



Subscriber access provided by University of Newcastle, Australia

Diverse Redoxome Reactivity Profiles of Carbon Nucleophiles

Vinayak Gupta, Jing Yang, Daniel C. Liebler, and Kate S. Carroll

J. Am. Chem. Soc., Just Accepted Manuscript • DOI: 10.1021/jacs.7b01791 • Publication Date (Web): 29 Mar 2017

Downloaded from http://pubs.acs.org on March 31, 2017

Just Accepted

Article

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036 Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Diverse Redoxome Reactivity Profiles of Carbon Nucleophiles

Vinayak Gupta,^{‡, [1]} Jing Yang,^{‡, [2]} Daniel C. Liebler,^[3] and Kate S. Carroll^{*, [1]}

^[1] Department of Chemistry, The Scripps Research Institute, Jupiter, Florida 33458, USA.

^[2] State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences, Beijing 102206, China.

^[3] Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA.

ABSTRACT: Targeted covalent inhibitors have emerged as a powerful approach in the drug discovery pipeline. Key to this process is the identification of signaling pathways (or receptors) specific to (or over-expressed in) disease cells. In this context, fragment-based ligand discovery (FBLD) has significantly expanded our view of the ligandable proteome and affords tool compounds for biological inquiry. To date, such covalent ligand discovery has almost exclusively employed cysteine-reactive small-molecule fragments. However, functional cysteine residues in proteins are often redox-sensitive and can undergo oxidation in cells. Such reactions are particularly relevant in diseases, like cancer, which are linked to excessive production of reactive oxygen species (ROS). Once oxidized, the sulfur atom of cysteine is much less reactive toward electrophilic groups used in the traditional FBLD paradigm. To address this limitation, we recently developed a novel library of diverse carbon-based nucleophile fragments that react selectively with cysteine sulfenic acid (Cys-SOH) formed in proteins *via* oxidation or hydrolysis reactions. Here, we report analysis of sulfenic acid-reactive C-nucleophile fragments screened against a colon cancer cell proteome. Covalent ligands were identified for >1280 *S*-sulfenylated cysteines present in 'druggable' proteins and orphan targets, revealing disparate reactivity profiles and target preferences. Among the unique ligand-protein interactions identified was that of a pyrrolidinedione nucleophile, PYD that reacted preferentially with protein tyrosine phosphatases (PTPs). Fragment-based covalent ligand discovery with C-nucleophiles affords an expansive snapshot of the ligandable 'redoxome' with significant implications for covalent inhibitor pharmacology and also affords new chemical tools to investigate redox-regulation of protein function.

INTRODUCTION

In recent years, targeting of disease proteins and pathways with covalent inhibitors has emerged as a viable option to overcome drug resistance, mutations and toxicity associated with non-covalent inhibitors.^{1,2} Advantages of targeted covalent inhibition (TCI) strategies include extended duration of action and the ability to target rare residues and shallow binding sites. The mechanism by which TCIs achieve desired inhibition necessitate two distinct steps - the first involves the reversible association of a high-affinity ligand with its biological target in such a manner that an electrophilic 'warhead' on the ligand is brought into an appropriately proximate position to form a covalent adduct with nucleophilic residue on the protein (more often than not, a cysteine). Majority of TCIs are being developed to target kinases for their role in cancer.³ Indeed, FDA approval of drugs such as 'Afatinib' and 'Ibrutinib' validated this strategy of incorporating Michael acceptor (electrophilic in nature) functionality in small-molecule inhibitors to covalently target cysteine residue in the active sites of diseased proteins (Fig. 1a). While extremely attractive, designing TCIs is difficult in practice because of the difficulties in striking the right balance between reactivity and selectivity. In other words, electrophilic warheads may suffer from chemoselectivity issues due to the presence of other nucleophilic species (such as lysine, histidine etc.). The electrophiles could

also be scavenged by ubiquitous low-molecular-weight nucleophiles such as glutathione (GSH). Moreover, current TCIs predominantly target protein cysteines and among many nucleophiles present in the cellular system, cysteine thiols/thiolates (Cys-SH/S) are primary targets for the redoxbased modulation of protein activity. Oxidation of a protein cysteine thiol (Cys-SH) to sulfenic acid (Cys-SOH) by reactive oxygen species (ROS) (e.g., H₂O₂) is termed Ssulfenylation (Fig. 1b). It is a reversible post-translational modification that plays a crucial role in regulating several protein functions.⁴⁻⁸ Over the past decade several groups have reported the role of Cys-SOH in regulation of proteins such as transcription factors, kinases, phosphatases, ion channels, peroxidases and cysteine proteases, sirtuins, human serum albumin, and many others.^{7,9-18} Aberrant Cys-SOH formations have been shown to be a biomarker of disease state as well.^{19,20} The aforementioned examples and many other reports have established protein S-sulfenylation as a global signaling mechanism and a potential drug target.²¹⁻²³

Due to over expression, mutations and constitutive activation in cancer cells, EGFR is targeted by a major class of covalent drugs that are aimed to covalently add at ⁷⁹⁷Cys-SH.^{1,3} However, ⁷⁹⁷Cys present in the catalytic pocket of EGFR is a target of signal-derived H_2O_2 oxidation which enhances the kinase activity of this enzyme.^{12,14,24} We recently showed that



2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

43

44

49

55

56

57

58 59

60

Figure 1. Targeting the cysteine thiol (Cys-SH) with covalent inhibitors. **(a)** FDA approved Afatinib forms a covalent adduct with EGFR ⁷⁹⁷Cys-SH *via* Michael addition. **(b)** Oxidation of Cys-SH results generates cysteine sulfenic acid (Cys-SOH). **(c)** Covalent inhibitors with electrophilic warheads (such as the Michael acceptor highlighted in grey) undergo facile reaction with Cys-SH, but not with Cys-SOH under physiological conditions.

chronic oxidation result in an EGFR subpopulation that is refractory to covalent drugs such as Afatinib (Fig. 1c).¹⁴ Inhibitors targeting ⁷⁹⁷Cys-SOH would be better suited for the purpose of preferential reactivity as well as enhanced selectivity. Another example of enzyme class that is redox regulated is protein tyrosine phosphatases (PTPs), including PTP1B for Type II diabetes and obesity, and SHP2 for cancer and others for rheumatoid arthritis, making them attractive therapeutic targets.²⁵⁻²⁹ In the past, we have reported redox-based trifunctional probes consisting of a warhead to covalently target Cys-SOH of PTPs, a ligand that directs the binding to the target and a reporter tag used for the identification, purification or direct visualization of the labeled PTP. These redox-based probes showed low micromolar inhibitory activity against phosphatase YopH.³⁰

Fragment-based ligand discovery (FBLD) and activitybased protein profiling (ABPP) approaches are powerful tools for the investigation of protein function and identification of drug leads for new therapeutics.^{31,32} Over the years, several FBLD and ABPP approaches have been reported that profile the reactivity of carbon electrophiles towards biological nucleophiles (such as Cys-SH) in cellular proteomes.³²⁻³⁴ However, ligands identified in these studies are incapable of reacting with cysteine sulfenic acid (Cys-SOH) under physiological conditions and cysteine oxidation is expected to impact the pharmacology of inhibitors based on covalent targeting of Cys-SH.¹⁴ For this reason, we have proposed a complementary strategy that uses clickable carbon nucleophiles in a chemical proteomic assay and studied the Cys sulfenylome in RKO colon adenocarcinoma cells. Based on the results of this work, the sulfenyl form of therapeutically important proteins could be exploited to develop a new class of irreversible covalent inhibitors targeting this unique Cys oxoform.

RESULTS AND DISCUSSION

Current approaches to study protein sulfenylation are based on a β -dicarbonyl scaffold that shows moderate reactivity compared to other biologically relevant reactions (such as disulfide formation).^{8,12,35-38} Moreover, due to the lack of scaffold diversity among present Cys-SOH probes, current proteomicsbased approaches (such as FBLD and ABPP) to identify small molecule ligands targeting sulfenylated proteins could not be implemented effectively.^{13,36,39-41} To remedy the moderate reactivity of current nucleophile probes and to increase the structure diversity, we recently developed a library of cyclic and linear C-nucleophiles that showed diverse reactivity profiles towards Cys-SOH in a small molecule dipeptide model as well as in a protein model.⁴²⁻⁴⁴ The next logical step is to identify the biological target preferences of these newly developed nucleophiles. To do so, we first broadly divided the collection of ~100 cyclic C-nucleophiles into eight distinct scaffolds cyclohexane-1,3-diones (1), (thio)barbituric acids (2), different sized β -dicarbonyl rings (3), dihydrothiophen-3(2H)-one 1,1-



Figure 2. Novel classes of nucleophile probes to profile cysteine oxidation (a) Structural classes of cyclic C-nucleophiles (1 - 8). In each structure, the nucleophilic carbon is highlighted in the red circle. Rate constants from ref. 42. (b) From the pool of cyclic C-nucleophiles, DYn-2 and four new C-nucleophiles were selected based on a range of reaction rate constants and on scaffold diversity. (c) Energy minimized 3-dimensional representation of the probes.

2

3

8

9 10

11

16

17 18

19

20

21

22 23 24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58 59

60



Figure 3. Application of new nucleophile probes for studying cellular sulfenylomes. (a) RKO whole cell lysates were prepared under non-denaturing conditions. Native lysate was divided into equal portions and incubated with one of five nucleophile probes (or vehicle) in separate reactions. Labeled proteins from each reaction were digested separately and the resulting peptides subjected to click reaction with UV-cleavable biotin-N₃. Enrichment and photo-release afforded different probe-labeled peptides samples for proteomic analysis.

dioxides/2-substituted isothiazolidin-4-one 1,1-dioxides (4) pyrrolidine-2,4-diones (5), 1,3-indandiones (6), piperidine-2,4-5(6H)-one (7), 1,2-thiazinan-5-one 1,1-dioxides/2H-1,2thiazin-5(6H)-one 1,1-dioxides/1H-benzo[c][1,2]thiazine-4(3H)-one 2,2-dioxides (8) (Fig. 2a). Nucleophiles belonging to these scaffolds showed reaction rates that ranged from 2- to 150-fold higher than 1 (Fig. 2a).⁴²⁻⁴⁴ Next, we chose cyclic Cnucleophiles from the above-described eight scaffolds based on differences in reactivity rates and structural diversity (Fig. 2b, Fig. S1). An alkyne was incorporated into the framework of these nucleophiles to provide a click chemistry handle for gel analysis, enrichment and mass spectrometric analysis. DYn-2 is an established probe for Cys-SOH detection.¹² Owing to the stability of resultant enolic carbanion (due to resonance stabilization), its reaction rate with Cys-SOH is modest (10 M⁻¹s⁻¹).⁴² Minimum energy calculations show that this resonance stabilization also enables DYn-2 to toggle between half-chair conformers (Fig. 2c). Thiazolidin-4-one 1,1-dioxide based TD ($k_{abs} = 20 \text{ M}^{-1}\text{s}^{-1}$) exists predominantly in a keto form with an envelope conformation. This is attributed to the sulfonamide S being out of the plane (due to sterics), thus distorting the tetrahedral geometry around the S atom (Fig. 2c). In the case of pyrrolidine-2,4-dione based PYD ($k_{obs} = 330 \text{ M}^{-1}$ ¹s⁻¹), the keto-enol equilibrium favors the keto form.⁴² Minimum energy calculations for PYD showed an almost complete preference for planar geometry (Fig. 2c). Piperidine-2,4-dione based PRD ($k_{obs} = 1155 \text{ M}^{-1}\text{s}^{-1}$) is structurally similar to DYn-2, but exhibits rate enhancement of more than two orders of magnitude, which is attributed to the instability of the carbanion in a keto-enol equilibrium that favors the keto form.⁴² Minimum energy calculation indicated that PRD preferentially exists in the boat conformation (Fig. 2c). Benzo[c][1,2]thiazine-based BTD ($k_{obs} = 1700 \text{ M}^{-1}\text{s}^{-1}$) showed the highest reaction rate towards Cys-SOH, which is attributed to predominance of the keto conformation and a distorted tetrahedron geometry around the S atom.⁴² Minimum energy calcu-

lation showed that BTD adopts a half-boat conformation with sulfone S out of the plane, similar to TD (Fig. 2c).

Application of these nucleophilies to RKO colon adenocarcinoma cells under basal conditions, followed by click chemistry with a biotin azide reporter tag and visualization of labeled proteins by SDS-PAGE demonstrated that the panel of nucleophiles exhibited a range of reactivities consistent with their rate constants (**Fig. S2**). As expected, the highest reactivity was observed for the BTD probe, which demonstrated substantial protein labeling at concentrations as low as 100 μ M. The PRD and PYD probes showed moderate to high levels of labeling while DYn-2 and TD displayed less extensive labeling at low concentrations (**Fig. S2**). From these results, it was apparent that if the protein site accessibility is not limiting, the comparative reactivity of these probes correlates well with their reaction rate constants towards Cys-SOH.

Subsequently, we further examined the peptide labeling profiles for these probes by analyzing their proteome reactivities using a mass spectrometry platform and a workflow referred to as SulfenM to site-specifically map the S-sulfenylome in RKO cells. The selectivity, accuracy and reproducibility of the workflow have been reported previously.^{13,39} The workflow consists of eight steps - lysing the RKO cells under native conditions (using lysis buffer containing non-ionic detergent and catalase) in the presence of C-nucelophile (Fig. 3a), reducing the native disulfides with DTT, alkylating the reduced cysteines, digesting the labeled proteins, conjugation of labeled peptides with a photocleavable azide biotin via 'click chemistry', capture of labeled peptides with streptavidin beads, photorelease by cleavage of the biotin linker with UV light and MS based proteomic analysis of the probe-labeled peptides (Fig. S3). A total of 1283 S-sulfenylated sites on 761 proteins were successfully labeled by the five C-nucleophiles (Fig. 3b-c). Interestingly, only nine proteins (ACLY, EEF2, GAPDH, HSPA8, PFN1, RPS20, RPS21, RPS27A, RPS3A)



Figure 4. Chemoproteomic analysis of RKO cellular S-sulfenylome. (a) Crystal structures of GAPDH (PDB: 4WNC) showing Cys152, Cys156 and Cys247 labeled by different probes. (b) CARM1 (PDB: 5Y4X) showing Cys421 that is labeled by TD. (c) Comparative sequence analysis of peptides labeled with different probes. (d) Heat map of Z-score normalized peak intensities of S-sulfenylated sites labeled with various probe. Values in each row direction have been mean centered and scaled. (Probe modified Cys occurring > 1 indicates different charge states on the corresponding peptide).

were labeled by all five nucleophiles. For example, in GAPDH, ¹⁵²Cys-SOH (by DYn-2, PYD, PRD and BTD), ¹⁵⁶Cys-SOH (by DYn-2, BTD) and ²⁴⁷Cys-SOH (TD, PYD, PRD, BTD) were successfully labeled. GAPDH is known as a major redox-sensitive protein, which has three Cys residues, one (¹⁵²Cys) in the exposed active site, and two (¹⁵⁶Cys, ²⁴⁷Cys) that are buried (**Fig. 4a**). While redox-sensitivity of ¹⁵²Cys is well established, ¹⁵⁶Cys and ²⁴⁷Cys were recently reported to be sensitive to storage-dependent oxidation. ⁴⁵ Even more interestingly, only two S-sulfenylated sites (¹⁷Cys-SOH on RPS21 and ¹²⁸Cys-SOH on PFN1) were labeled by all five

 nucleophiles (Fig. S4). An interesting target was PARK7 (also known as DJ-1), which is a highly conserved homodimeric small protein associated with Parkinson's disease.⁴⁶ Studies have demonstrated that DJ-1 protects cells against oxidative stress-mediated apoptosis, the mechanism of which is largely unknown.⁷ DJ-1 is also known to regulate the activity of phosphatase and tensin homolog (PTEN) via transnitrosylation reaction.⁴⁷ Structurally, DJ-1 has three potentially redox-active cysteines (⁴⁶Cys, ⁵³Cys and ¹⁰⁶Cys). However, ¹⁰⁶Cys, which is highly conserved, was shown to be highly sensitive to

2
3
1
4
5
6
7
2
8
9
10
11
11
12
13
14
45
15
16
17
10
10
19
20
21
21
ZZ
23
24
25
20
26
27
28
20
29
30
31
32
32
33
34
35
200
30
37
38
20
39
40
41
42
12
43
44
45
16
40
47
48
49
50
50
51
52
52
55
54
55
56
57
57
58

59

60

Probes	# of distinct labeled proteins	# of distinct labeled sites
DYn-2	150	183
TD	56	58
PYD	76	87
PRD	60	68
BTD	715	1186

oxidation and formed a stable Cys-SO₂H.⁷ We showed that DJ-1 was labeled by DYn-2 and PRD at ¹⁰⁶Cys and BTD at ⁴⁶Cys respectively. The rank order of the number of distinct probe-modified sites identified by these chemoproteomic analyses was BTD >> DYn-2 > PYD > PRD > TD (**Table 1**). Another interesting target was CARM1, in which ⁴²¹Cys was labeled by TD. CARM1 (also known as PRMT4) is a member of arginine-specific protein methyltransferases (PRMTs). While most PRMTs methylate Gly- and Arg-rich (GAR) motifs in their substrates, CARM1 displays unique substrate specificity by methylating Pro-, Gly-, and Met-rich motifs (PGM motifs).⁴⁸

Owing to its superior reaction rate towards Cys-SOH, BTD exhibited the highest level of S-sulfenylome reactivity. However, a very restricted reactivity profile was observed for other probes. Of particular interest was the relatively low number of sulfenylated proteins labeled by the PYD and PRD probes in RKO sulfenylome (Table 1), which contrasted sharply with the robust reactivity shown by these probes in native RKO cell lysate (Fig. S2). Potential reversibility of thioether adducts under reducing conditions under the proteomics work-up conditions could be responsible for unexpected labeling profile of PYD and PRD.44 To further explore structural features that may define the protein sulfenylation, we examined the flanking sequences of S-sulfenylated Cys residues modified by the five probes with the pLogo algorithm for the presence of consensus motifs (Fig. S5).⁴⁹ We found several consensus motifs for the nucleophile-labeled S-sulfenylome, such as KxC and KC. Of interest, Lys was significantly overrepresented at almost every adjacent position (from -6 to +6) of BTD-labeled Cys-SOH. This observed enrichment of Lys is intriguing and may indicate that positive charge in the vicinity of the reactive cysteine is critical for thiolate generation. However, if this were strictly the case, then Arg enrichment would also be observed. Another possibility is that due to its elevated reactivity, BTD was able to capture highly transient sulfenylated proteins. But the abundance of such BTD labeled peptides is low and probe-tagged peptides containing Lys ionize better, thus enabling detection of these peptides above the noise. We also compared the flanking sequences of S-sulfenvlated Cvs residues modified by the five probes using Two Sample Logo (2logP) algorithm (Fig. 4c).⁵⁰ Structurally, DYn-2 and PRD differ by just one atom (C4 in DYn-2 vs N in PRD). 2LogP analyses indicated that, PRD labeled peptides show higher frequency of Lys (at -2) and Leu (at +4) residues compared to DYn-2 modified peptides (Fig. 4c, panel 1). On the other hand, both PYD and PRD are β -ketolactams, but differ in the ring size, leading to differences in conformation but which also resulted in higher frequency of Lys (at -2) and Leu (+4) in the peptides modified by PRD (Fig. 4c, panel 2). KxC and CxxxL thus appeared to be the conserved sequence motifs in

PRD modified peptides. Indeed, a pLogo analysis of PRD modified peptides show a conserved KxC sequence motif as well (Fig. S5). Similarly, when sequences of peptides labeled by TD probe were compared with those modified by PYD (owing to both having 5-membered ring structure) and BTD (due to structural analogy), the KxC sequence motif was again found to be conserved (Fig. 4c, panel 3 - 4), an observation which also was seen in pLogo sequence analysis of TD (Fig. S5).

To further estimate the stoichiometries and selectivity of nucleophile labeling, we obtained peak intensities of all the detected S-sulfenylated peptides using MS1 filtering based label-free quantification as previously described (Fig. S6).^{13,39} The label-free quantification results from each nucleophilelabeling experiment were then standardized with Z-score normalization to enable further comparison. The quantitative comparison of peptide-nucleophile labeling is illustrated in a heat map (Fig. 4d). These data suggested that the nucleophiles display different trends of S-sulfenylome reactivity and selectivity in complex proteomes. For example, ⁴⁴²Cys-SOH of HSPD1 (Hsp60) showed a clear preference for BTD over DYn-2 (Fig. 4d). In some cases, S-sulfenylated sites on the same protein showed different probe preferences. DYn-2, PYD, PRD and BTD labeled ¹⁵²Cys-SOH of GAPDH with almost equal efficiencies, but BTD was more efficient in labeling surface exposed ²⁴⁷Cys-SOH compared to TD, PYD and PRD (Fig. 4d). These results could be biased by differences in ionization efficiency among different peptide-probe adducts. To evaluate this possibility, we compared the ionization of Cbz-Cys-Val-OMe peptide labeled with different nucleophiles. Each probe was exposed to limiting amount of Cbz-Cys(SOH)-Val-OMe under standard buffer conditions (Fig. S7a). Thioether products corresponding to each nucleophile were HPLC-purified, mixed together in equimolar quantities and analyzed by LC/MS. The relative ionization intensities of these adducts differed by one-fold between the lowest and highest ionizing species with lowest being Cbz-Cys(S-BTD)-Val-OMe and highest being Cbz-Cvs(S-DYn-2)-Val-OMe indicating limited effect of ionization efficiencies of peptideprobe adducts (Fig. S7b).

We next investigated whether the S-sulfenylomes labeled by the five probes had distinct structural features. Based on the calculations resulting from twenty proteins with structure files available in Protein Data Bank (PDB) that contained both buried and exposed Cys residues, we observed that DYn-2 and

Table 2. Surface distribution analysis of probe-labeled proteins. Calculations were based on twenty proteins with 3D structure files found in the Protein Data Bank (PDB) that contained both buried and exposed cysteine residues. E: exposed (RSA>25%); B: buried (RSA <10%); UM: Unmodified cysteine residues.

	E (%)	B (%)
DYn-2	35.3	52.9
TD	0.0	80.0
PYD	21.4	71.4
PRD	33.3	55.6
BTD	25.0	62.5
UM	19.6	58.7



Figure 5. PTPs as privileged targets of PYD. (a) Representative MS/MS spectra of PTPN1 (PTP1B) labeled by PYD probe. (b) PYD probe preferentially modifies the recombinant oxidized form of PTP1B. (c) Molecular docking (using AutoDock Vina) of PYD into catalytic pocket of PTPN1 (PDB: 10ET). (d) Abbreviated sequence alignment of PTPs that harbor a Cys-SOH labeled by PYD probe.

PRD prefer to label S-sulfenylated residues on protein surfaces. In contrast, PYD and TD display substantial reactivity towards the buried S-sulfenylated sites (Table 2). Of particular interest was the distinct reactivity of the PYD probe with the S-sulfenylated sites within the buried catalytic domains of several classical protein tyrosine phosphatase (PTPs) such as PTPN1 (PTP1B), PTPN7, PTPN11 and DUSP23 (Fig. 5a, Fig. S8). The preferential reactivity of PYD towards PTP1B was also verified by treatment of recombinant, oxidized phosphatase with each nucleophile followed by click chemistrymediated conjugation to an biotin-azide reporter tag and visualization by Strepavidin-HRP western blot (Fig. 5b). Molecular docking (using AutoDock Vina)⁵¹ of PYD into the catalytic pocket of PTP1B showed a good fit with the catalytic pocket, ostensibly owing to the almost planar conformation of this Cnucleophile (Fig. 5c). The CX₅R sequence is the signature motif of the PTP superfamily and forms the critical phosphatebinding loop (also known as the P-loop) in the active site (Fig. **5d)**. Irrespective of the variations in the X_5 segment, conformation of the P-loop is conserved (to keep catalytic Cys and Arg in close proximity) and can be easily superimposed among different PTP structures. PTPs are established targets of ROS because the signature motif CX₅R of this family contains an invariant cysteine that has an extremely low pk_{a} , which promotes its function as a nucleophile in catalysis and renders it highly susceptible to oxidative inactivation.²

CONCLUSION

1

2

3

4

5

6

7

8

9

10

11 12

13

14

15

16

17 18

19

20

21

22

23 24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59 60 In order to develop inhibitors that effectively target redoxactive cysteines in therapeutically important proteins, it is imperative to complement the structural knowledge of Ssulfenylated sites with an understanding of warhead reactivity. With suitable chemical tools, modern chemical proteomics approaches such as FBLD and ABPP present a way to achieve this task in a systematic and efficient manner. Not only do such approaches provide the technology that help identify sites of covalent modification, but they also differentiate between selective versus promiscuous covalent modifiers and thus aid

future drug design and development. Herein, we utilized a novel multi-probe approach and a highly efficient chemoproteomics platform to globally profile the reactivity of five Cnucleophile probes, including four novel molecules, in the RKO S-sulfenylome under normal physiological conditions. These five C-nucleophiles covalently labeled 1283 redoxsensitive Cys residues on 761 proteins that are endogenously S-sulfenylated. We show that distinct structural classes of Cnucleophiles targeted discrete protein populations and Ssulfenylated residues with little overlap. Whereas BTD labeled the greatest numbers of S-sulfenylated sites by far, the smaller size of the five-membered ring of TD and PYD appear to facilitate internal site labeling. In addition to physical size, the conformations of five-membered (envelope for TD and almost planar for PYD) versus six-membered (chair and boat) probes also differ, which also would favor buried site labeling by TD and PYD. Insights from this proteomics study have important implications in the development of covalent inhibitors that target cysteines within the thiol 'redoxome'. For example, PTPs, including PTP1B for Type II diabetes and obesity, and SHP2 for cancer and others for rheumatoid arthritis, are attractive therapeutic targets.²⁵⁻²⁸ Our study has identified the catalytic Cys in PTP1B (²¹⁵Cys), PTPN7 (³⁹⁶Cys), PTPN11 ⁴⁵⁹Cys) and DUSP23 (⁹⁵Cys), PTPs as favored targets of pyrrolidine-2,4-dione-based nucleophile, PYD. The enrichment of PYD-labeled PTPs suggests that PYD could be coupled to binding elements that capture interactions with adjacent pockets surrounding the active site to develop covalent inhibitors for this class of proteins. This strategy can be expanded and applied to other attractive targets. Our expanding library of novel warheads that target S-sulfenylated cysteines will enable high-throughput profiling to pair distinct C-nucleophilic scaffolds with various redox-regulated proteins in a combined FBLD/ABPP approach to develop effective covalent irreversible/reversible inhibitors based therapeutics.^{23,32,52} Finally, these probes can be used to study how oxidative cysteine modifications impact both cell signaling and drug pharmacology.

2

3

4

5

6

7

8

9

10

11

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

46

47

48

49

50

51

52

58

59

60

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental details and protocols including NMR data (PDF).

Proteomics data associated with each probe (XLSX).

AUTHOR INFORMATION

Corresponding Author

* Correspondence and requests for materials should be addressed to K.S.C. (kcarroll@scripps.edu)

12 Author Contributions

‡ These authors contributed equally.

Funding Sources

Part of this work is supported by the National Institutes of Health under awards No. R01 GM102187 and R01 CA174864 to K. S. C. J.Y. acknowledge the support of the National Key Research and Development Program (No. 2016YFA0501303), the National Natural Science Foundation of China (No. 31500666 and No. 81573395), the Beijing Natural Science Foundation (No. 5162009), and Beijing Nova Program.

ACKNOWLEDGMENT

V.G. thanks Dr. Mauro Lo Conte (Department of Chemistry, Scripps Florida) for biotin-N₃ reagent and Dr. Peng Wu (Department of Chemistry, Scripps La Jolla) & Dr. Hua Wang (Department of Chemistry, Scripps La Jolla) for providing BTTP 'click' ligand used in the 'click' chemistry for WB analysis.

REFERENCES

(1) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. Nat Rev Drug Discov 2011, 10, 307.

- (2) Baillie, T. A. Angew Chem Int Ed Engl 2016, 55, 13408.
- (3) Liu, Q.; Sabnis, Y.; Zhao, Z.; Zhang, T.; Buhrlage, S. J.;
- Jones, L. H.; Gray, N. S. Chem Biol 2013, 20, 146.
- (4) Paulsen, C. E.; Carroll, K. S. Chem Rev 2013, 113, 4633.
- (5) Reddie, K. G.; Carroll, K. S. Curr Opin Chem Biol 2008, 12, 746.
- (6) Paulsen, C. E.; Carroll, K. S. ACS Chem Biol 2010, 5, 47.
- (7) Lo Conte, M.; Carroll, K. S. J Biol Chem 2013, 288, 26480.
- (8) Gupta, V.; Carroll, K. S. *Biochim Biophys Acta* **2014**, *1840*, 847.
- (9) Denu, J. M.; Tanner, K. G. Biochemistry 1998, 37, 5633.
- 42 (10) Svoboda, L. K.; Reddie, K. G.; Zhang, L.; Vesely, E. D.;
 43 Williams, E. S.; Schumacher, S. M.; O'Connell, R. P.; Shaw, P.;
- Williams, E. S.; Schumacher, S. M.; O'Connell, R. P.; Shaw, R.;
 Day, S. M.; Anumonwo, J. M.; Carroll, K. S.; Martens, J. R. *Circ Res* 2012, *111*, 842.
 - (11) Kulathu, Y.; Garcia, F. J.; Mevissen, T. E.; Busch, M.; Arnaudo, N.; Carroll, K. S.; Barford, D.; Komander, D. *Nat Commun* **2013**, *4*, 1569.
 - (12) Paulsen, C. E.; Truong, T. H.; Garcia, F. J.; Homann, A.; Gupta, V.; Leonard, S. E.; Carroll, K. S. *Nat Chem Biol* **2011**, *8*, 57.
 - (13) Yang, J.; Gupta, V.; Carroll, K. S.; Liebler, D. C. Nat Commun 2014, 5, 4776.
- 53 (14) Truong, T. H.; Ung, P. M.; Palde, P. B.; Paulsen, C. E.;
 54 Schlessinger, A.; Carroll, K. S. *Cell Chem Biol* 2016, *23*, 837.
- 55 (15) Barrett, T. J.; Pattison, D. I.; Leonard, S. E.; Carroll, K. S.;
- Davies, M. J.; Hawkins, C. L. Free Radic Biol Med 2012, 52, 1075.

(16) Alili, L.; Sack, M.; von Montfort, C.; Giri, S.; Das, S.; Carroll, K. S.; Zanger, K.; Seal, S.; Brenneisen, P. *Antioxid Redox Signal* **2013**, *19*, 765.

(17) Chouchani, E. T.; Kazak, L.; Jedrychowski, M. P.; Lu, G. Z.; Erickson, B. K.; Szpyt, J.; Pierce, K. A.; Laznik-Bogoslavski, D.; Vetrivelan, R.; Clish, C. B.; Robinson, A. J.; Gygi, S. P.; Spiegelman, B. M. *Nature* **2016**, *532*, 112.

(18) Hourihan, J. M.; Moronetti Mazzeo, L. E.; Fernandez-Cardenas, L. P.; Blackwell, T. K. *Mol Cell* **2016**, *63*, 553.

(19) Zito, E.; Hansen, H. G.; Yeo, G. S.; Fujii, J.; Ron, D. Mol Cell 2012, 48, 39.

- (20) Seo, Y. H.; Carroll, K. S. P Natl Acad Sci USA 2009, 106, 16163.
- (21) Truong, T. H.; Carroll, K. S. Biochemistry 2012, 51, 9954.
- (22) Truong, T. H.; Carroll, K. S. Crit Rev Biochem Mol Biol 2013, 48, 332.

(23) Visscher, M.; Arkin, M. R.; Dansen, T. B. Curr Opin Chem Biol 2016, 30, 61.

- (24) Heppner, D. E.; Hristova, M.; Dustin, C. M.; Danyal, K.; Habibovic, A.; van der Vliet, A. *J Biol Chem* **2016**, *291*, 23282.
- (25) Elchebly, M.; Payette, P.; Michaliszyn, E.; Cromlish, W.;
- Collins, S.; Loy, A. L.; Normandin, D.; Cheng, A.; Himms-
- Hagen, J.; Chan, C. C.; Ramachandran, C.; Gresser, M. J.; Tremblay, M. L.; Kennedy, B. P. *Science* **1999**, *283*, 1544.
- (26) Saltiel, A. R.; Kahn, C. R. Nature 2001, 414, 799.
- (27) He, R. J.; Yu, Z. H.; Zhang, R. Y.; Zhang, Z. Y. Acta Pharmacol Sin **2014**, *35*, 1227.
- (28) Barr, A. J. Future Med Chem 2010, 2, 1563.
- (29) Tonks, N. K. Nat Rev Mol Cell Biol 2006, 7, 833.

(30) Leonard, S. E.; Garcia, F. J.; Goodsell, D. S.; Carroll, K. S. Angew Chem Int Ed Engl **2011**, *50*, 4423.

- (31) Murray, C. W.; Rees, D. C. Nat Chem 2009, 1, 187.
- (32) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.;
- Horning, B. D.; Gonzalez-Paez, G. E.; Chatterjee, S.; Lanning, B.

R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F. *Nature* **2016**, *534*, 570.

- (33) Weerapana, E.; Simon, G. M.; Cravatt, B. F. Nat Chem Biol 2008, 4, 405.
- (34) Kathman, S. G.; Xu, Z.; Statsyuk, A. V. J Med Chem 2014, 57, 4969.
- (35) Reddie, K. G.; Seo, Y. H.; Muse, W. B.; Leonard, S. E.; Carroll, K. S. *Mol Biosyst* **2008**, *4*, 521.
- (36) Leonard, S. E.; Reddie, K. G.; Carroll, K. S. ACS Chem Biol 2009, 4, 783.
- (37) Qian, J.; Klomsiri, C.; Wright, M. W.; King, S. B.; Tsang, A. W.; Poole, L. B.; Furdui, C. M. *Chem Commun* **2011**, *47*, 9203.
- (38) Qian, J.; Wani, R.; Klomsiri, C.; Poole, L. B.; Tsang, A. W.; Furdui, C. M. *Chem Commun* **2012**, *48*, 4091.
- (39) Yang, J.; Gupta, V.; Tallman, K. A.; Porter, N. A.; Carroll, K. S.; Liebler, D. C. *Nat Protoc* **2015**, *10*, 1022.
- (40) Gould, N. S.; Evans, P.; Martinez-Acedo, P.; Marino, S. M.; Gladyshev, V. N.; Carroll, K. S.; Ischiropoulos, H. *Chem Biol* **2015**, *22*, 965.
- (41) Akter, S.; Huang, J.; Bodra, N.; De Smet, B.; Wahni, K.; Rombaut, D.; Pauwels, J.; Gevaert, K.; Carroll, K.; Van
- Breusegem, F.; Messens, J. Mol Cell Proteomics 2015, 14, 1183.
- (42) Gupta, V.; Carroll, K. S. Chem Sci 2016, 7, 400.
- (43) Gupta, V.; Carroll, K. S. Chem Commun (Camb) 2016, 52, 3414.
- (44) Gupta, V.; Paritala, H.; Carroll, K. S. *Bioconjug Chem* **2016**, 27, 1411.
- (45) Reisz, J. A.; Wither, M. J.; Dzieciatkowska, M.; Nemkov, T.; Issaian, A.; Yoshida, T.; Dunham, A. J.; Hill, R. C.; Hansen, K. C.; D'Alessandro, A. *Blood* **2016**, *128*, e32.
- (46) Bonifati, V.; Rizzu, P.; van Baren, M. J.; Schaap, O.; Breedveld, G. J.; Krieger, E.; Dekker, M. C.; Squitieri, F.; Ibanez, P.; Joosse, M.; van Dongen, J. W.; Vanacore, N.; van Swieten, J.

- C.; Brice, A.; Meco, G.; van Duijn, C. M.; Oostra, B. A.; Heutink, P. Science 2003, 299, 256.
- (47) Choi, M. S.; Nakamura, T.; Cho, S. J.; Han, X. M.; Holland,
- E. A.; Qu, J.; Petsko, G. A.; Yates, J. R.; Liddington, R. C.;
- Lipton, S. A. J Neurosci 2014, 34, 15123.

- (48) Yang, Y.; Bedford, M. T. Nat Rev Cancer 2013, 13, 37.
- (49) O'Shea, J. P.; Chou, M. F.; Quader, S. A.; Ryan, J. K.;
- Church, G. M.; Schwartz, D. Nat Methods 2013, 10, 1211.
- (50) Vacic, V.; Iakoucheva, L. M.; Radivojac, P. Bioinformatics 2006, 22, 1536.
- (51) Trott, O.; Olson, A. J. J Comput Chem 2010, 31, 455.
- (52) Gruner, B. M.; Schulze, C. J.; Yang, D.; Ogasawara, D.; Dix,
- M. M.; Rogers, Z. N.; Chuang, C. H.; McFarland, C. D.; Chiou, S.
 - H.; Brown, J. M.; Cravatt, B. F.; Bogyo, M.; Winslow, M. M. Nat Methods 2016, 13, 883.

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

