger molecule suitable for CuAAC-labeling with an azide-bearing fluorophore. (B) Detailed procedure for CAPP strategy is as follows: (1) cleavage of the prey-bait complex (GFP-Biotin) via a H_2O_2-mediated oxidative reaction and (2) capturing the SA that contains prey proteins via (i) tagging the SA moiety generated *in situ* with ADMA and (ii) labeling the transferred ADMA tag with an azide that contains fluorophore. (C) Monitoring the workflow of the CAPP procedure by LC-MS/MS. The MS data of the following variants are shown: *GFP-Se-Biotin*, *GFP-Se−OH*, and *GFP-Se-ADMA* together with the desired peptide sequence, the observed (ob) and expected (ex) MW. (D) Structures of DMA and its alkynyl analogue ADMA, with function group in red.

ifications.[9] We used a copper-catalyzed azide−alkyne cycloaddition (CuAAC) reaction to conjugate a GFP that bears a site-specifically incorporated alkyne handle with an azido-biotin molecule carrying a γ-seleno substituted lysine linker at a molar ratio of 1:4. ESI-MS analysis verified the resulting GFP-Se-Biotin complex (Figures 2B,C and S14−15). H_2O_2-mediated cleavage was then applied to this model system for 1 h before SDS-PAGE being applied to the cleaved products. In order to successfully monitor the cleavage process, we used in-gel LC-MS/MS analysis to detect the desired SA-containing peptide from the cleaved product GFP-Se-OH (Figures 2B,C and S16−17).

In order to directly capture the cleaved, SA-containing proteins, we surveyed a series of aniline-based compounds that have been previously used as SA scavengers during oxidative elimination of the phenylselenide group (Figure S18).[10] To our delight, we found that 3,5-dimethoxyaniline (DMA) is an efficient capture reagent that can specifically react with the SA moiety generated *in situ* after oxidative cleavage of DiZSeK (Figures 2D and S19).[10] We then designed and synthesized an alkyne-containing DMA analogue-ADMA (Figures 2D and S20). To determine the reactivity of ADMA with SA generated *in situ* from GFP, we applied the CAPP strategy on denatured GFP-Se-Biotin by treating the protein sample simultaneously with 8 mM H_2O_2 and 500 μM ADMA at 30 °C for 2 h (Figure 2B). LC-MS/MS analysis confirmed that ADMA was successfully conjugated with SA and transferred onto the cleaved GFP template to generate GFP-Se-ADMA (Figures 2B,C and S16−17). To further test the specificity of our CAPP strategy, H_2O_2 and ADMA were immediately desalted, and the protein samples were labeled with Fluor 488-azide through CuAAC reaction (Figure 2B). In-gel fluorescence imaging clearly detected the fluorescently labeled bands (Figure S21), while no labeled bands were observed in the lane containing denatured WT-GFP (carrying two cysteine residues) under the same CAPP protocol (Figure S22). These observations indicate that proteinogenic amino acids including cysteine did not react with ADMA under our experimental conditions. The reaction procedure and the labeling yield were further optimized by in-gel fluorescence imaging assay (Figures S21−22). Together, our CAPP strategy exhibits high efficiency in capturing the *in situ* generated prey proteins after cleavage of the cross-linked prey-bait complexes.

We next tested the compatibility of our CAPP strategy for labeling and profiling the *in vivo* binding proteins of HdeA after protein photo-cross-linking. Indeed, we observed a band shift before and after the addition of H_2O_2 into the purified HdeA-substrate protein complex on SDS-PAGE (Figure S23), indicating the separation of HdeA and its binding proteins after the cleavage reaction. Furthermore, strong azide-biotin labeling signal was only detected in the presence of both H_2O_2 and ADMA (Figure S24), which agreed with our Western blot analysis after the bait protein was removed by affinity purification (Figure S25). Noteworthy, the subsequent capturing of the liberated prey proteins with a biotin tag allowed for the enrichment of prey proteins with streptavidin beads, which would significantly improve the detection limit of target proteins.

Finally, we demonstrated direct profiling of *in vivo* binding proteins of HdeA under acid stress by using the DiZSeK probe and CAPP strategy in conjunction with two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). Although 2D-PAGE is a method routinely used in proteomics,[11] coupling this method with protein photo-cross-linking remains challenging, mainly because of low cross-linking efficiency as well as the highly abundant bait proteins that are difficult to separate even with 2D gel electrophoresis. In contrast, our newly developed DiZSeK probe is not only able to completely remove the bait protein but can also label and/or enrich the binding proteins that are directly applicable for 2D-PAGE and the subsequent MS analysis. Following the previously described strategy, *E. coli* cells expressing the periplasm-residing HdeA-F35DiZSeK were treated with acid stress (pH 2.3) and photoirradiated for 10 min. The cross-linked protein complexes were affinity purified with Ni-NTA column before being applied to CAPP strategy (Figure S26). The prey protein pool labeled with Cy5-N_3 following the CAPP procedure was then analyzed by 2D-PAGE, which detected ∼80 dots on the fluorescent gel (Figure S27A). A coomassie gel with a near 4-fold higher loading of the protein samples without labeling with Cy5-N_3 was next used to afford sufficient samples for LC-MS/MS analysis (Figure 3). Although, this coomassie gel showed a similar pattern as the aforementioned fluorescent gel, these two 2D gels were not completely overlaid (Figure S27). The main reason was likely due to the different sample loading between the coomassie gel and the fluorescent gel, which may cause slightly different mobility change during gel-based analysis. Nevertheless, 2D-PAGE analysis on this coomassie gel alone is suitable for proof-of-concept analysis, and overlay with fluorescence gel is not necessary in this study.

dx.doi.org/10.1021/ja504371w | J. Am. Chem. Soc. 2014, 136, 11860−11863

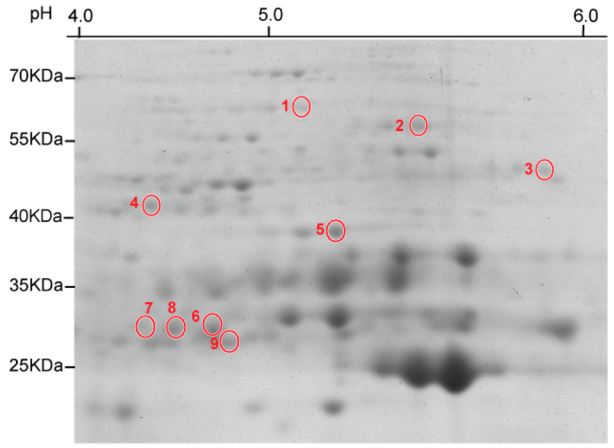


Figure 3. 2D-PAGE analysis of the photo-cross-linked HdeA binding proteins using DiZSeK and CAPP strategy on a coomassie gel. The HdeA binding proteins after photo-cross-linking with the DiZSeK probe were H_2O_2-cleaved, tagged with ADMA for 2D-PAGE analysis. LC-MS/MS-analyzed spots are indicated in red circles.

We randomly chose nine spots on the coomassie gel that overlaid with the spots on fluorescent gel (Figure 3; additional details regarding these binding proteins from each spot are in Table S1). LC-MS/MS analysis identified 13 binding proteins (Table S1, colored in orange) including 10 proteins (Table S2, colored in blue) that have been previously reported as HdeA clients, thus verifying the reliability of our method.[5] Three new proteins, PotF (putrescine-binding protein, residing in periplasm), FliY (cystine-binding periplamic protein) and DhsB (succinate dehydrogenase iron−sulfur subunit, residing on the inner membrane), were identified for the first time as potential HdeA binding proteins under acid stress (Table S2, colored in red), which expanded our knowledge of the *in vivo* substrate profiles for this acid chaperone. Taken together, the combination of DiZSeK cleavable photo-cross-linker, CAPP strategy, and 2D-PAGE offer a powerful tool for the systematic profiling of the interaction protein targets of a given protein in living cells.

In summary, we have employed a pyrrolysine-based genetic code expansion system to encode a Se-containing cleavable protein photo-cross-linker and developed a cleavage-and-capturing of interaction CAPP strategy. This cleavable photo-affinity amino acid not only covalently traps prey proteins under living conditions but also allows for the subsequent separation of bait and prey proteins via H_2O_2-mediated oxidative cleavage. The released prey proteins carrying the *in situ* generated SA moiety could be further captured by (i) tagging with an alkyne-bearing DMA molecule and (ii) labeling with an azide-containing fluorophore or biotin probe. This CAPP strategy, in conjunction with the 2D-PAGE proteomics and MS analysis, enhances the separation and enrichment efficiency for identifying native substrates of a given protein after photo-cross-linking. In this study, we demonstrated this concept by profiling the *in vivo* binding proteins of an *E. coli* acid chaperon HdeA under acid stress, which revealed potential new substrates that have not been identified from our previous work using noncleavable photo-cross-linker. In addition, the strategy we developed here could be used to improve the identification of the site of cross-linking on the prey proteins with MS analysis in future. Given the general applicability of this DiZSeK probe as well as the broad utilization of the Pyl-based genetic-code expansion system, this strategy may become a invaluable method in identifying protein−protein interactions in diverse living species.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details and supplemental data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

pengchen@pku.edu.cn

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Key Basic Research Foundation of China (2010CB912302 and 2012CB917301), National Natural Science Foundation of China (21225206 and 91313301). We thank supports from Peking University Principal Foundation. S.F. Reichard edited the manuscript.

■ REFERENCES

(1) (a) Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol. Biosyst.* **2008**, *4*, 473−480. (b) Preston, G. W.; Wilson, A. J. *Chem. Soc. Rev.* **2013**, *42*, 3289−3301. (c) Pham, N. D.; Parker, R. B.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2013**, *17*, 90−101.

(2) (a) Wu, H.; Ge, J.; Yang, P.-Y.; Wang, J.; Uttamchandani, M.; Yao, S. Q. *J. Am. Chem. Soc.* **2011**, *133*, 1946−1954. (b) Park, J.; Oh, S.; Park, S. B. *Angew. Chem., Int. Ed.* **2012**, *51*, 5447−5451.

(3) (a) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. *Nat. Methods.* **2005**, *2*, 201−206. (b) Hino, N.; Oyama, M.; Sato, A.; Mukai, T.; Iraha, F.; Hayashi, A.; Kozuka-Hata, H.; Yamamoto, T.; Yokoyama, S.; Sakamoto, K. *J. Mol. Biol.* **2011**, *406*, 343−353.

(4) (a) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem.* **2008**, *77*, 383−414. (b) Leriche, G.; Chisholm, L.; Wagner, A. *Bioorg. Med. Chem.* **2012**, *20*, 571−582. (c) Tamura, T.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2012**, *134*, 2216−2226.

(5) (b) Zhang, M.; Lin, S.; Song, X.; Liu, J.; Fu, Y.; Ge, X.; Fu, X.; Chang, Z.; Chen, P. R. *Nat. Chem. Biol.* **2011**, *7*, 671−677. (b) Lin, S.; Zhang, Z.; Xu, H.; Li, L.; Chen, S.; Li, J.; Hao, Z.; Chen, P. R. *J. Am. Chem. Soc.* **2011**, *133*, 20581−20587.

(6) (a) Xiang, Z.; Ren, H.; Hu, Y. S.; Coin, I.; Wei, J.; Cang, H.; Wang, L. *Nat. Methods* **2013**, *10*, 885−888. (b) Grammel, M.; Hang, H. C. *Nat. Chem. Biol.* **2013**, *9*, 475−484. (c) Chin, J. W. *Annu. Rev. Biochem.* **2014**, *83*, 379−408.

(7) Buchardt, O.; Elsner, H. I.; Nielsen, P. E.; Petersen, L. C.; Suenson, E. *Anal. Biochem.* **1986**, *158*, 87−92.

(8) (a) Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 6399−6401. (b) Wang, Z. U.; Wang, Y.-S.; Pai, P.-J.; Russell, W. K.; Russell, D. H.; Liu, W. R. *Biochemistry* **2012**, *51*, 5232−5234.

(9) Lin, S.; Yan, H.; Li, L.; Yang, M.; Peng, B.; Chen, S.; Li, W.; Chen, P. R. *Angew. Chem., Int. Ed.* **2013**, *52*, 13970−13974.

(10) Tietze, L.-F.; v.Kiedrowski, G.; Berger, B. *Tetrahedron. Lett.* **1982**, *23*, 51−54.

(11) Wulfkuhle, J. D.; McLean, K. C.; Paweletz, C. P.; Sgroi, D. C.; Trock, B. J.; Steeg, P. S.; Petricoin Iii, E. F. *Proteomics* **2001**, *1*, 1205−1215.