How to cite: *Angew. Chem. Int. Ed.* **2022**, *61*, e202112107
International Edition: doi.org/10.1002/anie.202112107
German Edition: doi.org/10.1002/ange.202112107

Bioconjugation

Tunable Amine-Reactive Electrophiles for Selective Profiling of Lysine

*Kuei-Chien Tang, Jian Cao, Lisa M. Boatner, Linwei Li, Jonathan Farhi, Kendall N. Houk, Jennifer Spangle, Keriann M. Backus, and Monika Raj**

Abstract: *Proteome profiling by activated esters identified >9000 ligandable lysines but they are limited as covalent inhibitors due to poor hydrolytic stability. Here we report our efforts to design and discover a new series of tunable amine-reactive electrophiles (TAREs) for selective and robust labeling of lysine. The major challenges in developing selective probes for lysine are the high nucleophilicity of cysteines and poor hydrolytic stability. Our work circumvents these challenges by a unique design of the TAREs that form stable adducts with lysine and on reaction with cysteine generate another reactive electrophiles for lysine. We highlight that TAREs exhibit substantially high hydrolytic stability as compared to the activated esters and are non-cytotoxic thus have the potential to act as covalent ligands. We applied these alternative TAREs for the intracellular labeling of proteins in different cell lines, and for the selective identification of lysines in the human proteome on a global scale.*

Introduction

Lysine shows rich chemistry through its nucleophilic amine group and is abundant in various active and allosteric sites. Lysines also catalyze multiple reactions and regulate various biological processes.[1–3] Moreover, lysines are frequent sites for posttranslational modifications and regulate the structure and functions of proteins.[4–7] Lysines have a high frequency in human proteins (≈6% of all residues).[8] Together, these properties make lysine residues desirable targets with covalent drugs; however, the low nucleophilicity of lysine as compared to cysteine makes selective targeting by covalent probes a highly daunting task. Several small molecule covalent ligands for selective proteome profiling of cysteine has been reported;[9–15] however most of the small molecule electrophiles for lysines such as dichlorotriazines,[16] imidoesters,[17] 2-acetyl- or 2-formyl-benzeneboronic acids,[18] isothiocyanates,[19,20] pyrazolecarboxamidines,[21,22] sulfonyl fluorides,[23,24] and vinyl sulfonamides[25] also react with other amino acids such as serine, tyrosine and cysteine. Sulfonyl acrylate reagent has been reported for the regio- and chemoselective labeling of lysine on pure protein by using low equivalent of the probe but it has a potential to react with cysteine during profiling.[26] Recently, activated esters such as NHS-ester[27] and STPyne[28] have been explored for ligandability of lysine in the human proteome. Although >9000 ligandable lysine sites have been discovered using NHS-esters[26] and STP-esters,[28] none of these probes act as a covalent ligand/inhibitor for lysine in cells because of poor hydrolytic stability (Figure 1a).[27,28] Therefore for the discovery of covalent ligands for lysines, new amine reactive chemotypes are needed that are stable to enzymatic and non-enzymatic hydrolysis, non-cytotoxic and generate stable covalent adducts with lysine. These lysine selective, hydrolytically stable covalent probes would significantly expand the druggable content of the human proteome. With these considerations in mind, we sought to develop new Tunable Amine-Reactive Electrophiles (TAREs) as covalent probes that would retain the advantages of covalent lysine targeting (stable adduct with lysine and high selectivity) without the potential liabilities associated with hydrolysis and cross-reactivity with other nucleophilic amino acids (Figure 1b). In this study, we elucidate specific structural features that underlie selective lysine addition to generate stable adducts, and we apply these principles to the design of lysine-targeted covalent ligands for live cell labeling and proteome profiling. While our paper was under review, a new study exploring varying probes for proteome profiling was published but detailed studies are needed on probes to determine their cytotoxicity, stability inside cells and their suitability to act as covalent ligands for lysine.[29] Here, we show that TAREs are non-cytotoxic, exhibit high hydrolytic stability, high activity in live cells and selectively enrich lysine in the human proteome with negligible enrichment of cysteine.

Results and Discussion

We sought to achieve selective enrichment of lysine by designing electrophilic heteroaromatic thioethers that on reaction with cysteine generate another lysine reactive electrophilic heteroaromatic thioethers that are capable of

[*] K. Tang, Prof. Dr. M. Raj
Present address: Department of Chemistry
Emory University, Atlanta, GA 30322 (USA)
E-mail: monika.raj@emory.edu
Homepage: https://raj.emorychem.science
Dr. J. Cao, L. M. Boatner, L. Li, Prof. Dr. K. N. Houk, Prof. Dr. K. M. Backus
Department of Chemistry and Biochemistry
College of Arts and Sciences, UCLA
Los Angeles, CA 90095 (USA)
L. M. Boatner, Prof. Dr. K. M. Backus
Present address: Department of Biological Chemistry
David Geffen School of Medicine, UCLA
Los Angeles, CA 90095 (USA)
J. Farhi, Prof. Dr. J. Spangle
Department of Radiation Oncology, Emory University
Atlanta, GA 30322 (USA)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202112107.

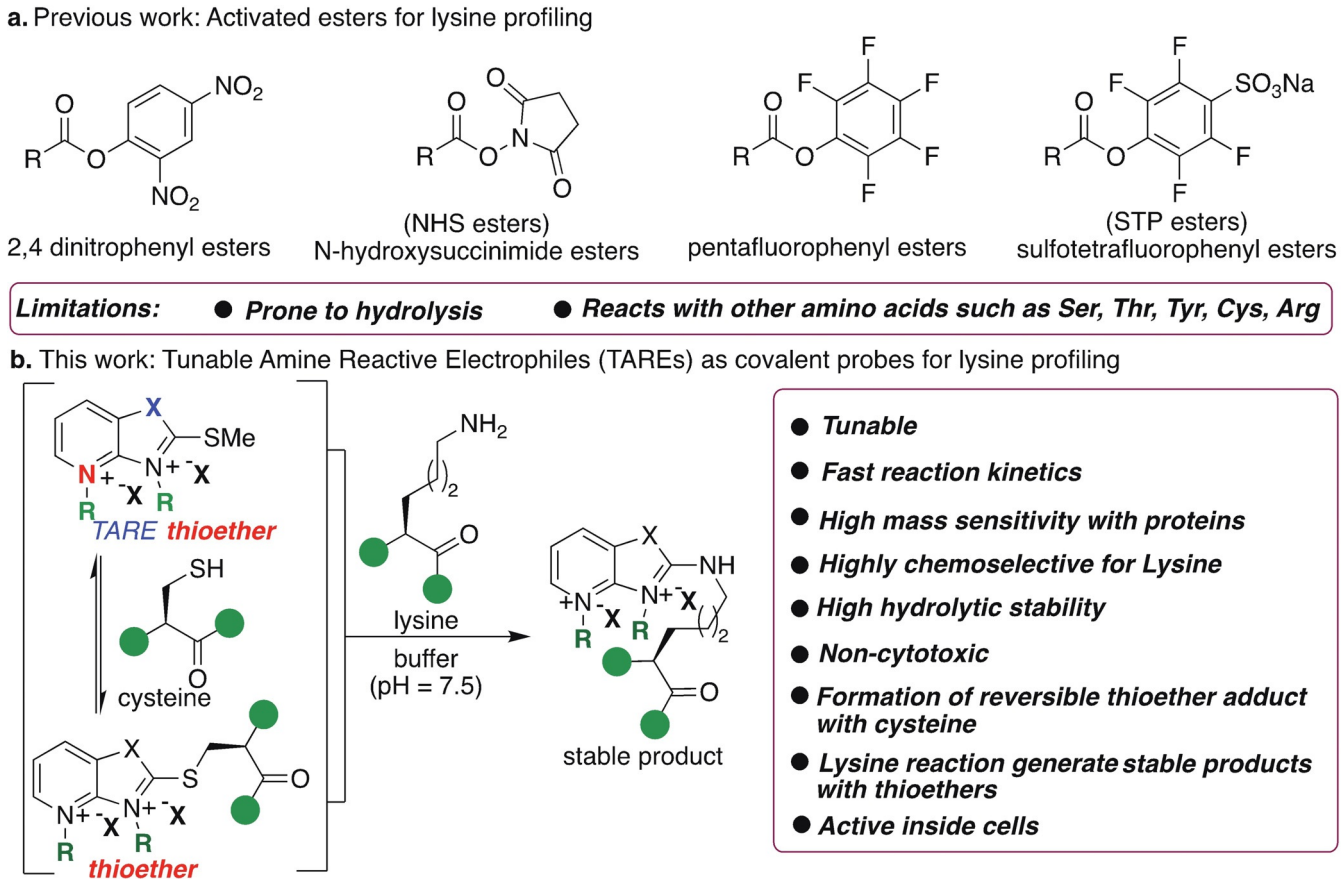


Figure 1. Tunable amine-reactive electrophiles (TAREs) for selective profiling of lysines. a) Previous work on activated esters for profiling lysine and their limitations. STP- and NHS-esters are commonly used for lysine profiling. b) TAREs, tunable and hydrolytically stable reagents for selective lysine proteome profiling. Nucleophilic Cys reacts with TAREs to generate another electrophilic thioether probes for selective reaction with lysine to generate stable products. Proposed reaction pathway between lysine and a TARE proceeds by nucleophilic attack of lysine on the heteroaromatic ring of thioether and undergoes SNAr substitution with lysine to generate stable adducts.

forming stable adducts with lysine (Figure 1b). We tune the heteroaromatic thioether probes to make them water soluble, selective for lysine, hydrolytically stable and easy analysis by mass spectrometry. We achieved this by (i) varying different heteroatoms O/S on heteroaromatic ring, (ii) varying the fused aromatic ring (benzene and pyridine) and (iii) by methylation of N-heteroatoms (Figure 2a). We initiated our studies by exploring the reactivity of a panel of TAREs (Figure 2a, Synthesis of probes, Supporting Information, Figure 1) with a peptide FKVCF **2a**, containing nucleophilic amino acids N-terminus, Lys and Cys, under phosphate buffer (NaP, pH 7.5, 10 mM) at room temperature (Figure 2b). We did not observe the formation of an adduct with heteroaromatic 2-methylthio benzooxazoline probe **1a** due to its poor solubility in a buffer and lower reactivity. To increase aqueous solubility, we replaced the benzene ring with pyridine, but the solubility of the probe 2-methylthio pyridoxazoline **1b** remained poor. We next carried out the methylation of the pyridine of the 2-methylthio pyridoxazoline **1b** to generate the 2-methylthio pyridinium oxazoline ion **1c** that substantially increased its reactivity and solubility in buffer (Figure 2a). The reaction of **1c** with a peptide FKVCF **2a** showed the modification of both Cys and Lys under the reaction conditions as analyzed by LC-MS/MS analysis (90% Figure 2b, Supporting Information, Figure 2). To further confirm the chemoselectivity for lysine and cysteine, reactions with probe **1c-yne** were performed with peptides PGYAHF and WDQRF containing other nucleophilic amino acid residues such as S, T, Y, W, E and H but no modification of any amino acids was observed (Supporting Information, Figure 2). To increase the reactivity and aqueous solubility of 2-methylthio benzooxazoline probe **1a**, we further methylated the heteroatom of the azoline ring to generate 2-methylthio benzoN-methyloxazolinium ion **1d**. The reaction of **1d** with a peptide FKVCF **2a** showed the modification of Lys amino acid only with 99% conversion to the modified lysine product. Interestingly, we did not observe any modification of the Cys and the N-terminal residue under the reaction conditions as determined by LC-MS/MS analysis (Figure 2b, Supporting Information, Figure 2). In contrast, the reaction of FKVCF **2a** with NHS and STP activated esters did not show any selectivity and labeled the N-terminus, cysteine and lysine within 3 h under the reaction conditions (Supporting Information, Figure 2).

To characterize the Lys coupling product with **1d**, a reaction with a model compound hexadecylamine was carried out on a large scale under the optimized conditions. The resulting product was isolated and Lys labeling by **1d** was

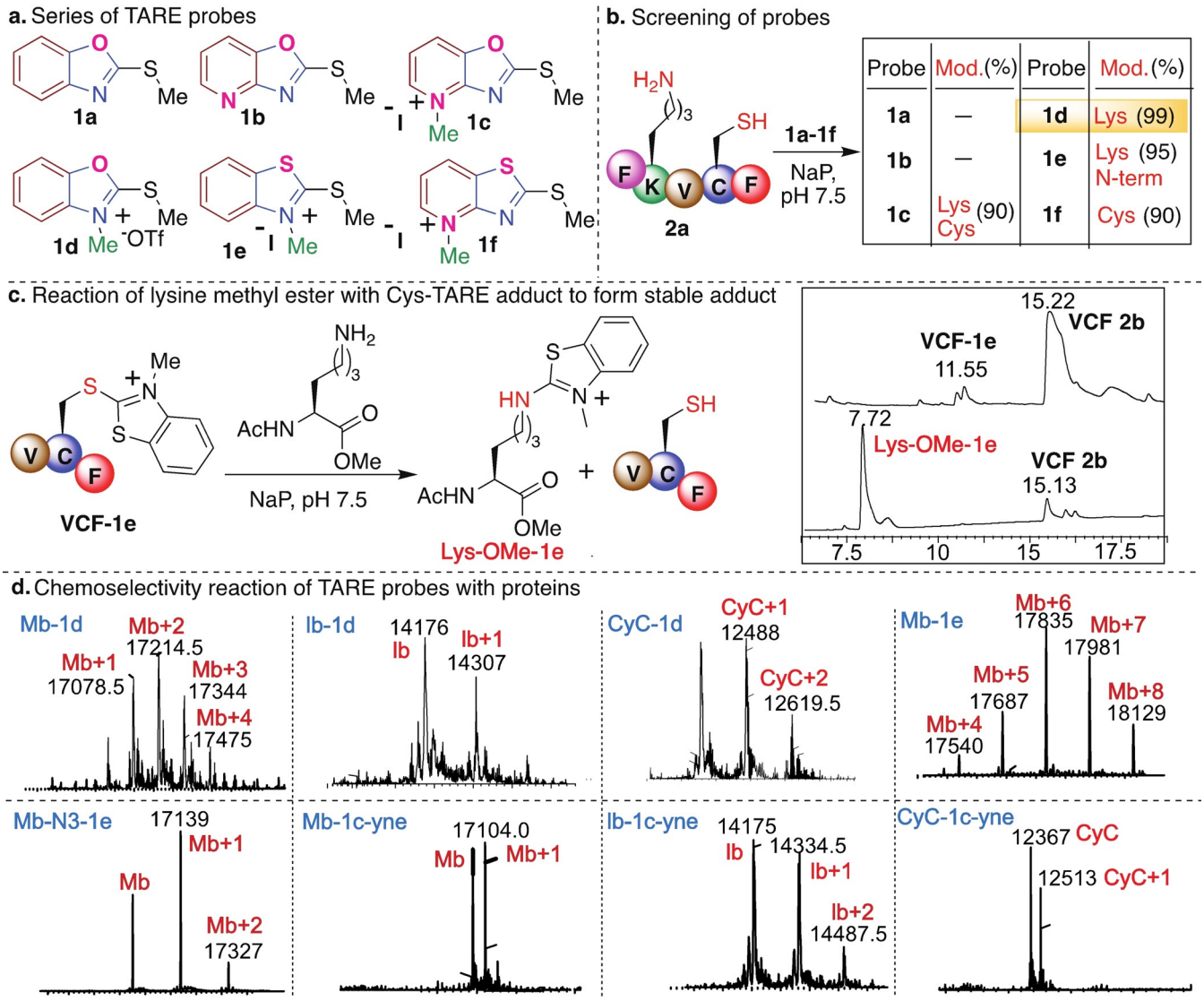


Figure 2. Tunable and stable TAREs for targeting lysine. a) Structures of a variety of TAREs with different modifications to interrogate their reactivity and selectivity for lysine. b) Screening of TAREs to determine high selectivity for lysine by using peptide FKVCF **2a** with all the nucleophilic residues. Probe **1d** showed high selectivity and reactivity for lysine. c) Reactivity of cysteine-TARE conjugate towards lysine to generate stable product. d) Modification of proteins with various TAREs. These proteins do not have any cysteine residue and TAREs showed high selectivity for lysine as analyzed by LC-MS/MS. Reaction Conditions: protein (3 mM in Nap (pH 7.5), probe **1d** (10 equiv, 30 mM), **1e** (100 equiv, 300 mM), room temperature for 1 h, detection wavelength 200 nm. For **Ib** and CyC modification, probe 1 equiv of **1d** (3 mM) was used for 1 h.

confirmed by NMR (^{1}H and ^{13}C) (Supporting Information, Figure 3). In contrast, no product was observed with a peptide Ac-VCF (without Lys), thus reconfirms the high selectivity of **1d** for Lys. We hypothesized that this might be due to the high reactivity of the **1d**-cysteine conjugate towards lysine to generate a stable product or for the released methanethiol to generate unreacted starting material. To confirm our hypothesis, we synthesized a probe 2-methylthio benzoN-methylthiozolinium ion **1e** of lower reactivity by replacing O heteroatom with S in the 2-methylthio benzoN-methyloxazolinium ion **1d**. We isolated the small amounts of **1e**-Cys-conjugate (**VCF-1e**) by carrying out the reaction of **1e** with a peptide AcVCF and incubated the cysteine-adduct **VCF-1e** with lysine methylester (Figure 2c, Supporting Information, Figure 4). As envisioned, the reaction resulted in the formation of a stable adduct with lysine methylester (**Lys-OMe-1e)** releasing the free peptide VCF **2b** as analyzed by LCMS (Figure 2c, Supporting Information, Figure 4). Since **1d-Cys** conjugate is highly reactive and difficult to isolate under the aqueous conditions, we synthesized **1d**-thio-conjugate under non-aqueous conditions and observed complete modification with lysine methylester within 1 h under physiological conditions (NaP, pH 7.5, 10 mM) at room temperature (Supporting Information, Figure 4). These experiments confirmed that thioethers generated by the reaction of Cys with TAREs **1c**–**1e** further react with lysine thus exhibit a potential for selective profiling of lysine in the human proteome. We also incubated **1d**-thio-conjugate under buffer (NaP, pH 7.5, 10 mM) at room temperature and observed complete hydrolysis in 1 h (Supporting Information, Figure 4) suggesting the

selectivity for lysine profiling. The reaction of **1e** with a peptide FKVCF **2a** modified both the N-terminus and Lys as analyzed by MS/MS (95% Figure 2b, Supporting Information, Figure 2). We also synthesized probe **1f** of lower reactivity by replacing O heteroatom with S in the azoline ring of 2-methylthio pyridinium oxazoline ion **1c**. As anticipated, we detected the modification at Cys of a peptide FKVCF **2a** with **1f** as analyzed by LC-MS/MS (90% Figure 2b, Supporting Information, Figure 2).

Previous studies with carbon electrophiles such as chloroacetamide, sulfonate esters, and aryl halides showed that the reactivity of electrophiles with peptides often vary from their reactivity with proteins due to the unique environment of reactive amino acids in proteins.[14,30,31] To evaluate the selectivity and reactivity of probes **1d** and **1e** with proteins, we carried out the reaction with Myoglobin (Mb) (without Cys Figure 2d).

The multiple modifications of lysines on Mb were observed with both the probes **1d** (10 equiv, 1 h) and **1e** (100 equiv, 12 h) as determined by MS/MS analysis on **Mb-1d** and **Mb-1e** protein conjugates (Figure 2d, Supporting Information, Figure 5, Supporting Information, Table 1 and Table 2). The modification of any other reactive amino acids on Mb such as Ser, Thr, Tyr, Trp, Glu, His was not observed even after using excess of probe **1e** (100 equiv) and for longer reaction time (12 h), thus further confirms high reactivity and selectivity of **1d** and **1e** for lysine. Next, we explored probes **1d** for the modification of other proteins such as lactalbumin (lb) and cytochrome C (CyC) and MS/MS analysis confirmed the modification of lysine amino acid only in **lb–1d** and **CyC-1d** protein conjugates (Figure 2d and Supporting Information, Figure 6). TAREs **1c–1f** are charged molecules and increase the ionization of the tagged fragments thus has a potential to enhance their detection sensitivity by MS. We showed that **1d** significantly increased the detection sensitivity of the labeled proteolytic fragments of cytochrome C as compared to the unlabeled fragments and STP and NHS-ester modified fragments (Supporting Information, Figure 7). The MS results clearly showed that the MS sensitivity of the fragments modified by **1d** probes is 100 fold higher than the mass sensitivity of the fragments modified by STP, NHS-ester and unlabeled fragments (Supporting Information, Figure 7).

The effective reaction for the enrichment of lysine fragments from a complex mixture should have the ability to attach affinity tags, thus we generated the alkyne and azide functionalized TAREs **N3-1e**, **1c-yne** and **1d-yne** (Synthesis of probes, Supporting Information, Figure 8). We then evaluated the reactivity of the azide- and alkyne-functionalized TAREs towards proteins such as Mb, lb and CyC. Similar to the model probe studies with **1c–1e**, azide- and alkyne-TAREs **N3-1e** (25 equiv 12 h) and **1c-yne** (1 equiv, 1 h) modified Lys of proteins Mb, lb and CyC under optimized conditions as confirmed by MS/MS (Figure 2d, Supporting Information, Figure 9). The modification of lysine residue is independent of the presence of the cysteine residue as observed by the labeling of significant amount of lysine residues (6–7 lysines) on pure proteins myoglobin, cytochrome C and lactalbumin without any cysteine using high equivalents of the probe **1c-yne** (50 equiv) (Supporting Information, Figure 9).

Next, we compared the reactivity of the two most reactive TAREs **1c-yne** and **1d-yne** by carrying out the rate studies using a peptide Ac-GKF (**GKF 2c**). The reactions were monitored after regular intervals of time using HPLC and MS (Figure 3a, Supporting Information, Figure 10). The reaction with **1d-yne** ($k = 307.52\ M^{-1}S^{-1}$) showed 3-fold higher reactivity than **1c-yne** ($k = 99.27\ M^{-1}S^{-1}$) (Supporting Information, Figure 10). The reaction rate of STPyne ($k = 190.92\ M^{-1}S^{-1}$) with a peptide GKF **2c** showed lower reactivity compared to **1d-yne** but high reactivity than **1c-yne** (Figure 3a, Supporting Information, Figure 10). Next, we sought to determine the stability of the reactive TAREs **1c-yne** and **1d-yne** towards hydrolysis and compare it with hydrolytic stability of NHS-ester and STPyne previously used for lysine profiling.[27,28] We incubated probes in aqueous phosphate buffer (pH 7.5) under ambient conditions and monitored after regular intervals of time by HPLC and MS. Although of almost similar reactivity, **1c-yne** probe is more stable towards hydrolysis as compared to **1d-yne** and only 30% degradation of **1c-yne** probe was observed after 6 h and 50% of **1c-yne** remained intact even after 12 h (Figure 3b, Supporting Information, Figure 11). In contrast, NHS-ester showed 90% degradation in 2 h and completely degraded within 4 h (Figure 3b, Supporting Information, Figure 11). STPyne showed 90% degradation in 6 h. These studies showed the high stability of TARE **1c-yne** as compared to NHS-ester and STPyne thus capable of acting as covalent inhibitors of lysine in cellular environment (Figure 3b, Supporting Information, Figure 11). The most reactive TARE probes **1c-yne** and **1d-yne** are bench stable for multiple weeks as a white solid, demonstrating robust properties for long-term storage.

Bolstered by the lysine-selectivity and elevated reactivity observed with our TARE-protein and peptide-based labeling studies, we next sought to investigate whether this chemotype would be suitable for activity-based chemoproteomic applications in complex cell lysates. Using a gel-based competitive activity-based protein profiling (ABPP) assay, we first assessed the relative cysteine- and lysine-reactivity of TAREs. HEK293T lysates were first treated with probes **1d**, **1c-yne** or **1d-yne** followed by labeling with a pan-cysteine or the pan-lysine reactive fluorescent probe, iodoacetamide-tetramethylrhodamine (IA-Rh) or NHS-tetramethylrhodamine (NHS-Rh), respectively (Figure 3c, Supporting Information, Figure 12). While **1c-yne** blocked both cysteine and lysine labeling (Figure 3e), gel-based ABPP analysis of compound **1d** suggested improved specificity for lysine (Figure 3d), as indicated by competition of labeling by NHS-Rh but not IA-Rh. By comparison, the canonical lysine-reactive probe STPyne (1 mM) competed labeling of both probes, consistent with off-target cysteine-reactivity, as reported previously.[32]

As competitive blockage of NHS-Rh labeling was only observed at relatively high concentrations of **1c** and **1d** (e.g. 5 mM for **1d**, Figure 3d,e, Supporting Information, Figure 12), our next step was to evaluate the concentration range compatible with probe labeling. For these studies, we turned to an alkyne derivative of **1d**, termed **1d-yne**. Cell

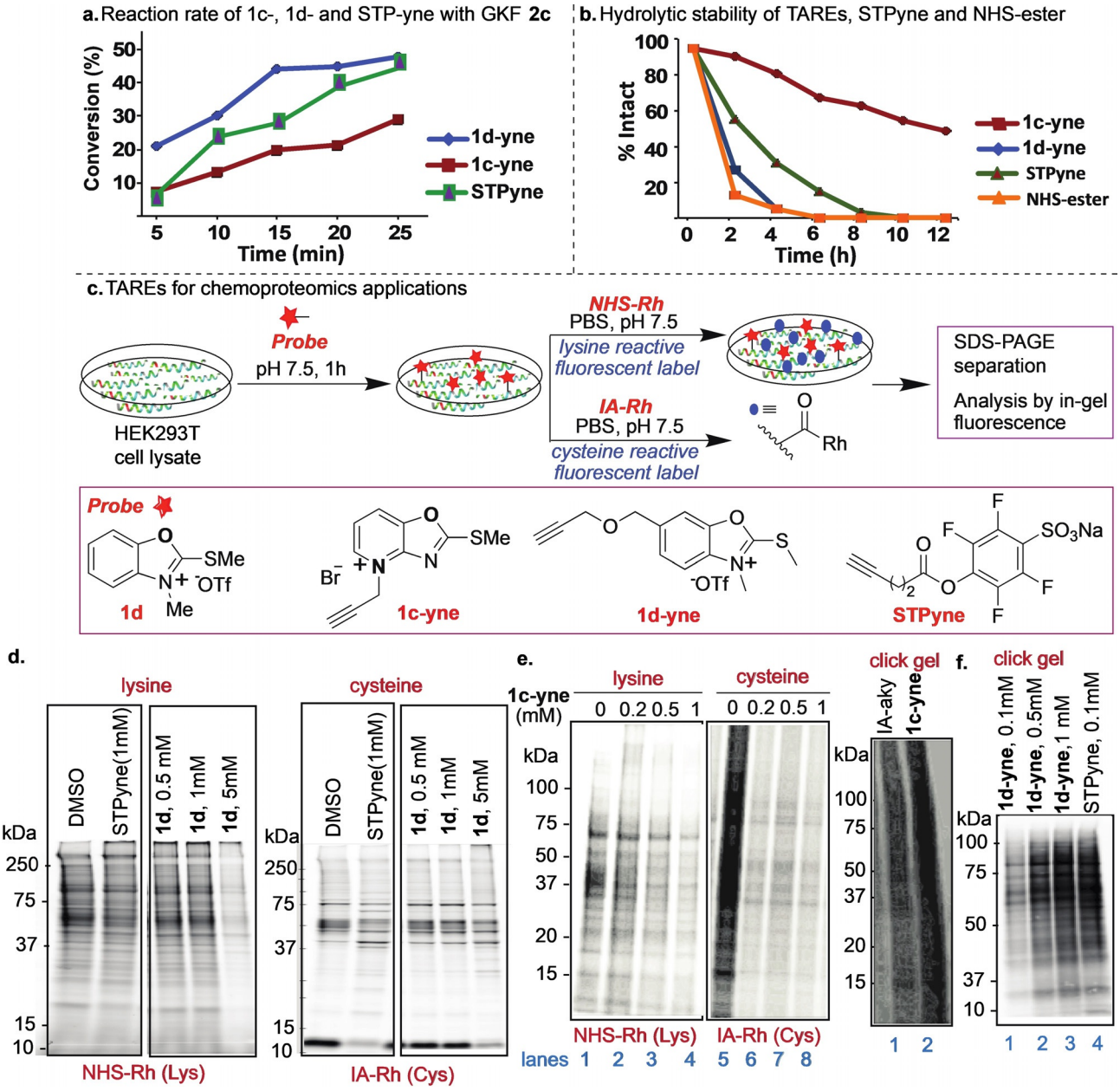


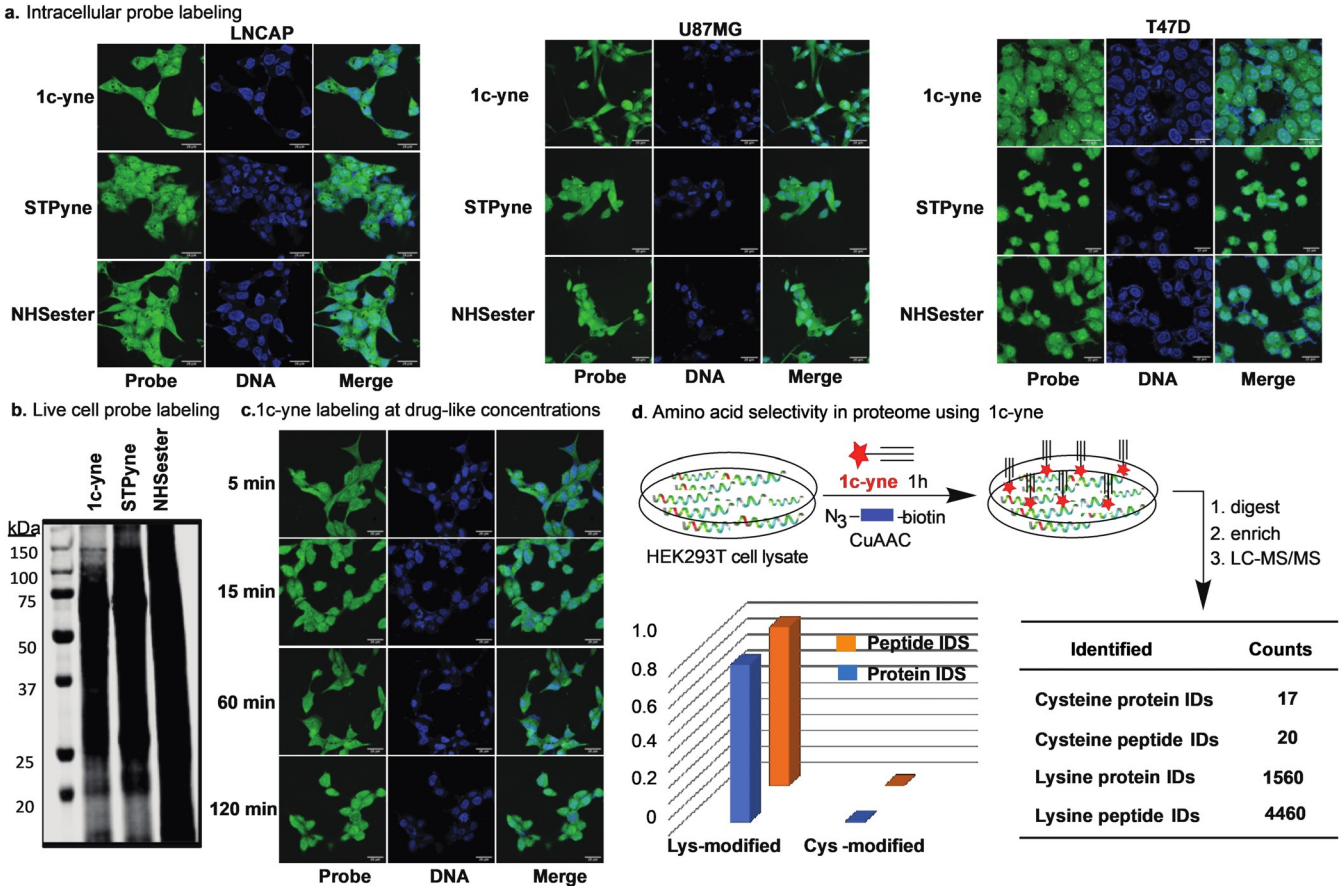
Figure 3. Chemoproteomic studies of TAREs by gel-based competitive protein profiling. a) Rate studies with peptide GKF **2c** showed high reactivity with **1d-yne**. Each time point represents an average of three independent experiments. b) Hydrolytic stability studies showed **1c-yne** is more hydrolytically stable as compared to other lysine reactive probes. c) General protocol for proteome profiling by different probes and structure of probes. d) In-gel fluorescence analysis of STPyne (1 mM) and **1d** probes at different concentrations of **1d** (0.5 mM to 5 mM) followed by detection with NHS-Rh for lysine reactivity and IA-Rh for cysteine reactivity. **1d** is more selective for lysine as compared to STPyne. e) In-gel fluorescence analysis of **1c-yne** at different concentrations (0 to 1 mM) showed reactivity with both lysine and cysteine. Click gel assay of **1c-yne** and cysteine reactive IA-aky probes with cell lysate using fluorescent-biotin azide to determine the total labeled proteins. High labeling of cell lysate was observed with **1c-yne**. f) Click gel assay of **1d-yne** at different concentrations (0.1 mM to 1 mM) and STPyne (0.1 mM) probes with cell lysate using biotin-azide and streptavidin blot to determine the total labeled proteins. Dose-dependent labeling was observed for **1d-yne** with banding pattern similar to STPyne.

lysates were subjected to labeling by **1d-yne** at the indicated concentrations followed by conjugation by CuAAC to biotin-azide, and the labeling was visualized by streptavidin blot (Figure 3 f, Supporting Information, Figure 12). Gratifyingly, robust concentration dependent labeling was observed with 1dyne and a banding pattern comparable to that of STPyne was consistent with lysine-directed reactivity. Similar labeling intensity was observed for 1 mM of **1d** when compared to 100 μM of STPyne, supporting that the **1d** TARE is a relatively attenuated warhead.

The ability of TAREs to profile the lysine proteome was further explored by treating living human cells with **1c-yne**. We incubated three cancer cell lines, LNCaP, U87MG and T47D, representing prostate, brain and breast cancers,

respectively, with increasing concentrations of **1c-yne**, NHS-ester, and STPyne. After 2 h, cells were fixed permeabilized, and washed to remove the unreacted probe. Subsequently, reacted probe was conjugated to a fluorophore via CuAAC to directly image the ability of each probe to label proteins within the cell. Confocal fluorescent microscopy confirmed that all three cell lines had taken up **1c-yne**, STPyne, and NHS-ester into both the cytosolic and nuclear compartments (Figure 4a, Supporting Information, Figure 13) at concentrations ranging from 5 μM to 100 μM (Supporting Information, Figure 13). Western blot analysis of LNCaP cells treated with 100 μM of **1c-yne**, NHS-ester, and STPyne and subsequent fluorophore labeling with CuAAC demonstrate protein labeling across molecular weights (Figure 4b, Supporting Information, Figure 14). The different band intensities may reflect the lysine selectivity of **1c-yne** compared to the cross reactivity of NHS-ester and STPyne probes with cysteine. Finally, we examined the rate at which **1c-yne** can penetrate and label unique cellular compartments in living cells. Intracellular labeling occurs within the first five minutes at **1c-yne** concentrations as low as 100 nM (Figure 4c). These results highlight the efficacy of **1c-yne** as a probe for rapid live cell labeling. Next, we carried out cell viability studies with 1c-yne using T47D cells at two different concentrations (5 μM and 20 μM) for 24 h. We did not observe any increase in apoptosis/necrosis compared to DMSO control as analyzed by flow cytometer (Supporting Information, Figure 15).

To further assess the proteome-wide reactivity profiles of the **1c** and **1d** TAREs, we turned to mass spectrometry-based chemoproteomics. Cell lysates were subjected to either probe **1c-yne** or **1d-yne**, labeled proteins were conjugated to biotin-azide via CuAAC and the samples were prepared and analyzed using our SP3 chemoproteomic sample preparation workflow (Supporting Information, Figure 16).[33] Briefly, TARE-labeled proteomes were subjected to single-pot solid-phase enhanced sample preparation (SP3) decontamination using carboxyl coated magnetic beads. SP3-resin tryptic digest biotinylated peptides were enriched by neutravidin and analyzed by LC-MS/MS. Surprisingly and in contrast with its apparent cysteine-reactivity observed by competitive gel-



Identified	Counts
Cysteine protein IDs	17
Cysteine peptide IDs	20
Lysine protein IDs	1560
Lysine peptide IDs	4460

Figure 4. Live cell labeling and amino-acid selectivity in proteome by probe **1c-yne** a) Human LNCAP, U87MG and T47D cells treated with 5 μM **1c-yne**, STPyne, or NHSester for 2 h followed by fixing the cells, washing of unreacted probes and then conjugation with azide fluorophore tags using CuAAC show labeling in multiple cellular compartments. b) Western blot fluorescent analysis of LNCAP cells incubated with 100 μM **1c-yne**, STPyne, or NHSester for 2 h and then conjugated with azide fluorophore tags using CuAAC demonstrates protein labeling across molecular weights. c) Drug-like concentrations of **1c-yne** (100 nM) demonstrate that robust intracellular labeling occurs within 5 min. d) Cellular lysates are labelled with **1c-yne** at different concentrations followed by conjugation with azide-biotin tags (blue) using CuAAC, enrichment of labeled proteins by neutravidin-conjugated beads and digested stepwise with trypsin to yield **1c-yne**-labeled peptides for LC-MS analysis. Percentage of unique peptides and proteins labeled on each nucleophilic amino acid by **1c-yne** in HEK293T proteome. Probe **1c-yne** preferentially enrich lysine residues in human cell proteomes. Data represent means ± standard deviation for two experiments.

based studies, **1c-yne** showed near complete selectivity for lysine residues (5124 total unique labeled lysine residues and 27 total unique labeled cysteine residues across two biological replicate experiments, Figure 4d, Supporting Information, Figure 16 and Supporting Information, Table 3). Similarly, chemoproteomic analysis of **1d-yne**-labeled lysates revealed near-complete lysine selectivity (1595 total unique lysine residues and 54 total unique labeled cysteine residues) across three compound concentrations analyzed.

Lysine specificity was further confirmed through re-search of the **1c-yne** datasets using the algorithm MSFragger,[34] which includes built in PTMProphet[35] for accurate mass modification localization, which identified 1560 unique labeled lysines and 17 unique labeled cysteines (Supporting Information, Figure 16, Supporting Information, Table 3). Consistent with our gel-based analysis, we observed a dose-dependent increase in peptides identified as the concentration of **1d-yne** was increased from 100 µM to 1 mM. **1c-yne** labeled substantially more peptides (5151) than **1d-yne** (1649), in contrast with the aforementioned kinetic analysis that revealed 3-fold higher reactivity for **1d–yne** ($k = 307.52\ M^{-1}S^{-1}$) as compared to **1c-yne** ($k = 99.27\ M^{-1}S^{-1}$) (Figure 3a). Surprisingly high reactivity of **1c-yne** in gel and proteomic studies as compared to **1d-yne** might be due to the high hydrolytic stability of **1c-yne** (95% intact in 2 h) as compared to **1d-yne** (28% intact in 2 h) under the reaction conditions (Figure 3b).

We were intrigued by the lack of detectable cysteine labeling observed for **1c-yne**, which was in contrast to our competitive gel-based studies. We speculated that liability of the cysteine adduct, potentially stemming from the reducing CuAAC conditions, might contribute to this observed difference. To confirm this hypothesis, we carried out labeling of a peptide GCF with **1c-yne** with and without TCEP under click chemistry conditions. We observed the modification of Cys in a peptide GCF in the absence of TCEP only, no modification of a peptide GCF was observed in the presence of TCEP under click chemistry conditions (Supporting Information, Figure 17). To further confirm, we isolated the Cys modified GCF-**1c-yne** adduct and incubated it in buffer containing TCEP. Within 5 mins, we observed complete decomposition of the GCF-**1c-yne** adduct to the unchanged peptide GCF (Supporting Information, Figure 18).

To further assess the observed differences in cysteine- and lysine-reactivity for the TAREs, we next performed DFT analysis on the reaction between **1c-yne** and methyl thiolate. Given its hydrolytic stability, observed reactivity with both cysteine and lysine residues, we selected compound **1c-yne** as optimal for our DFT studies, aiming to gain insight into the observed differences in the amino acid reactivity profile. The DFT calculations with **1c-yne** supported the reversibility of the reaction between **1c-yne** and thiol nucleophiles compared to the irreversibility of **1c-yne** modification by amine nucleophiles (Figure 5, Supporting Information, Figure 19).

Conclusion

We have systematically designed TAREs that showed high reactivity with both cysteine and lysine but reaction with cysteine generated another lysine reactive electrophile thus these probes are highly selective for the enrichment and formation of stable adducts with lysine only. These TARE chemotypes are stable under hydrolytic conditions as compared to other lysine reactive activated esters such as STPyne and NHS esters thus have the potential to act as covalent ligands for lysine in the cellular environment.

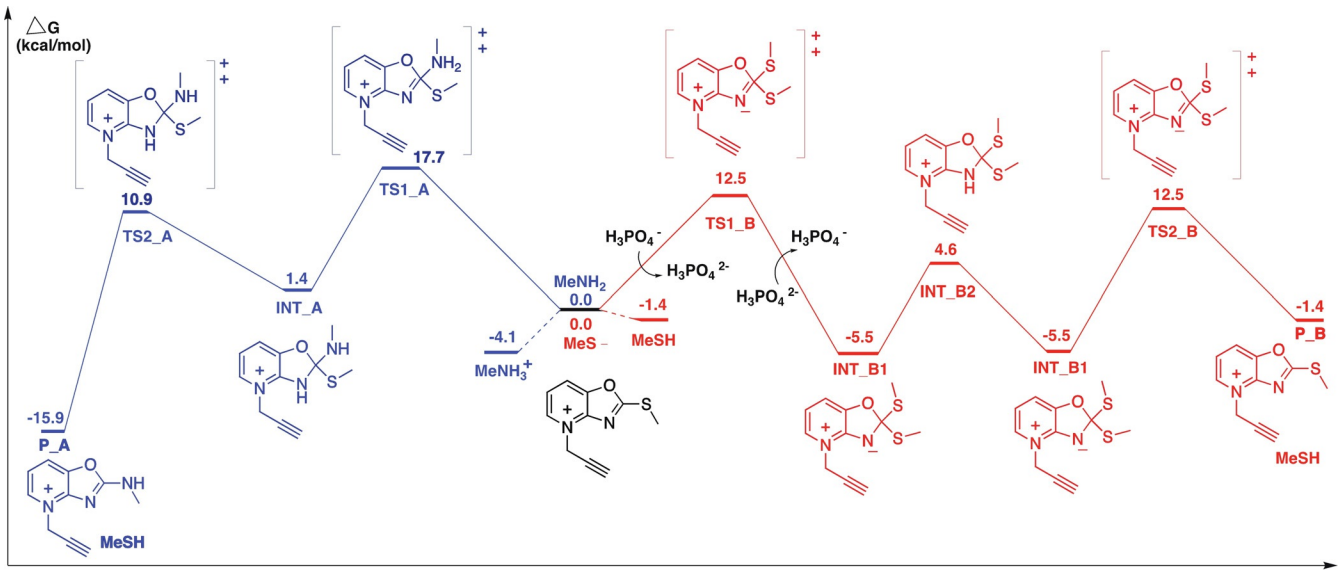


Figure 5. Free energy profiles for SNAr substitution of **1c-yne** with methylamine (blue) and methyl thiolate (red) in water, computed at the ωB97X-D/6-311++G(d,p) level of theory in SMD water. DFT method: ωB97X-D/6-311++G(d,p) SMD(H2O)//B3LYP/6-31+G(D) SMD(H2O), with Grimme correction for entropy and Head-Gordon correction for enthalpy in 298.15 K. All energies are in kcal mol^{-1}.

We show that TAREs are highly tunable, where reactivity and selectivity can be varied for desired applications by the addition of different heteroatoms and their methylation. As expected, due to the high electron withdrawing nature of O as compared to S and more electrophilicity and solubility of the charged methylated ions as compared to an uncharged moiety, **1d** and its analog **1d-yne** were the most reactive, soluble and selective for lysine. We also showed that **1c** is more hydrolytic stable as compared to other activated esters. There is no disparity between the reactivity of **1c**, **1d** and their alkyne derivatives **1c-yne** and **1d-yne**, and both of them are reactive at low micromolar concentrations. The labeling of both cytoplasmic and nuclear proteins in live cells demonstrate the high permeability, non-toxicity, cellular stability and cellular activity of TARE probes thus indicate its potential as a probe for rapid live cell labeling and covalent inhibitor. Identification of the sites of modification for TAREs showed that both **1c-yne** and **1d-yne** enriched only lysine peptide fragments from the cell lysate because the cysteine-adducts of both the probes are reactive electrophiles towards lysines and unstable towards TCEP thus unable to be enriched. DFT calculations on 1c-yne support our experimental results showing the reversibility of the reaction between **1c-yne** and thiol nucleophiles and the irreversibility of **1c-yne** modification by amine nucleophiles. 2-methylthio pyridiniumoxazoline ion **1c** and its alkyne analog **1c-yne** provides an aromatic, synthetically tractable, non-cytotoxic and hydrolytically stable electrophile to add to the arsenal of lysine-reactive groups available for protein modification. We anticipate that our design of new probes, easy synthesis to varying derivatives and detailed reactivity and selectivity studies with peptides, proteins, proteome and live cells will promote their use in multiple applications in the field of protein modification, bioconjugation, material science, activity-based protein profiling and covalent drug discovery for undruggable human proteins.

Acknowledgements

This research was supported by NIH (Grant No. 1R35GM133719-01) and Alfred P. Sloan Foundation granted to M.R, Beckman Young Investigator Award to K.M.B and NIGMS System and Integrative Biology 5T32GM008185-33 to L.M.B.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bioconjugation · chemoselective · mass sensitivity boosters · protein labeling · traceless

[1] M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner, J. W. Kozarich, *Biochemistry* **2007**, *46*, 350 – 358.
[2] A. C. Eliot, J. F. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383 – 415.
[3] G. Akcay, M. A. Belmonte, B. Aquila, C. Chuaqui, A. W. Hird, M. L. Lamb, P. B. Rawlins, N. Su, S. Tentarelli, N. P. Grimster, Q. Su, *Nat. Chem. Biol.* **2016**, *12*, 931 – 936.
[4] C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin, M. Mann, *Nat. Rev. Mol. Cell. Biol.* **2014**, *15*, 536 – 550.
[5] E. L. Greer, Y. Shi, *Nat. Rev. Genet.* **2012**, *13*, 343 – 357.
[6] K. Zhang, S. Y. Dent, *J. Cell. Biochem.* **2005**, *96*, 1137 – 1148.
[7] F. Mattiroli, T. K. Sixma, *Nat. Struct. Mol. Biol.* **2014**, *21*, 308 – 316.
[8] F. Tekaia, E. Yeramian, B. Dujon, *Gene.* **2002**, *297*, 51 – 60.
[9] D. A. Shannon, E. Weerapana, *Curr. Opin. Chem. Biol.* **2015**, *24*, 18 – 26.
[10] N. J. Pace, E. Weerapana, *ACS Chem. Biol.* **2013**, *8*, 283 – 296.
[11] Q. Liu, Y. Sabnis, Z. Zhao, T. Zhang, S. J. Buhrlage, L. H. Jones, N. S. Gray, *Chem. Biol.* **2013**, *20*, 146 – 159.
[12] J. M. Ostrem, K. M. Shokat, *Nat. Rev. Drug Discovery* **2016**, *15*, 771 – 785.
[13] N. London, R. M. Miller, S. Krishnan, K. Uchida, J. J. Irwin, O. Eidam, L. Gibold, P. Cimermancic, R. Bonnet, B. K. Shoichet, J. Taunton, *Nat. Chem. Biol.* **2014**, *10*, 1066 – 1072.
[14] D. A. Shannon, R. Banerjee, E. R. Webster, D. W. Bak, C. Wang, E. Weerapana, *J. Am. Chem. Soc.* **2014**, *136*, 3330 – 3333.
[15] E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790 – 795.
[16] L. A. Crawford, E. Weerapana, *Mol. Biosyst.* **2016**, *12*, 1768 – 1771.
[17] M. J. Hunter, M. L. Ludwig, *J. Am. Chem. Soc.* **1962**, *84*, 3491 – 3504.
[18] A. Bandyopadhyay, J. Gao, *J. Am. Chem. Soc.* **2016**, *138*, 2098 – 2101.
[19] X. Wang, A. J. Di Pasqua, S. Govind, E. McCracken, C. Hong, L. Mi, Y. Mao, J. Y.-C. Wu, Y. Tomita, J. C. Woodrick, R. L. Fine, F.-L. Chung, *J. Med. Chem.* **2011**, *54*, 809 – 816.
[20] Y. Zhang, T. W. Kensler, C. G. Cho, G. H. Posner, P. Talalay, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3147 – 3150.
[21] H.-J. Musiol, L. Moroder, *Org. Lett.* **2001**, *3*, 3859 – 3861.
[22] T. G. Kapp, M. Fottner, O. V. Maltsev, H. Kessler, *Angew. Chem. Int. Ed.* **2016**, *55*, 1540 – 1543; *Angew. Chem.* **2016**, *128*, 1564 – 1568.
[23] N. P. Grimster, S. Connelly, A. Baranczak, J. Dong, L. B. Krasnova, K. B. Sharpless, E. T. Powers, I. A. Wilson, J. W. Kelly, *J. Am. Chem. Soc.* **2013**, *135*, 5656 – 5668.
[24] Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame, J. Taunton, *J. Am. Chem. Soc.* **2017**, *139*, 680 – 685.
[25] S. Asano, J. T. Patterson, T. Gaj, C. F. Barbas 3rd, *Angew. Chem. Int. Ed.* **2014**, *53*, 11783 – 11786; *Angew. Chem.* **2014**, *126*, 11977 – 11980.
[26] M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés, G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, *140*, 4004 – 4017.
[27] C. C. Ward, J. I. Kleinman, D. K. Nomura, *ACS Chem. Biol.* **2017**, *12*, 1478 – 1483.
[28] S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia, B. F. Cravatt, *Nat. Chem.* **2017**, *9*, 1181 – 1190.
[29] M. E. Abbasov, M. E. Kavanagh, T. A. Ichu, M. R. Lazear, Y. Tao, V. M. Crowley, C. W. am Ende, S. M. Hacker, J. Ho, M. M. Dix, R. Suciu, M. M. Hayward, L. L. Kiessling, B. F. Cravatt, *Nat. Chem.* **2021**, *13*, 1081 – 1092.
[30] G. C. Adam, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2002**, *20*, 805 – 809.
[31] E. Weerapana, G. M. Simon, B. F. Cravatt, *Nat. Chem. Biol.* **2008**, *4*, 405 – 407.

[32] J. Cao, L. M. Boatner, H. S. Desai, N. R. Burton, E. Armenta, N. J. Chan, J. O. Castellon, K. M. Backus, *Anal. Chem.* **2021**, *93*, 2610 – 2618.
[33] T. Yan, H. S. Desai, L. M. Boatner, S. L. Yen, J. Cao, M. F. Palafox, Y. Jami-Alahmadi, K. M. Backus, *ChemBioChem* **2021**, *22*, 1841 – 1851.
[34] A. T. Kong, F. V. Leprevost, D. M. Avtonomov, D. Mellacheruvu, A. I. Nesvizhskii, *Nat. Methods* **2017**, *14*, 513 – 520.
[35] D. D. Shteynberg, E. W. Deutsch, D. S. Campbell, M. R. Hoopmann, U. Kusebauch, D. Lee, L. Mendoza, M. K. Midha, Z. Sun, A. D. Whetton, R. L. Moritz, *J. Proteome Res.* **2019**, *18*, 4262 – 4272.

Manuscript received: September 7, 2021
Revised manuscript received: October 19, 2021
Accepted manuscript online: November 11, 2021
Version of record online: December 16, 2021