Disparate proteome reactivity profiles of carbon electrophiles

Eranthie Weerapana, **Gabriel M. Simon**, and **Benjamin F. Cravatt**[*]
The Skaggs Institute for Chemical Biology and Department of Chemical Physiology, The Scripps Research Institute, 10550 N. Torrey Pines Road. La Jolla, CA 92037

Abstract

Insights into the proteome reactivity of electrophiles are crucial for designing activity-based probes for enzymes lacking cognate affinity labels. Here, we show that different classes of carbon electrophiles exhibit markedly distinct amino acid labeling profiles in proteomes, ranging from selective reactivity with cysteine to adducts with several amino acids. These data thus specify electrophilic chemotypes with restricted and permissive reactivity profiles to guide the tailored design of next-generation functional proteomics probes.

The field of activity-based protein profiling (ABPP) applies reactive chemical probes to profile the functional state of enzymes in native proteomes[1]. Original ABPP probes incorporated well-defined affinity labels as reactive groups to target enzyme classes such as the serine[2] and cysteine[3] hydrolases. Many enzymes, however, do not possess cognate affinity labels, and the design of ABPP probes for these proteins remains challenging. Structural insights into the substrate-binding pocket of enzyme classes can reveal nucleophilic residues for targeting with appropriate electrophiles. Recent work in the design of protein kinase probes positioned α-fluoromethyl ketone and acyl-phosphate electrophiles within an adenosine triphosphate (ATP) scaffold to exploit the nucleophilicity of proximal cysteine[4] and lysine[5] residues respectively. Differentiating among electrophilic chemotypes that show restricted and permissive amino acid reactivity profiles should streamline such endeavors to design ABPP probes for a wide range of enzyme classes.

A variety of electrophiles are available for incorporation into ABPP probes. The proteome reactivity profiles of iodoacetamide and maleimide reactive groups have been extensively investigated[6]. Here, we expand on these studies by investigating the reactivity of a panel of carbon electrophiles (Fig. 1a), comprising a phenylsulfonate ester (SE, **1**), linear- (EP, **2**) and spiro-epoxides (SP, **5**), an α-chloroacetamide (CA, **3**) and an α,β-unsaturated ketone (UK, **4**) in complex proteomes. An alkyne was incorporated into these electrophilic frameworks to provide a click chemistry handle for gel and mass spectrometric analysis[7]. Application of these electrophiles to a soluble mouse proteome, followed by click chemistry with a rhodamine azide (Rh-N$_3$) reporter tag and visualization of labeled proteins by SDS-PAGE and in-gel fluorescence scanning, demonstrated that the panel of electrophiles exhibit a range of protein reactivities (see Supplementary Information Fig. 1). Highest reactivity was observed for the UK probe, which demonstrated substantial protein labeling at concentrations as low as 1 μM. The CA and SE electrophiles demonstrated moderate levels of reactivity, whereas, the EP and SP probes displayed little to no protein labeling even at concentrations up to 20 μM.

We then examined in greater depth the protein and amino-acid labeling profiles for the three probes that displayed the highest levels of proteome reactivity (SE, CA and UK). To address this question, we utilized a mass spectrometry platform referred to as tandem orthogonal proteolysis (TOP)–ABPP for simultaneous identification of protein targets and exact sites of probe modification[8]. The probes were applied to four different mouse tissue proteomes (soluble fractions of heart, kidney, and liver, and the membrane fraction of liver; 50 μM probe, 2 hrs, PBS, pH 7.4; n = 2 per tissue). The tandem MS datasets generated by TOP-ABPP were analyzed by the SEQUEST algorithm, specifying a differential modification corresponding to the masses of each probe on nine potentially nucleophilic amino acids. Assignments were screened for peptides uniquely labeled on a single amino acid residue. A very restricted reactivity profile was observed for the CA and UK probes, which selectively labeled cysteine residues in the proteomes (Fig. 1b). In contrast, the SE probe demonstrated unique labeling events on several amino acids, including aspartate, glutamate, cysteine, tyrosine and histidine residues (Fig. 1b).

We next asked whether the distinct proteome labeling profiles of the SE, CA and UK probes could be discerned from their reactivities with isolated amino acids in solution. Each probe was exposed to 20 equivalents of amino acid derivatives under buffer conditions that mimic the proteomic environment. Product formation was monitored by LC/MS and revealed similar reactivity profiles for the SE and CA probes (Fig. 1c), with cysteine adducts representing the predominant products and minor amounts of the corresponding histidine, lysine, and aspartate adducts also being observed. Of particular interest was the relatively low reactivity of the SE probe with carboxylic acid groups in solution (Fig. 1c), which contrasted sharply with the large number of aspartate and glutamate labeling sites observed for this probe in proteomes (Fig. 1b). In contrast, the UK probe displayed significant reactivity with cysteine, lysine and histidine groups in solution (Fig. 1c), but only showed adducts with cysteine residues on proteins within complex proteomes (Fig. 1b). These data thus indicate that electrophilic probes can display unanticipated trends of reactivity in proteomes that are not easily extrapolated from solution studies.

To further explore the disparate reactivity of these probes, we analyzed the labeling profile of aldehyde dehydrogenase-1 (ALDH-1), an enzyme containing both a cysteine nucleophile (C303) and catalytic glutamate base (E269). Each catalytic residue was mutated to alanine and the probe reactivity profiles of mutant enzymes were compared to wild type (WT) ALDH-1 following transient transfection in COS-7 cells. Consistent with previous studies[9], the SE probe was found to label both WT- and C303A mutant ALDH-1, but not the E269A mutant (Fig. 2a). Interestingly, the cysteine-reactive CA probe showed the opposite profile, labeling both WT and E269A ALDH-1, but not the C303A mutant (Fig. 2b). These data thus demonstrate that different classes of carbon electrophiles can display mutually exclusive amino acid reactivity profiles, even when placed within the same enzyme active site.

A survey of the residues labeled by the SE, CA and UK probes revealed a strong enrichment for functional residues that play roles in catalysis, substrate binding, and post-translational regulation. Representative examples of these functional residues are listed in Table 1 (see Supplementary Information, Table 1 for an expansive list of functional residues labeled and Supplementary Information Tables 2, 3 and 4 for complete lists of labeling sites for SE, CA, and UK probes, respectively). SE-labeled proteins include the previously reported enoyl coenzyme A hydratase-1 and the acyl-CoA dehydrogenases, which were labeled on catalytic glutamate and aspartate bases, respectively[9]. In addition, the expansive tissue profiling demonstrated several tyrosine specific labeling events, including the active site tyrosine of corticosteroid 11-beta dehydrogenase and a tyrosine residue from the dual specificity tyrosine-phosphorylation regulated kinase 1A that is known to be dynamically phosphorylated[10]. The proteome coverage of the CA and UK probes spanned a variety of

enzymes with cysteine nucleophiles including fatty acid synthase, UDP-glucose-6-dehydrogenase, and multiple nitrilases. Additionally, labeling was observed on several metal-coordinating residues, such as cysteines that bind iron and zinc in aconitate hydratase, betaine-homocysteine S-methyltransferase and alcohol dehydrogenase. Cysteine residues within nucleotide binding domains were also labeled, including GTP- and NAD-binding residues in phosphoenolpyruvate carboxykinase, lactate dehydrogenase and isocitrate dehydrogenase. Interestingly, more than half of the proteins targeted by the CA and UK probes were exclusively labeled by one of these two agents (Supplementary Information Fig. 2 and Supplementary information Table 5), indicating that different cysteine-reactive electrophiles target specific subsets of the proteome. All three probes labeled cysteine residues involved in oxidative regulation, exemplified by hemoglobin, glyceraldehyde-3-phosphate dehydrogenase and peroxiredoxin, which were modified on known sites of oxidation or *S*-nitrosylation[11].

The proteome coverage of the panel of carbon electrophiles included several proteins previously inaccessible to the current suite of ABPP probes. Interestingly, some of these proteins, such as the chloride intracellular channel (CLIC) family, were labeled on conserved residues implicated in the post-translational regulation of activity[12]. CLIC proteins are ion channels that are able to assume both soluble and membrane-bound forms and are distantly related to the glutathione S-transferase superfamily[12, 13]. Our mass spectrometric data revealed that the CA probe labeled CLIC4 at C35, a conserved residue among CLIC family members that has been shown to be a site for nitrosylation and oxidative regulation[12]. To test whether the CA probe might serve as a general profiling tool for CLICs, we overexpressed three members of the mouse CLIC family, CLIC1, CLIC4 and CLIC5 with a C-terminal myc/His tag in COS-7 cells. All three CLICs were labeled by the CA probe in transfected cell proteomes (Fig. 2c, panel 1). A C35S mutant of CLIC4 was not labeled by the CA probe (Fig 2c), confirming that reactivity occurred specifically on the conserved C35 residue. Pre-treatment of cell lysates with oxidizing agents such as nitric oxide or oxidized glutathione resulted in substantial reduction of labeling of the CLICs, supporting that the conserved cysteine residue targeted by the CA probe is susceptible to oxidation (Fig. 2d, Supplementary Information Fig. 3). Notably, the CA probe also proved capable of labeling CLICs in living cells (Fig. 2c, panel 2), which suggests a potential route to monitor the post-translational regulation of these proteins *in vivo*.

In order to expand the number of protein classes addressable by ABPP, structural knowledge of active sites needs to be paired with an understanding of small-molecule reactivity. Here, we show that distinct classes of reactive carbon electrophiles demonstrate widely divergent amino acid preferences in proteomes. The promiscuity of the SE probe designates it as a highly versatile electrophile for ABPP, as well as potentially related chemical biology endeavors, such as ligand-guided protein surface labeling[14], which aims to convert reversible ligands into covalent probes by proximity-induced reactivity with nucleophilic amino acids neighboring protein active sites. One could envision improving the target selectivity of SE probes by combining this electrophile with high-affinity binding groups for individual proteins of interest. In contrast, the CA and UK probes, by displaying selective reactivity with cysteine residues, constitute powerful electrophiles for enzymes that require this amino acid for function and/or post-translational regulation. Furthermore, the bias toward cysteine reactivity demonstrated by the CA and UK groups suggests a similar reactivity profile for structurally related, but less electrophilic groups, such as the acrylamide and α-fluoromethyl ketone. These latter electrophiles could prove particularly useful for generating target-selective ABPP probes, as has been demonstrated for the EGF receptor[15] and RSK kinases[4], respectively. More generally, the preponderance of functional residues labeled by carbon electrophiles in proteomes suggests that these sites display sufficiently enhanced nucleophilicity, possibly dictated by local protein microenvironment,

to be experimentally discriminated from the large excess of nonfunctional residues in the proteome. Future ABPP studies incorporating the SE, CA, and UK electrophiles, as well as others, into substrate and/or inhibitor scaffolds should facilitate the development of functional proteomics probes for a wide range of proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Evans MJ, Cravatt BF. Chem Rev 2006;106:3279–301. [PubMed: 16895328]
2. Kidd D, Liu Y, Cravatt BF. Biochemistry 2001;40:4005–15. [PubMed: 11300781]
3. Kato D, et al. Nat Chem Biol 2005;1:33–8. [PubMed: 16407991]
4. Cohen MS, Zhang C, Shokat KM, Taunton J. Science 2005;308:1318–21. [PubMed: 15919995]
5. Patricelli MP, et al. Biochemistry 2007;46:350–8. [PubMed: 17209545]
6. Shin NY, Liu Q, Stamer SL, Liebler DC. Chem Res Toxicol 2007;20:859–67. [PubMed: 17480101]
7. Speers AE, Cravatt BF. Chem Biol 2004;11:535–46. [PubMed: 15123248]
8. Speers AE, Cravatt BF. J Am Chem Soc 2005;127:10018–9. [PubMed: 16011363]
9. Adam GC, Burbaum J, Kozarich JW, Patricelli MP, Cravatt BF. J Am Chem Soc 2004;126:1363–8. [PubMed: 14759193]
10. Rush J, et al. Nat Biotechnol 2005;23:94–101. [PubMed: 15592455]
11. Forrester MT, Foster MW, Stamler JS. J Biol Chem 2007;282:13977–83. [PubMed: 17376775]
12. Littler DR, et al. FEBS J 2005;272:4996–5007. [PubMed: 16176272]
13. Cromer BA, Morton CJ, Board PG, Parker MW. Eur Biophys J 2002;31:356–64. [PubMed: 12202911]
14. Chen G, et al. J Am Chem Soc 2003;125:8130–3. [PubMed: 12837082]
15. Fry DW, et al. Proc Natl Acad Sci U S A 1998;95:12022–7. [PubMed: 9751783]
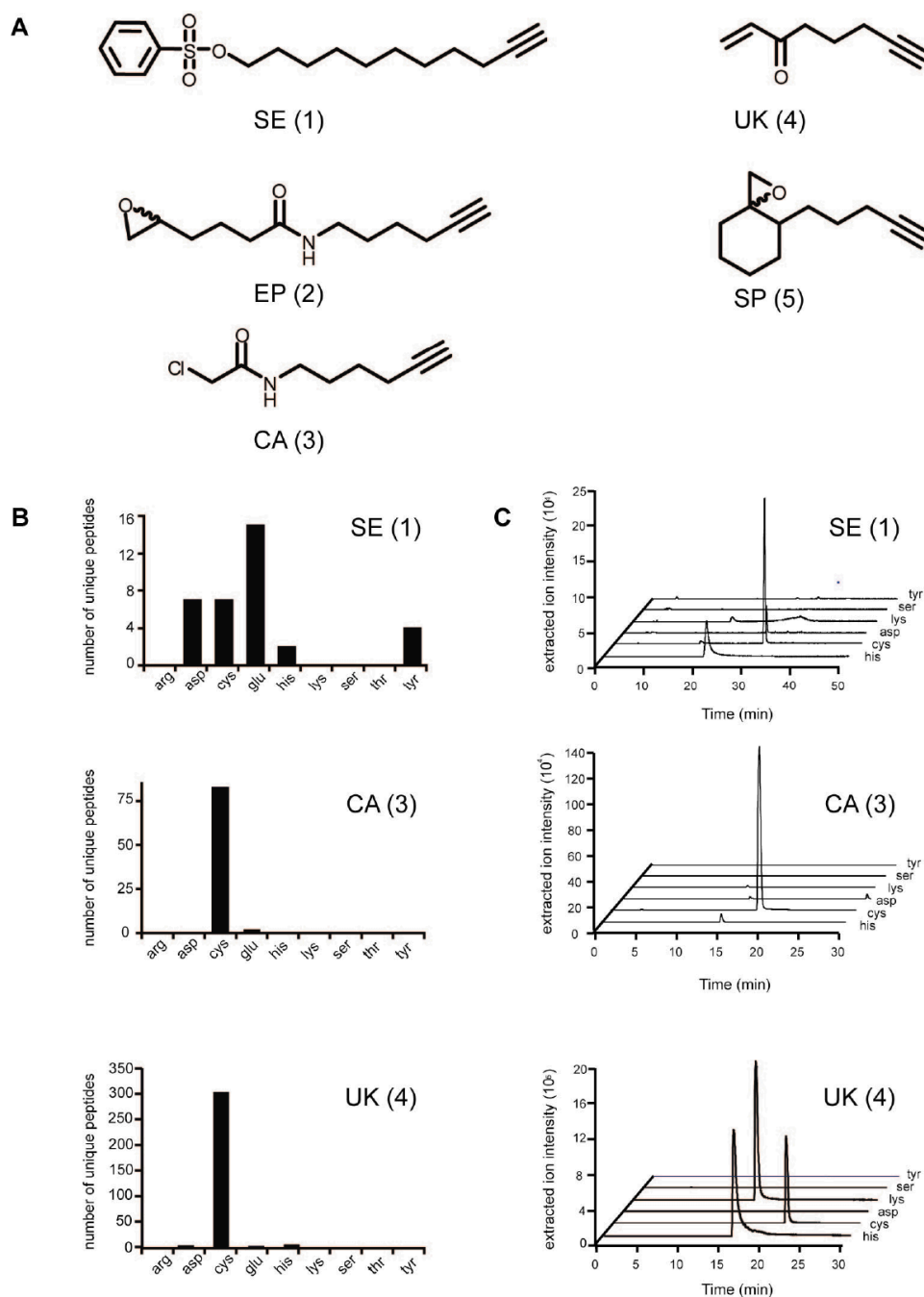
Figure 1.
Proteome and solution reactivities of carbon electrophile probes. (**a**), Panel of probes utilized in this study. (**b**), Proteome reactivity profiles for SE, CA and UK probes demonstrating the number of assigned peptides with unique labeling sites on nine nucleophilic amino acid residues; top panel, SE, middle panel, CA, bottom panel, UK. (**c**), Solution reactivity profiles for SE, CA and UK probes; top panel, SE, middle panel, CA, bottom panel, UK. Representative extracted ion chromatograms are shown for product formation upon reacting the probes with 20 equivalents of each amino acid derivative in PBS for 12 hrs.
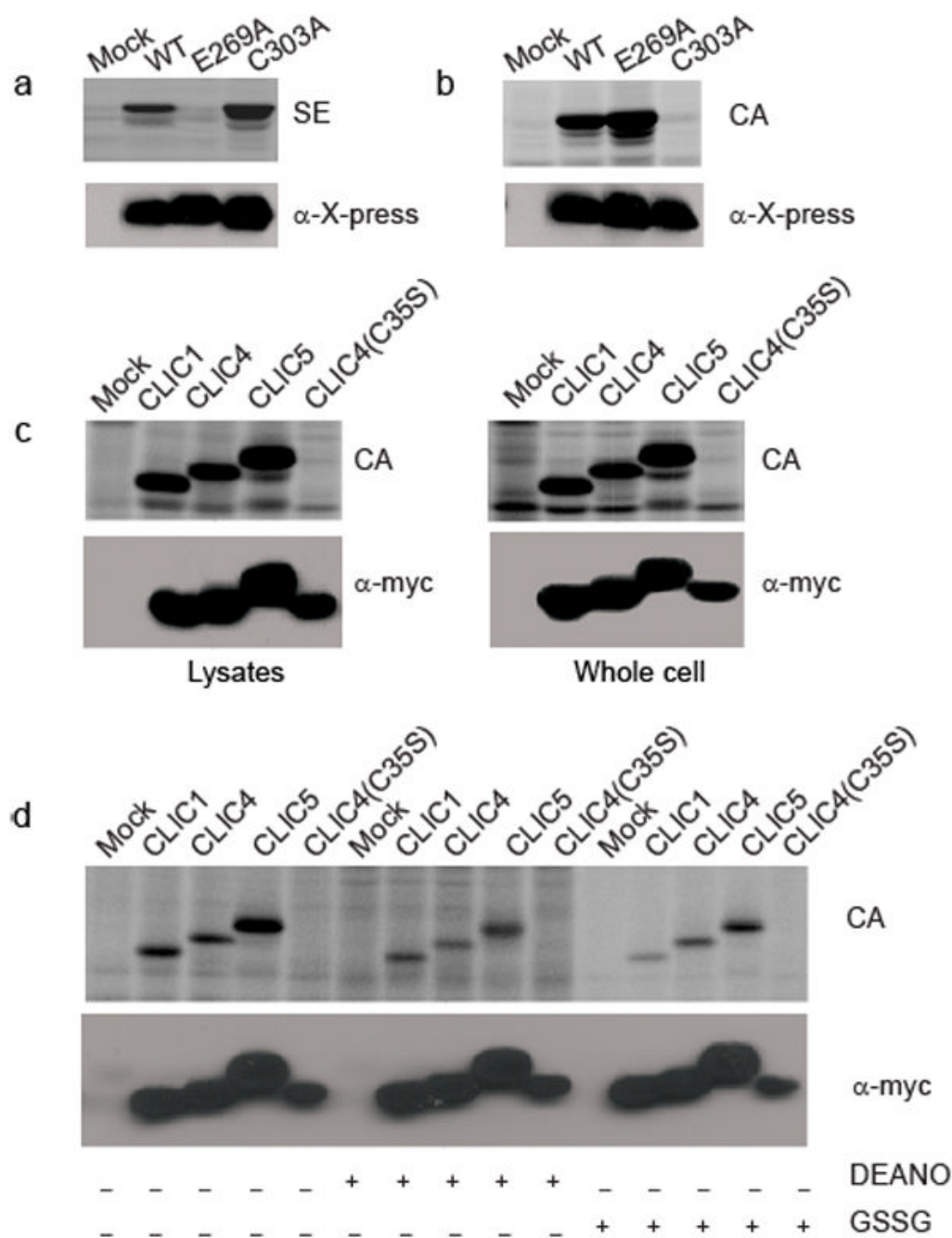
Figure 2.
Labeling of ALDH, and the CLIC protein family with carbon electrophile probes. (**a**),
Labeling of WT, E269A, and C303A ALDH1 enzymes with the CA probe. (**b**),. Labeling of
WT, E269A, and C303A ALDH1 enzymes with the SE probe. *Top panel*, fluorescent gel
images shown in grayscale demonstrating the selective labeling of the E269A and C303A
mutants with the CA and SE probes, respectively. *Bottom panel*, western blots confirming
equivalent expression of WT, E269A and C303A ALDH1 enzymes using α-X-press
antibodies. (**c**), Labeling of CLICs with the CA probe in lysates and in whole cells. *Top
panel*, fluorescent gel images shown in grayscale. *Bottom panel*, western blots confirming
expression of CLICs using α-myc antibodies. (**d**), Nitric oxide and oxidized glutathione

treatment of CLICs. Lysates were treated with either 5 mM of the nitric oxide donor, diethylamine nitric oxide, sodium salt (DEANO) or 2 mM of oxidized glutathione (GSSG). *Top panel*, fluorescent gel images shown in grayscale. *Bottom panel,* western blots confirming expression of CLICs using α-myc antibodies.