JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Investigating the Proteome Reactivity and Selectivity of Aryl Halides

D. Alexander Shannon,[†] Ranjan Banerjee,[†] Elizabeth R. Webster,[†] Daniel W. Bak,[†] Chu Wang,[‡] and Eranthie Weerapana*,[†]

[†]Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, Massachusetts 02467, United States

[‡]Department of Chemical Biology, College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, and Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

Supporting Information

ABSTRACT: Protein-reactive electrophiles are critical to chemical proteomic applications including activity-based protein profiling, site-selective protein modification, and covalent inhibitor development. Here, we explore the protein reactivity of a panel of aryl halides that function through a nucleophilic aromatic substitution (S_NAr) mechanism. We show that the reactivity of these electrophiles can be finely tuned by varying the substituents on the aryl ring. We identify *p*-chloro- and fluoronitrobenzenes and dichlorotriazines as covalent protein modifiers at low micromolar concentrations. Interestingly, investigating the site of labeling of these electrophiles within complex proteomes identified *p*-chloronitrobenzene as highly cysteine selective, whereas the dichlorotriazine favored reactivity with lysines. These studies illustrate the diverse reactivity and amino-acid selectivity of aryl halides and enable the future application of this class of electrophiles in chemical proteomics.

The covalent modification of proteins by small molecules has widespread applications in drug design,[1,2] activity-based protein profiling,[3,4] and imaging.[5] These applications rely on the availability of an arsenal of electrophiles with tunable reactivity and selectivity.[6] Well-characterized electrophiles include haloacetamide,[7−9] maleimide,[9] and α,β-unsaturated ketones,[7] which have been shown to be highly selective for the thiol group of cysteine residues. Electrophiles that target other amino acids include sulfonate esters (aspartate, glutamate, and tyrosine),[7] fluorophosphonates (serine),[10] sulfonyl fluorides (serine, tyrosine),[11,12] and acyl phosphates (lysine).[13] In addition to functional group specificity, electrophiles also demonstrate diverse reactivities. Many of the electrophiles listed above demonstrate high reactivity and covalently modify proteins in the absence of a binding motif. The epoxysuccinates,[14] spiroepoxides,[15] carbamates,[16] acyloxymethyl ketones,[17] phenoxymethyl ketones,[18] and β-lactams[16] are milder and often require a binding element to facilitate covalent modification. Due to the widespread applications of covalent protein modification, there is a constant need for novel electrophiles with distinct and tunable reactivities and selectivities.

The proteome reactivity of aryl chloride-based electrophiles has been poorly characterized, with a few scattered examples in the literature. *p*-Chloronitrobenzenes have been incorporated into cysteine-reactive peroxisome proliferator-activated receptor (PPARγ)[19] and β-tubulin-modifying compounds.[20] 4-Halopyridines have been shown to covalently modify the active-site cysteine of dimethylarginine dimethylaminohydrolase.[21,22] Lastly, perfluoroaryl groups have been used for covalent peptide stapling.[23] In all these cases, arylation of the cysteine residue is presumed to progress via a nucleophilic aromatic substitution (S_NAr) mechanism, similar to that occurring during the enzymatic conjugation of glutathione to activated aryl groups by glutathione *S*-transferases.[24] Here, we synthesized a panel of alkyne-functionalized aryl halides and systematically evaluated the reactivity and selectivity of these electrophiles within the context of a complex proteome. The alkyne handle facilitated the use of copper-assisted azide−alkyne cycloaddition (CuAAC) for gel and mass spectrometry (MS)-based analysis of protein reactivity.[25,26]

We initiated our studies by exploring the proteome reactivity of a panel of chloronitrobenzenes (CNBs; Figure 1A). These

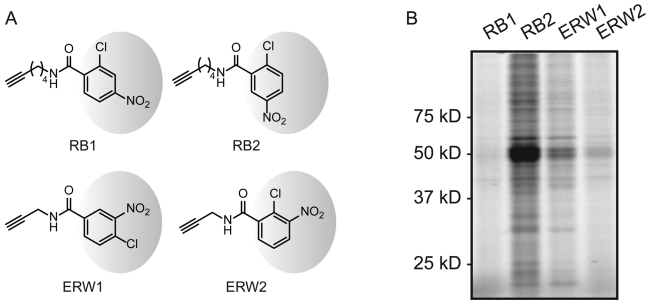


Figure 1. Evaluating the proteome reactivity of chloronitrobenzenes (CNBs). (A) The panel of alkyne-functionalized CNBs used in this study. (B) In-gel fluorescence evaluation of the reactivity of CNBs in proteomes after incorporation of Rh−N_3 using CuAAC. The corresponding Coomassie-stained gel can be found in the Supporting Information (Figure S1).

electrophiles are anticipated to react via an S_NAr mechanism, whereby the rate-limiting step is formation of the Meisenheimer or σ-complex.[27,28] Therefore, we hypothesized that the reactivity of these aryl halides can be tuned by varying the position of the electron-withdrawing nitro substituent. To test this hypothesis, we synthesized four alkyne-functionalized CNB probes with the nitro group ortho (ERW1 and ERW2), meta

Received: November 19, 2013
Published: February 18, 2014

   dx.doi.org/10.1021/ja4116204 | *J. Am. Chem. Soc.* 2014, 136, 3330−3333

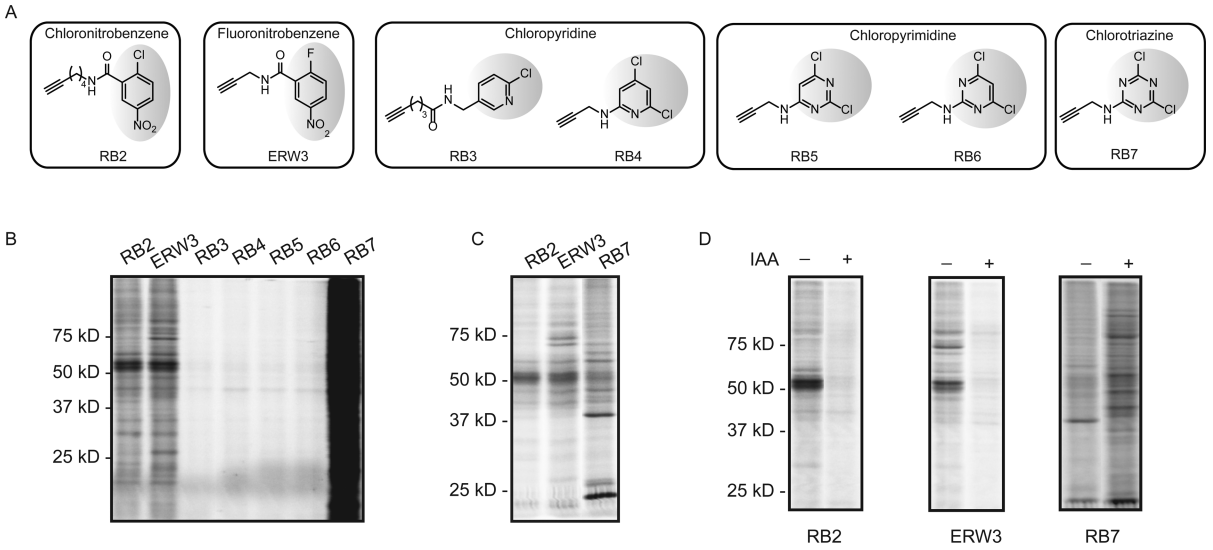


Figure 2. Investigating the proteome reactivity of aryl halides. (A) Panel of aryl halides (chloronitrobenzene, fluoronitrobenzene, chloropyridine, chloropyrimidine, and chlorotriazine) analyzed in this study. (B) In-gel fluorescence image of probe labeling in HeLa cell lysates treated with 100 μM of each probe, tagged with Rh-N_3 using click chemistry, and separated by SDS-PAGE. (C) Comparison of probe labeling of RB7 (5 μM) with RB2 and ERW3 (20 μM). (D) Effect of iodoacetamide (IAA) treatment (1 mM) prior to labeling with RB2, ERW3, and RB7. Coomassie-stained gels can be found in the Supporting Information (Figures S2–S4).

(RB1), and para (RB2) to the chloro substituent (Figure 1A; Scheme S1). We incubated HeLa cell lysates with each probe at 100 μM concentrations for 1 h, after which we appended a fluorescent rhodamine-azide (Rh-N_3) using CuAAC, separated the proteins on SDS-PAGE, and visualized protein labeling using in-gel fluorescence (Figure 1B). As anticipated, the *p*-nitro-substituted CNB was the most reactive (Figure 1B, lane 2) due to resonance stabilization of the Meisenheimer complex by the electron-withdrawing nitro group. Shifting the nitro substituent to the meta position (RB1) fully abrogated reactivity (Figure 1B, lane 1), whereas the ortho-substituted ERW1 and ERW2 showed protein reactivity (Figure 1A, lanes 3 and 4), albeit less than for RB2. These data demonstrate that the proteome reactivity of CNBs can be finely adjusted by modulating the electronics and sterics of the aryl ring system, thereby providing an ideal tunable electrophile for chemical proteomic applications.

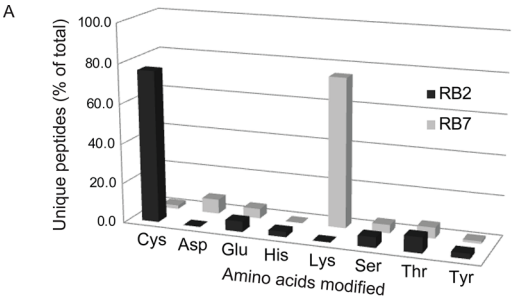
To compare the reactivity of CNBs to that of other aryl halides, we expanded our panel to include a fluoronitrobenzene (ERW3), chloropyridines (RB3 and RB4), chloropyrimidines (RB5 and RB6), and a chlorotriazine (RB7) (Figure 2A). The extent of resonance stabilization provided by the electron-withdrawing nitro substituent or the nitrogen(s) present within the ring itself will determine the reactivity of each of these aryl halides.[27,28] Since we identified RB2 as the most reactive of the CNBs, we compared each probe to RB2 in terms of their reactivity (Figure 2B, lane 1). Despite the presence of the poorer fluorine leaving group, ERW3 exhibited reactivity similar to that of RB2 (Figure 2B, lane 2). The chloropyridines (RB3 and RB4; Figure 2B, lanes 3 and 4) and the chloropyrimidines (RB5 and RB6; Figure 2B, lanes 5 and 6) showed significantly decreased reactivity relative to RB2. Interestingly, the dichlorotriazine RB7 exhibited very potent labeling (Figure 2B, lane 7), suggesting that the stabilization of the S_NAr transition state due to the presence of the third ring nitrogen is greater than that provided by the *p*-nitro group of RB2.

This initial screening identified RB2, ERW3, and RB7 as aryl halides with high proteome reactivity. To better visualize the pattern of labeling across these three probes, we reduced the concentration of the highly reactive RB7 probe to 5 μM (Figure 2C, lane 3) and compared it to RB2 and ERW3 at 20 μM (Figure 2C, lanes 1 and 2). This comparison revealed that the majority of the bands for RB7 do not coincide with those of RB2 and ERW3, suggesting that this dichlorotriazine probe was targeting a subset of proteins distinct from the halonitrobenzenes. We then proceeded to interrogate the cysteine selectivity of these electrophiles by competition with a known highly reactive, cysteine-selective electrophile, iodoacetamide (IAA). To achieve this competition, we pretreated HeLa lysates with IAA (0 and 2 mM), followed by labeling with RB2 (20 μM), ERW3 (20 μM), or RB7 (1 μM) (Figure 2D). IAA treatment almost completely abolished RB2 and ERW3 labeling, suggesting the selective modification of cysteine residues by these two electrophiles. In contrast, the majority of RB7-labeled proteins remained after IAA treatment. In fact, labeling of some bands was enhanced by IAA, likely due to partial denaturation and exposure of previously inaccessible sites under high IAA concentrations. These data thereby allude to the fact that RB7 predominantly targets amino acids other than cysteine within the proteome. We also administered the less reactive probes (RB1, RB3, RB4, RB5, and RB6) at high millimolar concentrations and show that labeling by these probes is similar to that of RB2 and therefore cysteine-selective (Figures S5–S9).

To confirm the cysteine selectivity of RB2 and identify the amino-acid targets of RB7, we employed a MS platform geared to identify the site of probe labeling within a complex proteome. This method, termed tandem orthogonal proteolysis activity-based protein profiling (TOP-ABPP),[29,30] allows for the specific enrichment of probe-labeled peptides for identification of the site of labeling. Mouse liver and HeLa lysates were treated with RB2 or RB7 (100 μM) and analyzed by TOP-ABPP. The resulting MS data were analyzed for probe modifications on all nucleophilic amino acids (cysteine, aspartate, glutamate, histidine, lysine, serine, threonine, and tyrosine) using the SEQUEST algorithm.[31] The percentage of

dx.doi.org/10.1021/ja4116204 | *J. Am. Chem. Soc.* 2014, 136, 3330–3333

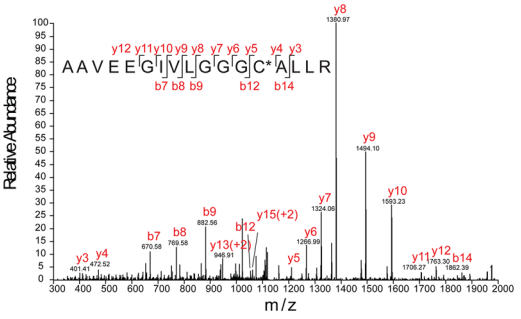
unique peptides modified for each reactive amino acid was plotted for both RB2 and RB7 (Figure 3A). In agreement with



B

Probe	Gene	Protein Name	Peptide
RB2	*Mm* ALDH6A1	Aldehyde dehydrogenase family 6A1	R.C*MALSTAILVGEAK.K
	Mm PCK1	Phosphoenolpyruvate carboxykinase	K.YLAAAFPSAC*GK.T
	Hs HSPD1	60 kDa heat shock protein	R.AAVEEGIVLGGGC*ALLR.C
	Hs EEF2	Elongation factor 2	K.STLTDSLVC*K.A
	Hs RPS3	40S ribosomal protein S3	R.GLC*AIAQAESLR.Y
RB7	*Mm* GNMT	Glycine N-methyltransferase	K.SDLTK*DITTSVLTVNNK.A
	Mm GSTA3	Glutathione S-transferase A3	K.LVQTK*AILNYIASK.Y
	Hs EEF1A1	Elongation factor 1-alpha 1	R.VETGVLK*PGMVVTFAPVNVTTEVK.S
	Hs VIL2	Ezrin	R.AK*FYPEDVAEELIQDITQK.L
	Hs HSPA1	70 kDa heat shock protein 1	R.LIGDAAK*NQVALNPQNTVFDAK.R

C



D

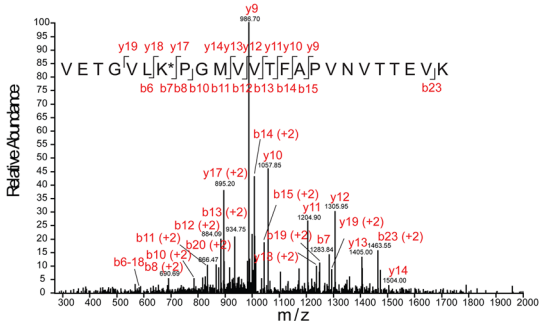


Figure 3. Characterizing amino-acid selectivity in proteomes. (A) Percentage of unique peptides labeled on each nucleophilic amino acid by RB2 and RB7 (100 μM) in mouse liver (RB2, n = 3; RB7, n = 5) and HeLa (RB2, n = 2; RB7, n = 4) proteomes. (B) A subset of RB2- and RB7-labeled peptides identified in the proteomic studies using TOP-ABPP. (C) Annotated MS2 of RB2-labeled *Hs*HSPD1 peptide. (D) Annotated MS2 of RB7-labeled *Hs*EEF1A1 peptide.

our IAA competition study (Figure 2C), RB2 primarily labels cysteine residues (Figure 3A). In contrast, RB7 shows very high reactivity with lysine, with minimal cysteine modification observed (Figure 3A).

Closer investigation of the probe-labeled peptides identified for RB2 and RB7 showed targeting of diverse proteins (Figure 3B; Tables S1−S3). RB2 targets known active-site nucleophiles of aldehyde dehydrogenase (ALDH6A1) and phosphoenol pyruvate carboxykinase (PCK1), alluding to the use of this electrophile to target functional cysteines in the proteome. To further explore the targets of RB7, we performed an isoTOP-ABPP analysis,[8] whereby we compared labeling at low (10 μM) and high (100 μM) concentrations of RB7 to rank the labeled lysine residues in order of reactivity (Table S4). Within these reactive lysines, we identify many that are annotated as sites of acetylation, as well as known active sites (Glud1) or ATP-binding (Nme1/2) sites. The acetylation sites are generally surface exposed, whereas the active-site residues are found to be buried within the protein core (Figures S12−S14). We subsequently validated selected labeled peptides by manually calculating theoretical y- and b-ions and annotating the associated MS2 spectra for both RB2 and RB7-labeled peptides (Figures 3C,D and S5). Additionally, we performed probe labeling *in situ* with both RB2 and RB7 probes and found that these electrophiles show similar protein labeling patterns when administered to living cells (Figure S10).

Previous studies with carbon electrophiles such as chloroacetamides and sulfonate esters demonstrated that the solution reactivity of electrophiles is often not predictive of reactivity observed in a proteome.[7] This is likely due to the unique protein microenvironment that serves to modulate the pK_a and reactivity of amino-acid side chains. We sought to determine if the lysine selectivity observed for RB7 in proteomes is reflected in solution reactivity with free amino acids. To achieve this, we incubated RB2 and RB7 with a 20-fold excess of C- and N-terminal protected lysine and cysteine amino-acid derivatives and analyzed the formation of the corresponding adducts by LC-MS. We observed formation of the RB2-Cys adduct but could not detect the corresponding lysine adduct (Figure 4A). In contrast, RB7 formed adducts

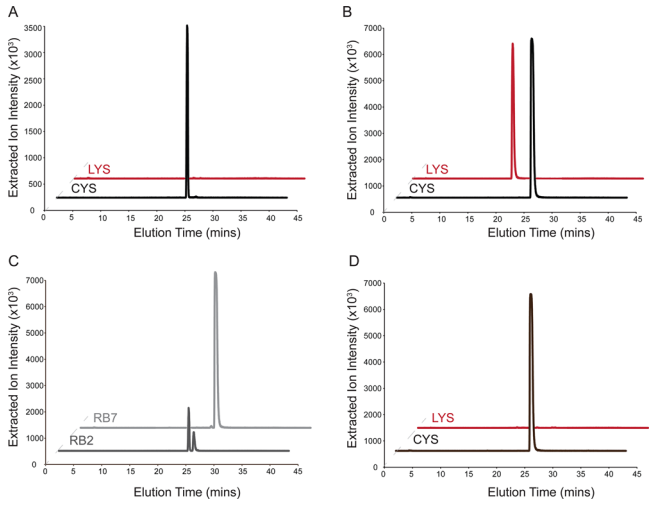


Figure 4. Reactivity of RB2 and RB7 with protected amino-acid derivatives in solution (12-h incubation in PBS). (A) RB2 with 20-fold excess cysteine or lysine derivatives. (B) RB7 with 20-fold excess cysteine or lysine derivatives. (C). RB2 and RB7 were mixed together and treated with 20-fold excess cysteine derivative. (D) RB7 was concurrently treated with 10-fold excess of both cysteine and lysine derivatives.

with both cysteine and lysine (Figure 4B). To further evaluate the cysteine reactivity of RB2 and RB7, both probes were simultaneously incubated with excess cysteine (Figure 4C). This experiment showed predominantly the RB7-Cys adduct, implying that RB7 was more reactive than RB2. Lastly, when a limiting amount of RB7 was incubated with excess cysteine and lysine derivatives, only the RB7-Cys adduct was observed

dx.doi.org/10.1021/ja4116204 | *J. Am. Chem. Soc.* 2014, 136, 3330−3333

(Figure 4D), indicating that while RB7 can react with free amines at biologically relevant pH, the thiol adduct is the favored product. These studies serve to highlight the disparity often observed between the solution reactivity and the proteome reactivity of electrophiles. In this case, RB2 prefers cysteine in solution and in proteomes, whereas RB7 appears to preferentially react with cysteines in solution but with lysine in the context of a proteome.

In summary, we have systematically evaluated the reactivity and amino-acid selectivity of a panel of aryl halides that function through an S_NAr mechanism. We show that chloronitrobenzenes are highly tunable electrophiles, where reactivity can be matched for a desired application by modifications to the steric and electronic properties of the aryl ring system. As expected, due to resonance stabilization of the S_NAr intermediate, the *p*-chloronitrobenzenes were the most reactive. We also evaluated the proteome reactivity of fluoronitrobenzenes, halopyridines, halopyrimidines, and dichlorotriazines, which have been poorly characterized and utilized in proteomic applications. We show that the *p*-halonitrobenzenes and dichlorotriazine are reactive at low micromolar concentrations. Identification of the sites of modification for these aryl halides in proteomes demonstrated that the *p*-chloronitrobenzene preferentially reacts with cysteine residues whereas dichlorotriazine prefers lysine. The dichlorotriazine provides an aromatic, synthetically tractable, and hydrolytically stable electrophile to add to the arsenal of lysine-reactive groups available for protein modification. Lastly, we show that the observed amino-acid selectivity is not universally mimicked when we expose the electrophiles to free amino acids in solution, underscoring the importance of studying reactivity and selectivity in the context of a proteome. We anticipate that our studies into the proteome reactivity of aryl halides will promote the educated use of these electrophiles in chemical proteomics, where the reactivity and amino-acid selectivity can be judiciously matched to the desired application.

■ ASSOCIATED CONTENT

Supporting Information

Detailed synthetic and experimental procedures, NMR and HR-MS characterization of probes, supplementary figures, and tables containing complete results of TOP-ABPP experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

eranthie@bc.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

E.W. is a Damon Runyon-Rachleff Innovator (DRR-18-12). We are grateful for financial support from the Smith Family Foundation and Boston College. We thank members of the Weerapana laboratory for comments and critical reading of the manuscript and Bo Li for acquiring crystallography data.

■ REFERENCES

(1) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug Discov.* **2011**, *10*, 307.
(2) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. *Future Med. Chem.* **2010**, *2*, 949.
(3) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem.* **2008**, *77*, 383.
(4) Fonovic, M.; Bogyo, M. *Exp. Rev. Proteomics* **2008**, *5*, 721.
(5) Edgington, L. E.; Verdoes, M.; Bogyo, M. *Curr. Opin. Chem. Biol.* **2011**, *15*, 798.
(6) Haedke, U.; Kuttler, E. V.; Vosyka, O.; Yang, Y.; Verhelst, S. H. *Curr. Opin. Chem. Biol.* **2013**, *17*, 102.
(7) Weerapana, E.; Simon, G. M.; Cravatt, B. F. *Nat. Chem. Biol.* **2008**, *4*, 405.
(8) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. *Nature* **2010**, *468*, 790.
(9) Shin, N. Y.; Liu, Q.; Stamer, S. L.; Liebler, D. C. *Chem. Res. Toxicol.* **2007**, *20*, 859.
(10) Kidd, D.; Liu, Y.; Cravatt, B. F. *Biochemistry* **2001**, *40*, 4005.
(11) Gu, C.; Shannon, D. A.; Colby, T.; Wang, Z.; Shabab, M.; Kumari, S.; Villamor, J. G.; McLaughlin, C. J.; Weerapana, E.; Kaiser, M.; Cravatt, B. F.; van der Hoorn, R. A. *Chem. Biol.* **2013**, *20*, 541.
(12) Shannon, D. A.; Gu, C.; McLaughlin, C. J.; Kaiser, M.; van der Hoorn, R. A.; Weerapana, E. *Chembiochem* **2012**, *13*, 2327.
(13) Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W. *Biochemistry* **2007**, *46*, 350.
(14) Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogyo, M. *Chem. Biol.* **2000**, *7*, 569.
(15) Evans, M. J.; Morris, G. M.; Wu, J.; Olson, A. J.; Sorensen, E. J.; Cravatt, B. F. *Mol. Biosyst.* **2007**, *3*, 495.
(16) Bachovchin, D. A.; Ji, T.; Li, W.; Simon, G. M.; Blankman, J. L.; Adibekian, A.; Hoover, H.; Niessen, S.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 20941.
(17) Kato, D.; Boatright, K. M.; Berger, A. B.; Nazif, T.; Blum, G.; Ryan, C.; Chehade, K. A.; Salvesen, G. S.; Bogyo, M. *Nat. Chem. Biol.* **2005**, *1*, 33.
(18) Verdoes, M.; Oresic Bender, K.; Segal, E.; van der Linden, W. A.; Syed, S.; Withana, N. P.; Sanman, L. E.; Bogyo, M. *J. Am. Chem. Soc.* **2013**, *135*, 14726.
(19) Leesnitzer, L. M.; Parks, D. J.; Bledsoe, R. K.; Cobb, J. E.; Collins, J. L.; Consler, T. G.; Davis, R. G.; Hull-Ryde, E. A.; Lenhard, J. M.; Patel, L.; Plunket, K. D.; Shenk, J. L.; Stimmel, J. B.; Therapontos, C.; Willson, T. M.; Blanchard, S. G. *Biochemistry* **2002**, *41*, 6640.
(20) Banerjee, R.; Pace, N. J.; Brown, D. R.; Weerapana, E. *J. Am. Chem. Soc.* **2013**, *135*, 2497.
(21) Johnson, C. M.; Monzingo, A. F.; Ke, Z.; Yoon, D. W.; Linsky, T. W.; Guo, H.; Robertus, J. D.; Fast, W. *J. Am. Chem. Soc.* **2011**, *133*, 10951.
(22) Johnson, C. M.; Linsky, T. W.; Yoon, D. W.; Person, M. D.; Fast, W. *J. Am. Chem. Soc.* **2011**, *133*, 1553.
(23) Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y. S.; Pentelute, B. L. *J. Am. Chem. Soc.* **2013**, *135*, 5946.
(24) Ji, X.; Armstrong, R. N.; Gilliland, G. L. *Biochemistry* **1993**, *32*, 12949.
(25) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.
(26) Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535.
(27) Miller, J. *J. Am. Chem. Soc.* **1963**, *85*, 1628.
(28) Bernasconi, C. F. *Chimia* **1980**, *34*, 1.
(29) Speers, A. E.; Cravatt, B. F. *J. Am. Chem. Soc.* **2005**, *127*, 10018.
(30) Weerapana, E.; Speers, A. E.; Cravatt, B. F. *Nat. Protoc.* **2007**, *2*, 1414.
(31) Eng, J. K.; Mccormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.