Redox Biology 69 (2024) 102969

Contents lists available at ScienceDirect

Redox Biology

journal homepage: www.elsevier.com/locate/redox

Identification of sulfhydryl-containing proteins and further evaluation of the selenium-tagged redox homeostasis-regulating proteins

Zhongyao Jiang [a], Yue Tang [c,**], Jun Lu [a], Chang Xu [a], Yaxin Niu [a], Guanglu Zhang [a], Yanmei Yang [a], Xiufen Cheng [a], Lili Tong [a], Zhenzhen Chen [a,***], Bo Tang [a,b,*]

[a] College of Chemistry, Chemical Engineering and Materials Science, Key Laboratory of Molecular and Nano Probes, Minis-try of Education, Institute of Molecular and Nano Science, Collaborative Innovation Center of Functionalized Probes for Chemical Imaging in Universities of Shandong, Shandong Normal University, Jinan, 250014, PR China

[b] Laoshan Laboratory, Qingdao, 266200, PR China

[c] Department of Emergency Medicine, Shandong Provincial Clinical Research Center for Emergency and Critical Care Medicine, Qilu Hospital of Shandong University, Jinan, PR China

ARTICLE INFO

Keywords:
Chemoproteomic profiling
Sulfhydryl-containing proteins
selenium isotopic signature
Cysteine residues
Selenium-containing probes

ABSTRACT

Chemoproteomic profiling of sulfhydryl-containing proteins has consistently been an attractive research hotspot. However, there remains a dearth of probes that are specifically designed for sulfhydryl-containing proteins, possessing sufficient reactivity, specificity, distinctive isotopic signature, as well as efficient labeling and evaluation capabilities for proteins implicated in the regulation of redox homeostasis. Here, the specific selenium-containing probes (Se-probes) in this work displayed high specificity and reactivity toward cysteine thiols on small molecules, peptides and purified proteins and showed very good competitive effect of proteins labeling in gel-ABPP. We identified more than 6000 candidate proteins. In TOP-ABPP, we investigated the peptide labeled by Se-probes, which revealed a distinct isotopic envelope pattern of selenium in both the primary and secondary mass spectra. This unique pattern can provide compelling evidence for identifying redox regulatory proteins and other target peptides. Furthermore, our examiation of post-translational modification (PTMs) of the cysteine site residues showed that oxidation PTMs was predominantly observed. We anticipate that Se-probes will enable broader and deeper proteome-wide profiling of sulfhydryl-containing proteins, provide an ideal tool for focusing on proteins that regulate redox homeostasis and advance the development of innovative selenium-based pharmaceuticals.

1. Introduction

Sulfhydryl-containing proteins are crucial to biological and physiological processes owing to their thiol intrinsic nucleophilicity and redox sensitivity. Despite its relatively rare content in the proteome (~2 %), cysteines present in 97 % of all human proteins and multiple classes of proteins including proteases, oxidoreductases, kinases, and acyltransferases, deubiquitinases contain cysteines [1–3]. Moreover, the cysteine residues can undergo a broad range of PTMs including *S*-nitrosylation, *S*-sulfenylation, *S*-sulfinylation, *S*-sulfonylation, palmitoylation, and prenylation etc., which dramatically expands the functional diversity of proteins [4,5]. Thus, the chemical labeling and proteome-wide profiling of sulfhydryl-containing proteins has always been an attractive hotspot for identification of new druggable and therapeutic covalent ligands.

A broad variety of electrophilic probes for proteins-sulfhydryl profiling have been reported in the past decade [6–17]. Iodoacetamide alkyne (IA-alkyne) coupling activity-based protein profiling (ABPP) platform has been widely used for profiling of cysteines as targets of natural products, electrophilic drugs, libraries of covalent fragments and

* Corresponding author. College of Chemistry, Chemical Engineering and Materials Science, Key Laboratory of Molecular and Nano Probes, Minis-try of Education, Institute of Molecular and Nano Science, Collaborative Innovation Center of Functionalized Probes for Chemical Imaging in Universities of Shandong, Shandong Normal University, Jinan, 250014, PR China.

** Corresponding author.

*** Corresponding author.

E-mail addresses: yuetang0531@hotmail.com (Y. Tang), zzchen@sdnu.edu.cn (Z. Chen), tangb@sdnu.edu.cn (B. Tang).

https://doi.org/10.1016/j.redox.2023.102969
Received 18 October 2023; Received in revised form 16 November 2023; Accepted 16 November 2023
Available online 27 November 2023

electrophilic metabolites [18–22]. Nonetheless, each of these developed electrophilic probes also carries certain limitations, such as insufficient reaction, lower reaction rate or non-specificity. For example, IA-alkyne, *p*-chloronitrobenzene and NPM can also react with other nucleophilic amino acids such as lysine, tyrosine and histidine [8,11,23,24]. The reaction of IA-alkyne and other probes with protein thiols has been shown to be slow and often incomplete [25,26]. On the other side, the incorporation of some special atoms (such as Br or F) into proteomic probe structures will show increase the confidence in protein target and amino acid site assignment in mass-spectrometry-based identification workflows [10,27]. Therefore, a probe in which a particular atom itself is part of the structure of the reactive group, as well as possessing sufficient reactivity, high specificity, and high rapidity for comprehensive proteome-wide analysis, would be an ideal and versatile probe for proteins-sulfhydryl profiling.

In the last decade or so, our group has developed a variety of small molecule probes and analysis methods based on Se–N bond for imaging or mass spectrometry analysis of low molecular weight thiols or ebselen-targets in cells or other biological samples [28–33]. The thiol derivatization reaction through Se–N bond cleavage and Se–S bond formation is highly selective, rapid, and efficient at room temperature, which helps the transient state capture of sulfhydryl-containing proteins, effectively reduces the oxidation modification of thiol during probe incubation, avoids interference from other nucleophilic amino acids and enriched proteins releasing by reducing agents [34,35]. In addition, the formed Se–S bond in derivatized peptide exhibits tunable MS-dissociation behaviors, such as collision induced dissociation (CID), electron transfer dissociation (ETD) and ultraviolet photodissociation (UVPD) [34–37]. These different dissociation behaviors will not only be used to validate proteomic results but also contributes to the development of chemical proteomics. More to the point, the selenium tag possessing six stable isotopes with detectable distribution can be introduced into the target thiol-peptides [34,35]. The unique isotopic envelope pattern in mass spectrum constitutes an additional advantage with a more convincing adduct-cysteine site assignment in MS-based identification workflows [38]. Due to the aforementioned advantages, we posit that the Se–N bond would serve as a more optimal and appropriate warhead for identifying sulfhydryl-containing proteins.

It is noteworthy that numerous redox-regulated proteins heavily depend on the crucial cysteine or selenocysteine for their core functionalities, exemplified by thioredoxin (TXN) and glutathione peroxidase (GPX) [39,40]. Despite the availability of various chemical tools for labelling and monitoring redox proteins, there remains a dearth of a probe that can be more focused into effectively labelling and identifing redox-associated proteins and their cysteine sites for subsequent assessment in mass spectrometry. The redox chemistry emerges as the most prominent scientific distinction between selenium and sulfur. Selenium demonstrates a heightened level of redox activity in comparison to sulfur [41]. The organoselenium compounds (OSeCs) containing Se–N or Se–Se bond can also function as ‘‘hidden selenol’’ molecules [42,43]. Therefore, it is postulated that probes containing Se–N or Se–Se bond may exhibit enhanced specificity towards redox-regulated proteins.

Here, we utilized Se–N bond and Se–Se bond as new classes of cysteine-directed electrophiles, designed and synthesized Se-probes (NPSSyne, NPSPyne and DPDSyne) for proteome-wide profiling of cysteines. The high rapidity, specificity and reactivity of the Se-probes toward cysteine thiols were tested on small molecules, peptides and purified proteins. We identified more than 6000 candidate thiol-containing proteins as well as 1270 cysteine sites directly from tandem orthogonal proteolysis-ABPP (TOP-ABPP), and 4434 cysteine sites indirectly from rdTOP-ABPP in HeLa lysates. Subsequent bioinformatics analysis revealed that a considerable number of proteins involved in regulating redox homeostasis, including TXN, GPX, thioredoxin reductase (TXNRD) and peroxidase (PRDX), among the identified targets. Further analysis of the MS1 and MS2 spectra of the relevant site-peptides, the selenium isotopic signature could facilitate peptide sequencing and validate proteome-wide profiling of cysteine sites including various redox-regulated proteins. Moreover, we successfully determined the functional and classificatory attributes of the sulfhydryl-containing proteins, along with its cysteine domain, secondary structure, and PTMs. We envision that the Se-probes will be ideal chemical tools with selenium isotopic signature for detecting and capturing proteome-wide thiol-containing proteins and cysteine sites in many physiological and/or pathological conditions.

2. Materials and methods

2.1. Reagents and antibodies

Iodoacetamide (IAA), glutathione (GSH), DL-dithiothreitol (DTT), tris(2-chloroethyl) phosphate (TCEP), copper sulfate ($CuSO_4$), S-methyl methanethiosulfonate (MMTS) and calcium chloride ($CaCl_2$) were purchased from Sigma-Aldrich unless otherwise specified. Fmoc-Cys and N-ethylmaleimide (NEM) was obtained from Heowns and J&K Scientific, respectively. Biotin-PEG3-azide, DADPS-Biotin-Azide and click chemistry auxiliary reagents (TBTA and BTTAA) was purchased from Click chemistry tools. Cy3-Azide and desthiobiotin-PEG3-Azide was obtained from Okeanos and Confluore Biotechnology, respectively. Four peptides containing one cysteine, GCSWDYKN was synthesized from GL Biochem and the others were purchased from SciLight Biotechnology. All cell culture related reagents, such as fetal bovine serum, Dulbecco’s modified Eagle’s medium (DMEM) and penicillin-streptomycin were purchased from Biological Industries (BI). Three kinds of proteins containing one free cysteine residues, β-lactoglobulin A, bovine serum albumins and papain were obtained from Sigma-Aldrich. All antibodies used for immunoblotting were purchased from Abcam except HRP-labeled Streptavidin (Beyotime Biotechnology), and other chemical or biological reagents were obtained from commercial suppliers without any manipulation.

2.2. Cell Culture and Preparation of Cell Lysates

HeLa cells were obtained from China Center for Type Culture Collection (Wuhan, China) and cultured at 37 °C under a humidified 5% CO_2 atmosphere in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum, and 1% penicillin-streptomycin. For in-gel fluorescence scanning, western blots and mass spectrometry experiments, cells are grown to 80-90% confluence and were collected by cell scraper and pelleted by centrifugation at 1 000 rpm for 5 min at 25 °C, followed by washing with PBS (1 mL) three times and the cell pellets were flash frozen in liquid nitrogen and stored in -80 °C for use. Cell pellets were then resuspended followed by ultrasonication and centrifugation at 4 °C for 30 min at 20 000 g. The resulting supernatant (soluble cell lysate) was collected and protein concentration was determined via BCA assay (Pierce, Thermo Scientific).

2.3. Small molecule experiments

All of the reagents used were freshly prepared, and the experiments were carried out at room temperature. Except for reactions performed at different pH, all reactions used PBS (pH=7.4) or ultrapure water. Se-probes were dissolved in anhydrous DMSO and the final concentration of NPSPyne, NPSSyne and DiSeNDyne used for Fmoc-Cys labeling was 100 μM, 100 μM and 50 μM, respectively. For LC-MS analysis, Fmoc-Cys or GSH aqueous solutions were added to Se-probes solutions with a final concentration of 20 μM. 5-Fold molar excess of Se-probes was added into the small thiol molecule samples and incubated for 3 min in the dark, after that, NEM (200 μM) was added to the mixed solution and incubated for 5 min, followed by LC-MS analysis. 2.5 μM TBTA was utilized as the internal standard. To test the influence of reaction time, the reaction mixture mentioned above was tested by LC-MS at different reaction time

points. As for pH experiments, the procedures were the same except that the concentration of each compound was changed, and the subsequent desalination operation was performed. The final concentration of Fmoc-Cys and Se-probe was 100 μM and 500 μM, respectively. And the final concentration of NEM was 2 mM. At the designated time, precooled acetonitrile (600 μL) was added into reaction mixture (100 μL) and the mixtures were left at -20 °C for 1h. Then the mixtures were centrifuged (20000 g for 10 min) at 4°C to separate the aqueous and organic phase. 450 μL of the organic solution were aspirated and dried in vacuum. For LCQ Fleet analysis, the samples were dissolved in 120 μL ultrapure water containing 1 μM of the internal standard (TBTA). Full-scan mass spectra were acquired over the m/z range from 100 to 1000 using the ion trap mass analyzer with positive ion mode. Extracted ion chromatogram of relevant product were integrated and the overall intensity were compared.

2.4. In-Gel Fluorescence Scanning and western blotting

The soluble proteomes were normalized to 2 mg/mL, incubated with different concentrations of Se-probes or IA-alkyne at 25 °C for 1 h. The labeled proteomes were precipitated by adding PBS and methanol/chloroform (v/v = 4:1) mixture with 1:2.5 volumetric ratio, and centrifuged at 10000 g for 10 min. The supernatants were discarded, and the proteomes were washed using cold methanol and resuspended in 0.4% (w/v) SDS/PBS. For in-gel fluorescence scanning, Cy3-azide (200 μM) was added into the proteomes for click chemistry (100 μM TBTA, 1 mM $CuSO_4$ and 1 mM TCEP or 200 μM BTTAA, 100 μM $CuSO_4$ and 5 mM ascorbate sodium). Samples were allowed to react for 1h at room temperature and quenched with 5× loading buffer. The proteomes were boiled using sample buffer at 90 °C for 5 min, resolved on 10% SDS-PAGE gels and visualized by fluorescence using Cy3 channel. The gels were then stained by Coomassie brilliant blue (CBB) to demonstrate equal loading. For western blotting, Biotin-PEG3-azide (200 μM) was added into the proteomes for click chemistry (100 μM TBTA, 1 mM $CuSO_4$ and 1 mM TCEP). Samples were allowed to react for 1h at room temperature and precipitated by methanol/chloroform (v/v = 4:1) as described above. Then the pull-down and western blotting were conducted following the standard experimental steps. All antibodies were used according to the manufacturer's protocols and blotting was imaged using a ChemiDoc™ Touch Imaging System (Bio-Rad).

2.5. Labeling of purified proteins

Human albumin (HSA, 2 mg/mL), β-lactoglobulin A (2 mg/mL) and papain (2 mg/mL) were prepared in PBS. Se-probes (0.5 mM) or IA-alkyne (0.5 mM) was added into each protein solutions and the mixtures were incubated at 25 °C for 1 h. After induction, proteins were filtered with NAP-5 columns (GE Healthcare) to remove small molecule species. A fraction of β-lactoglobulin A was diluted and desalted with Zeba Spin Desalting column (Thermo Fisher Scientific) for intact mass analysis by LCQ Fleet. The labeled proteins were conjugated with Cy3-N_3 under two CuAAC conditions as described above. After click chemistry, a fraction of β-lactoglobulin A was diluted and desalted for mass analysis and a fraction of HSA was precipitated by acetone for color contrast. The remaining proteins was subjected to SDS-PAGE and visualized by fluorescence using Cy3 channel.

2.6. Labeling of peptides containing cysteines

Four peptides (GLCTVAML, TLACFVLAAV, GCSWDYKN, HCLGKWLGHPDKF, 10 μM) were dissolved in PBS. NPSSyne (100 μM) was added into the mixed solutions and was incubated at 25 °C for 1 h. After that, a fraction of the mixture was desalted with ZipTip C18 (Millipore) for LC-MS/MS. The rest of the mixture is divided evenly into two parts. Both were conjugated with DADPS-Biotin-Azide and Biotin-PEG3-azide via CuAAC (200 μM BTTAA, 100 μM $CuSO_4$ and 5 mM ascorbate sodium), respectively. The mixtures were then subjected to streptavidin enrichment for 2 h and the peptides were eluted by 5% HCOOH or TCEP. The samples were dried under vacuum and reconstituted in 10 μL of 0.1% (v/v) formic acid in water for further analysis of LC-MS/MS.

2.7. RD-ABPP for Quantifying proteins labeled by Se-probes

Cell lysates (1mL, 2 mg/mL) were labeled with Se-probes (100 μM) or DMSO (control) and allowed to incubate at 25 °C for 1 hour. The proteomes were precipitated with 2.5 mL methanol/chloroform (v: v=4:1), washed two times with 1 mL of ice-cold MeOH and resuspended in 0.4% SDS/PBS. The resuspended samples were added with 200 μM Biotin-PEG3-azide, 1 mM CuSO4, 100 μM TBTA ligand, and 1 mM TCEP to react for 1 h. The proteomes were precipitated again with methanol/chloroform, resuspended in 1.2% SDS/PBS, diluted to 0.2% SDS/PBS and subjected to streptavidin enrichment for 3h with rotation. The beads were then washed with 1×5 mL of 0.2% SDS/PBS, 3×5 mL of PBS and 3×5 mL of water, and were denatured in 6 M urea/TEAB (100 mM), reduced with 10 mM dithiothreitol (DTT) at 37 °C for 30min and alkylated with 20 mM IAA at 35 °C for 30 min in dark. The redundant IAA in mixtures was quenched with isometric DTT at 37 °C for 30 min. To this bead-mixture was added certain amount of TEAB (100 mM) to dilute solution to 2 M urea/TEAB and 4 μL of trypsin (Promega, 20 μg reconstituted in 40 μL of the TEAB buffer) and incubated at 37 °C for 16 h. Then 4% of light or heavy formaldehyde was added to Se-probe or control proteomes, with simultaneous addition of 0.6 M sodium cyanoborohydride. The reaction was lasted for 1 h at room temperature and sequentially quenched by 1% ammonia and 5% formic acid. Finally, heavy and light labeled peptides were combined, concentrated by speed vacuum, and resuspended into 100μL of water. Then the samples were centrifuged (100,000g, 10min, 4°C) and fractionated with a fast sequence workflow by dual reverse phase high performance liquid chromatography (RP-HPLC)[46]. The samples were dried under vacuum, combined into 10 fractions and reconstituted in 10 μL of 0.1% (v/v) formic acid in water. Finally, all fractions were performed on LC-M/MS.

2.8. TOP-ABPP for identification of cysteine sites labeled by Se-probes

Cell lysates (1.5 mL, 2 mg/mL) were labeled with Se-probes (100 μM) and allowed to incubate at 25 °C for 1 hour. The proteomes were then treated with 200 μM DADPS-Biotin-Azide, 1 mM $CuSO_4$, 100 μM TBTA ligand, and 1 mM TCEP for 1 h. After that, the proteomes were precipitated with methanol/chloroform, resuspended in 1.2% SDS/PBS, and diluted to 0.2% SDS/PBS. The samples were then subjected to streptavidin enrichment, and on-bead trypsin digestion according to the published TOP-ABPP protocol.[47] The reducing reagents such as DTT or TCEP should not be used in the bead-mixtures. The trypsin digest was separated from the beads through 1400g centrifuge, and the beads were washed with 3 × 300 μL PBS, 3 × 300 μL H_2O. The Si-O bond cleavage was carried out by incubating the beads with 3 × 100 μL 5% formic acid for 1 h at room temperature with gentle rotation. The supernatant was collected and stored at -20 °C.

2.9. IA-alkyne-based rdTOP-ABPP analysis of Se-SMs-modified cysteines

Cell lysates (1mL, 2 mg/mL) were labeled with Se-SMs (100 μM) or DMSO (control) and allowed to incubate at 25 °C for 1 hour. After that, the lysates were incubated with IA-alkyne (100 μM) at 25 °C for 1 h. The resulting lysates were precipitated with 2.5 mL methanol/chloroform (v: v=4:1), washed two times with 1 mL of ice-cold MeOH and resuspended in 0.4% SDS/PBS. The samples were then treated with Desthiobiotin-PEG3-Azide, followed with streptavidin enrichment and on-bead trypsin digestion as described above. The beads were collected and washed by 1 mL of distilled water and 1 ml of 100 mM TEAB buffer for twice before being suspended in 200 μL of 100 mM TEAB buffer. Then 8

μL of 4% $D^{13}CDO$ (Sigma-Aldrich) or HCHO (Sigma-Aldrich) was added to the vehicle and Se-SMs-treated samples to be light and heavy labeled, respectively. Then 8 μL of 0.6 M $NaBH_3CN$ was added to the samples. After being washed by 100 mM TEAB buffer twice, the two samples were mixed and washed with 1 mL distilled water twice, followed by on-bead elution with 3×200 mL 1% TFA in 70% CH_3CN at room temperature. The released peptides were concentrated using a vacuum centrifuge, reconstituted in 10 μL of 0.1% (v/v) formic acid in water and then subjected for LC–MS/MS analysis.

2.10. LC-MS/MS Analysis

LC-MS/MS was performed on a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled with an EASY nLC 1200 system. The peptides were separated using an increasing gradient of aqueous acetonitrile with 0.1% formic acid. Mobile phase A was 0.1% FA in H_2O, and mobile phase B was 0.1% FA in 80% ACN. The flow rate was 3 μL/min for loading and 0.3 μL/min for eluting. Labeled peptide samples were loaded onto an EASY-Spray column (Thermo Fisher Scientific, ES900) packed with 15 cm × 3 μm C18 resin. Under the positive-ion mode, full-scan mass spectra were acquired over the m/z range from 350 to 2000 using the Orbitrap mass analyzer with mass resolution of 70 000. MS/ MS fragmentation is performed in a data-dependent mode, in which the 20 most intense ions were selected from each full-scan mass spectrum for fragmentation by high-energy collision induced dissociation (HCD). MS/MS spectra were acquired with a resolution of 17,500 using the Orbitrap analyzer. Other important parameters: isolation window, 1.6 m/z units; default charge, 2+; normalized collision energy, 28%; maximum IT, 50 ms; dynamic exclusion, 20.0 s.

2.11. Peptide identification, quantification and and bioinformatic analysis

The obtained raw data was inputted into the Proteome Discoverer software (Thermofisher Scientific, version 2.1.1) for the specific protein information searching in Human UniProt database (release-2021_02). The parameters of "Sequest HT" were set as follows: Trypsin (Full) was selected as the digestion enzyme. Max. Missed Cleavage Sites was 2, Min. Peptide Length was 6 and Max. Peptide Length was 144. Precusor Mass Tolerance was 10 ppm and Fragment Mass Tolerance was 0.02 Da. Cysteine carbamidomethylation (+57.0215 Da), N-terminus dimethyl labeling, and lysine residue dimethyl labeling were set in Static Modification section with tag masses of +28.03130 Da (light), +34.06312Da Da (heavy), respectively. Dynamic modifications on differential oxidation of methionine are set at 15.9949 Da and cysteines are set at 179.94782 Da (for alkyne-PhSe tag) in analysis of peptides containing the cysteine, 323.05368 Da (for acid-cleavable tag) in TOP-ABPP or 551.34313 Da (for desthiobiotin-eluotropic tag). "Protein FDR Validator" were set as follows: Target FDR (Strict)=0.01, Target FDR (Relaxed)=0.05. "Peptide and Proteins Quantifier" were set as follows: Peptide to use: Unique+Razor peptide; Maximum allowed fold Change:15; Top N peptides used for area C :3; Use single-peak quan channels: False; Normalization mode: None. The crystal structures were downloaded from AlphaFold and visualized using VMD. GO enrichment analysis results, KEGG and Disease analysis were also obtained in The Database for Annotation, Visualization and Integrated Discovery (DAVID).*2.12 Statistical analysis.*Results are expressed as mean ± s.d. Fold change in relation to control groups of three independent cell culture and subsequent procedures. We analyzed the data in GraphPad Prism (GraphPad Software), using the unpaired, two tailed t-test module. Statistical significance was considered when a P value was below 0.05.

3. Results

3.1. Chemoselective labeling of sulfhydryl on purified small thiol molecules, proteins and peptides

First, NPSSyne and NPSPyne containing the Se–N bond with similar structures were designed and synthesized as labeling probes (Fig. 1A). The structures of phthalimide and succinimide in the probes can help to form relatively stable Se–N bond with phenylseleno (PhSe). Considering the hydrolyzation tendency of both probes in aqueous solution to produce diselenide, and the possibility of Se–S bond formation through reaction between thiol and diselenide similar to thiol-disulfide exchange, we also designed and synthesized the diselenide DPDSyne as a labeling probe (Fig. 1A). We assessed the reactivity of NPSSyne and NPSPyne towards Fmoc-Cys, a small, stable molecule that contains a thiol group (Fig. S1A). The corresponding MS1 chromatographic peak of the labeling product emerged in the presence of excess NPSPyne or NPSSyne, and the Fmoc-Cys peak was invisible after 3 min, indicating the complete labeling of Fmoc-Cys (Fig. 1B) by the Se-probes. Notably, the MS of the labeling product displayed conspicuous selenium isotopes with detectable distribution (Fig. S1B). The similar results were also obtained through labeling of GSH (20 μM) by NPSPyne (100 μM, Fig. S2). In addition, DPDSyne, the hydrolysate of NPSPyne and NPSSyne, were also observed by UV chromatogram ($t_R \approx 19.90$, Fig. S1C), which confirmed previous research that NPSS and NPSP are susceptible to hydrolysis and can rapidly convert into diselenide [44, 45]. To assess the labeling efficacy of Fmoc-Cys by DPDSyne, DPDSyne (50 μM) was added into Fmoc-Cys (20 μM) samples and incubated for 3 min, N-Ethylmaleimide (NEM, 500 μM)) was then added. The MS1 and UV chromatograms in Fig. S3 revealed that only a small fraction of Fmoc-Cys was labeled by DPDSyne, with NEM predominantly consuming Fmoc-Cys, which indicated that the reaction between DPDSyne and Fmoc-Cys was considerably slower than that of NPSPyne and NPSSyne. These findings suggest that NPSSyne and NPSPyne with Se–N bond reacts with the thiol, despite the hydrolyzation and diselenide formation.

To further prove that NPSPyne and NPSSyne still keep the same character of Se–N bond as that in NPSP, three representative purified proteins which contain only one free thiol group, including human albumin (HSA), β-lactoglobulin A and papain were utilized as substrates to explore the reactivities of the Se-probes. After incubation of three purified proteins by NPSSyne or NPSPyne for 1 h in PBS, the mixture was clicked with Cy3-azide, and the labeled proteins were separated by SDS-PAGE and analyzed by in-gel fluorescence scanning. Meanwhile, a widely used electrophilic probe for cysteine, IA-Alkyne, was used as a control. As shown in Fig. 1C, bands with stronger fluorescence were observed from the samples treated with NPSSyne or NPSPyne than IAA-alkyne treated samples via two different CuAAC conditions, indicating that the Se-probes are more efficient and faster (Figs. S4A and S4D). The observation of deep-color acetone-precipitation of HSA labeling by the Se-probes and subsequent conjugation with Cy3-azide was consistent with in-gel images Fig. S4B). Furthermore, ESI-MS spectra of β-lactoglobulin A labeled by NPSPyne and NPSSyne were obtained and the mass difference between tagged and untreated protein ions is 179 Da in the deconvoluted spectra, indicating the attachment of one alkyne-PhSe-tag to the sole free cysteine site of β-lactoglobulin A (Fig. 1D). Meanwhile, ESI-MS spectrum of β-lactoglobulin A labeling by NPSSyne and conjugation with Cy3-azide was successfully obtained (Fig. S4C). Hence, purified protein experiments demonstrated that the reactivity of Se–N bond in the Se-probes remains the same as that in ebselen and NPSP.

The mixture of four peptides (GLCTVAML, TLACFVLAAV, GCSWDYKN, HCLGKWLGHPDKF) was labeled with excess NPSPyne (STP1), followed by biological orthogonal reaction, enrichment, elution, HCOOH acid cleavage (STP2) or TCEP reductive cleavage (STP3) (Scheme S1). As shown in Fig. 1E and F, MS1 of all target peptides in STP1 and STP2 displayed the unique isotopic envelope pattern

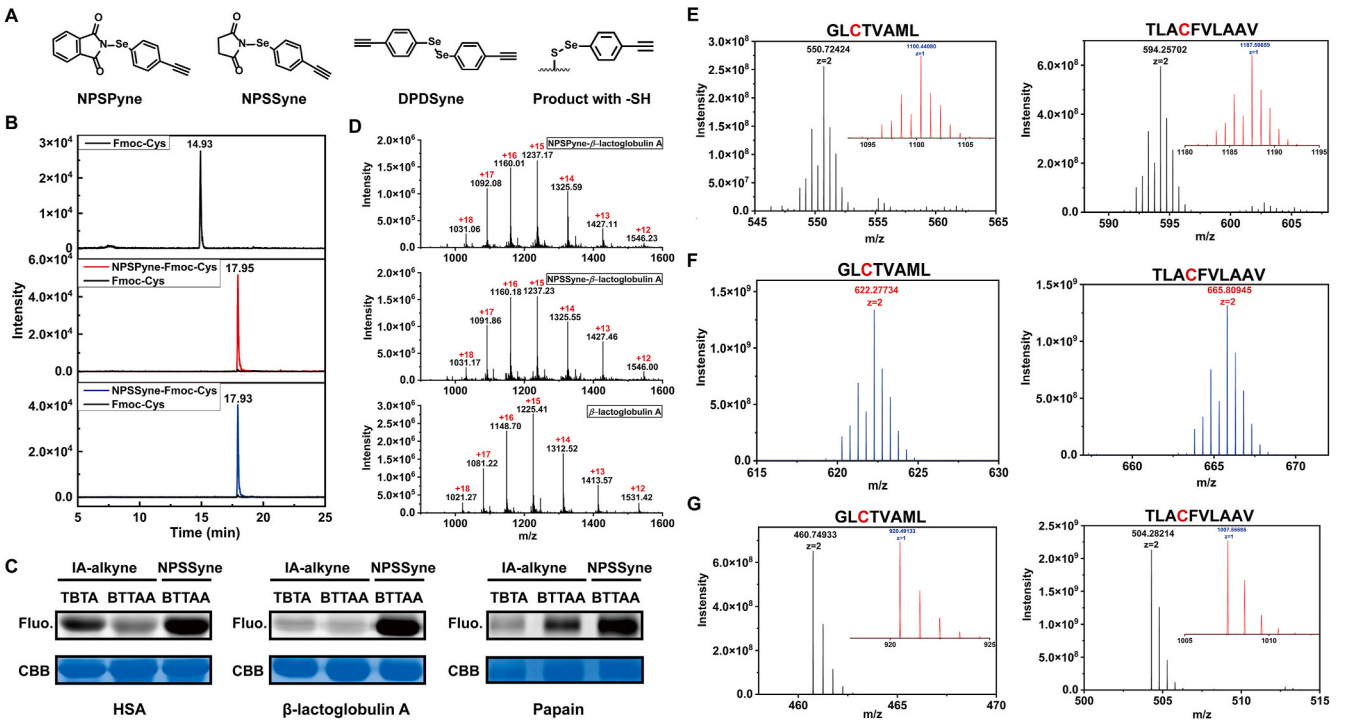

Fig. 1. Chemoselective labeling of thiol on small thiol molecules, purified proteins and peptides. (A) Structures of three Se-probes and their same reaction product with thiol. (B) Extracted MS1 chromatograms of reaction products between NPSPyne or NPSSyne (100 μM) and Fmoc-Cys (20 μM) monitored by LC-MS (tR = 17.95, 17.93). (C) Labeling of HSA, β-lactoglobulin A and papain (2 mg/mL) by IA-alkyne and NPSSyne (500 μM) detected through SDS-PAGE after Cy3-azide coupling via CuAAC. Fluo.: in-gel fluorescence scanning. CBB: Coomassie gel. CuAAC conditions: TBTA-CuSO4-TCEP, BTTAA-CuSO4-VcNa. (D) Full mass-spectra showing β-lactoglobulin A labeling by NPSPyne and NPSSyne. (E)–(G) MS1 of GLC*TVAML (left) and TLAC*FVLAAV (right) in STP1, STP2 and STP3, respectively. The upper right insertion represents isotope distribution of z = 1 for this peptide.

attributing to alkyne-PhSe-tag on cysteines, and the *m/z* difference of peptides between single isotope peaks in STP1, STP2 and untreated peptides match Δm/z of alkyne-PhSe-tag (in STP1, Figs. S6A and S6D) and conjugate-PhSe-tag (in STP2, Figs. S6B and S6E). Both isotopic envelope pattern and *m/z* of enriched peptides reduced by TCEP were the same as these of untreated peptides (Fig. 2F and Figs. S6C and S6F).

Take the case of GLCTVAML as an example, the *m/z* difference of peptides in MS1 between single isotope (z = 2) in STP1, STP2 and in STP3 were Δ89.97491 Da (alkyne-PhSe-tag) and 161.52801 Da (conjugate-PhSe-tag). The mass differences of the corresponding fragment ions were consistent with the simulation of MS/MS spectrum (Figs. S5–6). Hence it can be concluded that the peptides were labeled by NPSSyne at cysteine

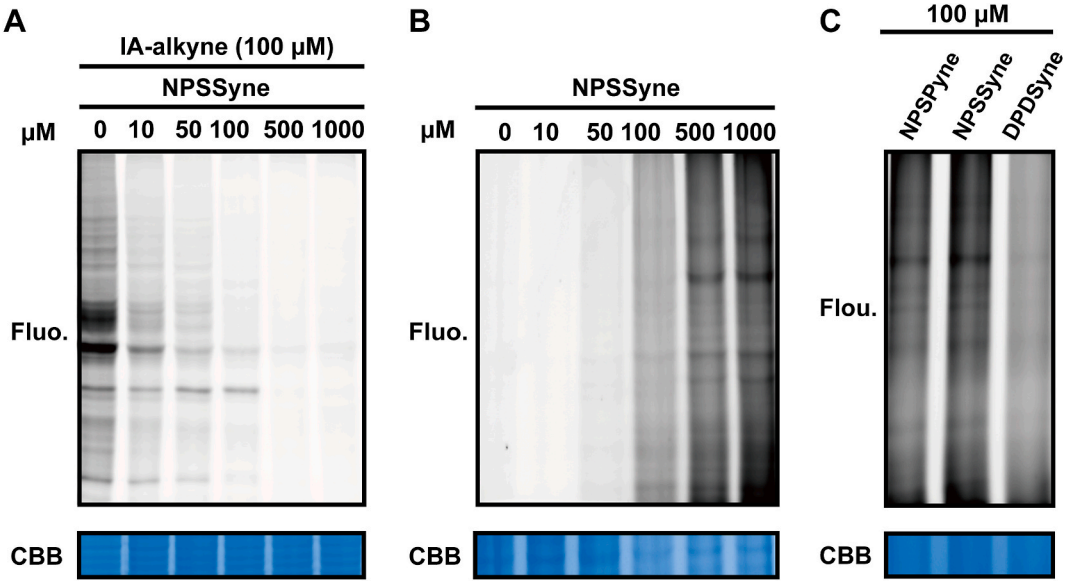

Fig. 2. Se-probes labeling characterization by gel-ABPP in HeLa proteomes. (A) Gel-ABPP of HeLa cell lysates treated with indicated concentrations of NPSSyne for 1 h and IA-alkyne (100 μM) for another 1 h at 25 °C, conjugated with Cy3-azide via CuAAC and visualized on SDS-PAGE. Loading buffer + β-mercaptoethanol. (B) Gel-ABPP of HeLa cell lysates treated with indicated concentrations of NPSSyne for 1 h at 25 °C, conjugated with Cy3-azide via CuAAC and visualized on SDS-PAGE. (C) Comparison of HeLa cell lysates labeled by three Se-probes, with each conjugated with Cy3-azide and visualized by in-gel fluorescence.

sites and PhSe-tag would be valuable in pinpointing the location of cysteine residues of peptides which would also facilitate peptide sequencing. Overall, our data suggest that Se-probes can be used to selectively derivatize the cysteine residues in proteins and peptides, and the introduced selenium tags will be beneficial to identify cysteine sites in complex protein samples.

3.2. Characterization of Se-probes labeling by gel-ABPP in HeLa proteomes

Subsequently, we conducted a comprehensive analysis of cysteine proteins in HeLa cell lysates using gel-based ABPP, and optimized the experimental conditions for subsequent investigations. Competition experiments were conducted by treating HeLa cell lysates with varying concentrations of Se-probes, followed by the addition of specific amounts of IA-alkyne (100 μM). The labeled proteins were then conjugated to Cy3-azide via CuAAC for in-gel fluorescence analysis. β-Mercaptoethanol (2-ME) were selectively added in the loading buffer in order to break the formed Se–S bonds, besides the intended S–S breakage. Fig. 2A and Fig. S7A demonstrated that increasing concentrations of NPSSyne and NPSPyne led to reduced protein IA-alkyne labeling in the competition experiments. Similar competitive effects were observed with NPSP and NPSS (Figs. S7B–C). Notably, the competitive ability of Se-probes on protein labeling was comparable to that of classical sulfhydryl derivatization reagents, including Iodoacetamide (IAA), NEM, S-Methyl methanethiolsulfonate (MMST), and 2-(methylsulfonyl)benzothiazole (MSBT), as shown in Fig. S7D. The aforementioned phenomena indicate that Se-probes could label a significant number of cysteines in proteomes. On the other hand, NPSSyne and NPSPyne exhibit dose-dependent labeling of cysteine proteins in HeLa cell lysates, from both the in-gel fluorescence images, and the $CH_3OH/CHCl_3$ precipitation with SDS/PBS resuspension of cell lysates, ranging from 0 to 1000 μM (Fig. 2B and S8). Time-dependent experiments demonstrated that the Se-probes exhibited fast labeling for targeted proteins in as little as 10 min and kept a stable signal in 2 h (Fig. S9). Significantly, the majority of proteome-wide protein bands labeled by NPSPyne and NPSSyne in gel were imperceptible upon the addition of 2-ME in the loading buffer, providing additional evidence of Se–S bond formation in the Se-probes derivatized cysteine residues, which can be disrupted by 2-ME (Fig. S10A). In agreement with low Fmoc-Cys labeling efficiency by DPDSyne, the labeling of cysteine proteins throughout the proteome by NPSPyne and NPSSyne was marginally greater than that of DPDSyne (Fig. 2C). There was no significant difference whether NAP-5 columns are used or not to remove small molecule species, which demonstrated the small thiol molecules do not have an impact on the subsequent proteomic MS-workflow (Fig. S10B). When different cell types including HeLa, HCT116, A549, and L-02 cells were investigated, the Se-probe labeled protein pattern demonstrated a

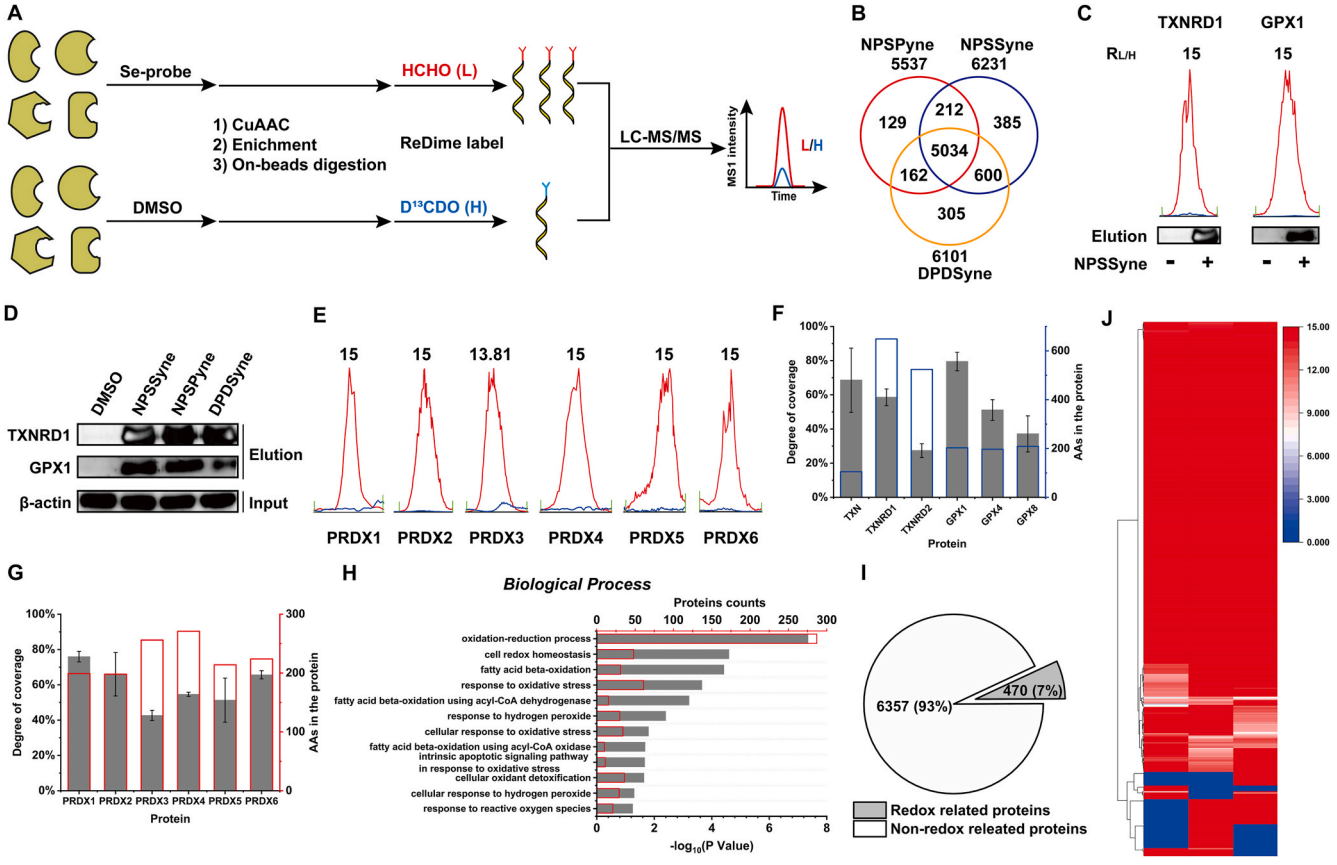

Fig. 3. Quantitative MS-based profiling of potential target proteins in HeLa proteomes with Se-probes. (A) Schematic workflow for MS-based profiling of potential target proteins including probe vs. control (DMSO). (B) Venn diagram showing the number of potential target proteins quantified in probe vs. DMSO ($R_{L/H} \geq 10$) labeled by NPSPyne, NPSSyne and DiSeNDyne. (C) Extracted MS1 chromatograms of tryptic peptides from representative proteins TXNRD1 and GPX1 with enrichment ratios ($R_{L/H}$), and Western blots shown in HeLa proteomes. (D) Western blots confirming selected proteins enrichment by Se-probes. (E) Chromatographic peaks of selected proteins (PRDX1-6) labeled by NPSSyne. Red and blue are traces of light and heavy dimethylated peptides. The protein names and quantified ratios are shown below and above, respectively. (F) and (G) Coverage of representative proteins associated with intracellular redox in NPSSyne enrichment experiments. (H) Bioinformatics analysis of target proteins of Se-probes based on involvement in specific biological processes. (I) Proportion of the potential target proteins related to intracellular redox in terms of biological process. (J) The quantified ratios for the 470 proteins related to intracellular redox are characterized on heat map. A value of 0 in the color level indicates no quantitative information obtained. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

high degree of consistency across various cell-lysates. However, the L-02 cell-lysate displayed distinctive protein targets (Fig. S11). Based on the above findings, we selected 100 μM Se-probes, with a 1-h incubation period and TBTA-$CuSO_4$-TCEP conditions, to label and identify proteins and cysteine sites for subsequent MS-based investigations.

3.3. Mass spectrometry (MS)-based quantitative profiling of Se-probes-reactive proteins

We next performed MS-based quantitative profiling of Se-probes-reactive proteins by ABPP using stable isotope reductive dimethylation method (ReDiMe, RD-ABPP, Fig. 2A and Table S1) [46]. We assigned the proteins with an enrichment ratio over 10 in at least two of the three biological replicates as the candidate reactive Se-probes target proteins. 5537, 6231 and 6101 potential target proteins were identified

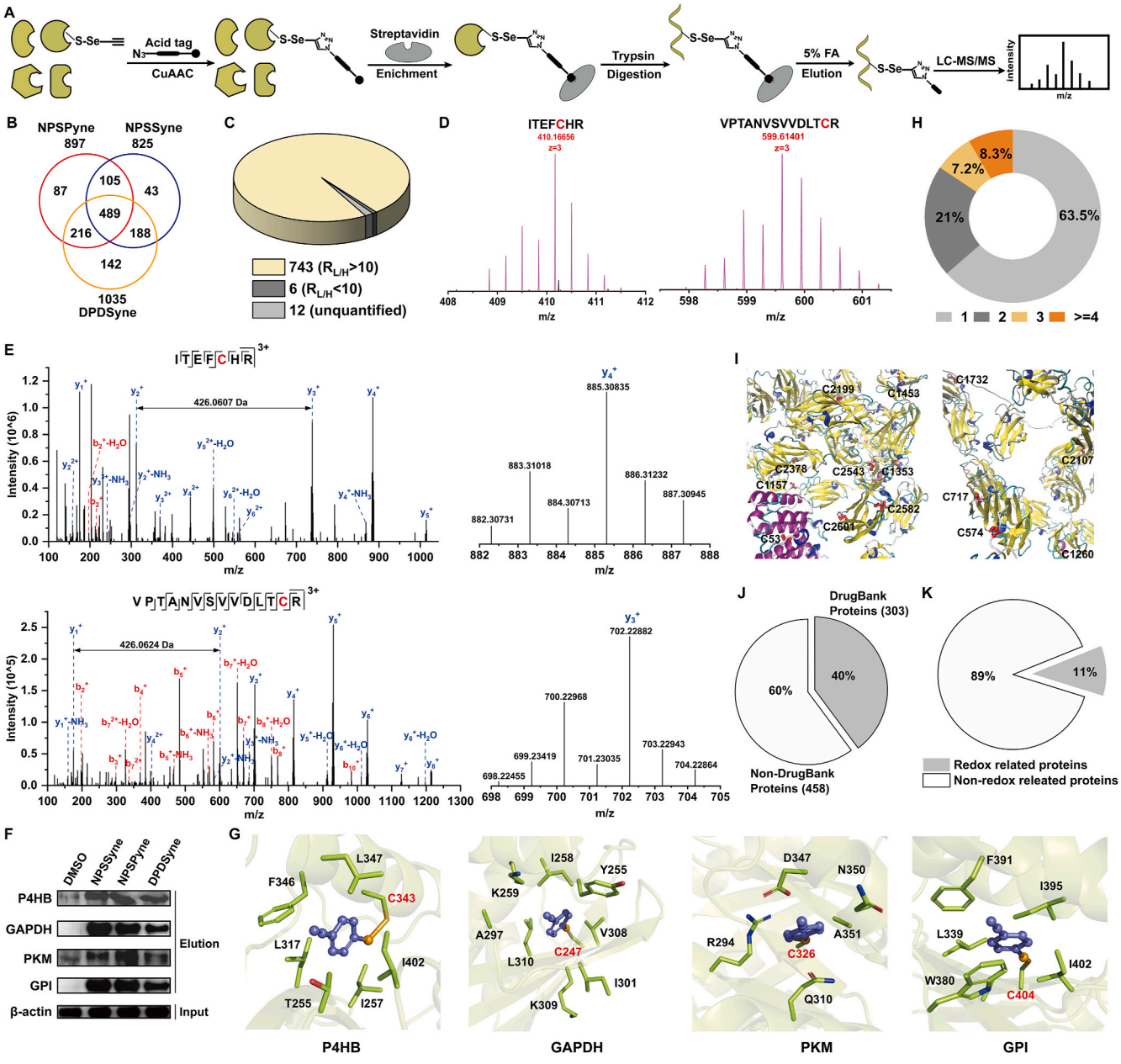

Fig. 4. Identification of cysteine sites from HeLa cell lysates induced with Se-probes by TOP-ABPP. (A) Schematic workflow for MS-based identification of cysteine residue sites using a tandem orthogonal proteolysis strategy. (B) Venn diagram showing the number of cysteine sites labeled by NPSPyne, NPSSyne and DiSeNDyne. (C) Categorization of 761 proteins with identified sites by three Se-probes based on distribution of their quantification ratios, $R_{H/L}$. (D) MS1 with the corresponding selenium isotope distribution of ITEFC*HR (z = 3) and VPTANVSVVDLTC*R (z = 3) shown as the light purple line on the left and right, respectively. (E) MS/MS spectra of ITEFC*HR (z = 3, top) from P4HB and VPTANVSVVDLTC*R (z = 3, down) from GAPDH. The *m/z* difference of 426.0607 Da between y2+ and y3+ (top) and 426.0624 Da between y1+ and y2+ (down) supports the expected modification on the cysteine residue. (F) Western blots confirming selected proteins P4HB, GAPDH, PKM and GPI enrichment by Se-probes. (G) Example structures of the proteins P4HB (PDB: 1BJX), GAPDH (PDB: 1U8F), PKM (PDB: 1T5A) and GPI (PDB: 1IAT) with the reactive cysteine residues modified by Se-probes. (H) Distribution of the number of the identified cysteine sites per protein. (I) The specific location of every cysteine site labeled by Se-probes in FLAN protein secondary structure. (J) Fraction of proteins possessing labeled cysteine sites found in DrugBank database. (K) Proportion of the target proteins related to intracellular redox in terms of biological process in site experiment results. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

by NPSPyne, NPSSyne and DPDSyne through this quantitative, MS-based chemical proteomic platform, respectively (Fig. 3B and Figs. S12A–B). The proteins labeled by any of the three Se-probes were identified as prospective sulfydryl-reactive target proteins, encompassing a total of 6827 proteins, of which 5034 were common to all three probes (Fig. 3B). The Western blot analysis demonstrated a concordance between the labeling characteristics and the prior in-gel fluorescence findings, thereby validating the specific labeling of modified proteomes by Se-probes (Fig. S12C). Noticeably, several important proteins involving in redox homeostasis such as thioredoxin (TXN), thioredoxin reductase-2 (TXNRD1-2), glutathione peroxidase 1 (GPX1), phospholipid hydroperoxide glutathione peroxidase (GPX4) and glutathione peroxidase 8 (GPX8) were identified (Fig. 3C and Figs. S12C and S13A). For validation, two candidate proteins TXNRD1 and GPX1 were investigated by affinity purification and immunoblotting experiments, and the results were consistent with their distribution on the mean ratio plot and the extracted MS1 chromatograms of representative peptides (Fig. 3D and Fig. S13B) obtained by mass spectrometry. Besides, peroxiredoxins (PRDX1-6), a family of proteins that are extremely effective at scavenging peroxides, were also identified by the three Se-probes (Fig. 3E and Fig. S13C). These results indicated that Se-probes could affect the antioxidant defense in cells by targeting important redox-regulating enzymes.

Next, we counted the amino acid coverage of potential targets enriched by the Se-probes (Fig. S14A). About 42 % proteins had an amino acid coverage of more than 20 %, which suggests that to some extent, our approach can indeed enrich as many proteins as possible. The coverage of representative proteins mentioned above were all more than 30 % except TXNRD2, and the coverage of most peroxiredoxins reached more than 40 % (Fig. 3F–G and Figs. S14B–C). Gene ontology analysis by DAVID (Database for Annotation, Visualization and Integrated Discovery) revealed that the Se-probes reactive proteins were centered on important biological processes such as cell-cell adhesion, rRNA-processing as well as translational initiation (Fig. 3H and Table S1). Further analysis revealed that 7 % of the total candidate proteins (470 proteins) were involved in various intracellular redox biological processes, which further demonstrated that Se-probes exert their role in redox-regulating activities (Fig. 3I–J and Fig. S15, Table S1).

3.4. *Identification of cysteine sites using tandem orthogonal proteolysis strategy (TOP-ABPP)*

We next sought to identify exact modified cysteine sites by Se-probes using TOP-ABPP strategy and the acid cleavable azide-biotin tag was used to enrich and release the modified peptides for LC-MS/MS analysis (Fig. 4A, Scheme S2A, Fig. S16A and Table S2) [47]. The cysteine residue sites that appeared once in three replicates labeled by NPSPyne, NPSSyne and DPDSyne were regarded as exact sites of modification by Se-probes. 897, 825 and 1035 sites were identified by NPSPyne, NPSSyne and DPDSyne by TOP-ABPP. In total, 1270 sites of 761 proteins were identified by Se-probes (Fig. 3B and Fig. S17). A majority ($\geq$97 %) of the proteins containing Se-probes reactive sites were quantified ($R_{L/H} \geq 10$) from RD-ABPP and only 12 proteins belonged to unquantified targets, suggesting the reliability of the RD-ABPP quantitative results (Fig. 4C and Fig. S17). Besides, more than 93 % of the proteins with identified sites by Se-probes had amino acid coverages greater than 20 % (Fig. S18).

We randomly selected several cysteine site-containing peptides from TOP-ABPP results to test whether Se-tag were introduced into the corresponding peptides. As shown in Fig. 4D and Figs. S19–20, unique isotopic envelope pattern of ITEFCHR (z = 3) and VPTANVSVVDLTCR (z = 3 and z = 2) were displayed, which were essentially identical to their simulated peaks. The MS/MS spectrum generated by higher-energy collisional dissociation (HCD) fragmentation unambiguously confirmed the adduct of Se-tag (Fig. 4E and Fig. S20C) on Cys343 of Protein disulfide-isomerase (PH4B) and on Cys247 of Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The MS1 and MS/MS spectra of other selected cysteine site-containing peptides also exhibited similar coincidence, such as LFQCLLHR from X-ray repair cross-complementing protein 5 (XRCC5X_C493, Fig. S21), IADLCHTFIK from translation initiation factor eIF-2B subunit alpha (EIF2B1_C115, Fig. S22), TFTAWCNSHLR from alpha-actinin-1 (ACTIN-1_C60) or/and alpha-actinin-4 (ACTIN-4_C41, Fig. S23), AGKPVICATQMLESMIK from pyruvate kinase PKM (PKM_C326, Fig. S24) and MIPCDFLIPVQTQHPIRK from glucose-6-phosphate isomerase (GPI_C404, Fig. S25). The enrichment of P4HB, GAPDH, PKM and GPI by Se-probes were further proved by affinity purification and immunoblotting (Fig. 4F and Fig. S26A). The results of the molecular docking analysis indicate that the organic selenium probes have the ability to covalently label cysteine residues located within the inner cavity of these proteins, thereby forming Se–S bonds (Fig. 4G and Fig. S26B).

Unique isotopic envelope patterns for the fragment ions of the corresponding peptides in MS/MS spectra (Fig. 4E and Figs. S19C, S20D, S21D, S23D, S25D) were also observed, demonstrating that the Se-tag from Se-probes derivatization can survive in the HCD process, which undoubtedly promote the credibility of the cysteine residue sites identification. Moreover, even though there is no unique isotopic envelope pattern presented in MS1 due to low relative abundance of the target peptide, the fragment ions containing Se-tag in MS/MS spectra can provide complementary evidence for identification results. For example, although a unique isotopic envelope pattern for FINYVKCFR in MS1 was not found, the *m/z* difference of 426.0625 Da between y2+ and y3+ and the specific isotope distribution of y4+, y5+and y6+ were sufficient to confirm the cysteine site (nucleophosmin, NPM1_C275) labeled by the Se-probes (Fig. S27). The complementary nature of the isotopic patterns observed in MS1 and MS2 holds promise as a significant indicator for the future identification and validation of selenium drug targets.

We further analyzed the number of identified cysteine sites on each protein and found that nearly two thirds (63.5 %) of the proteins contained only one modified cysteine, while 8.3 % identified proteins had at least four modified cysteine sites (Fig. 4H and Fig. S28A). For example, Filamin-A (FLAN), which plays an important role in conformity cell mechanics and signal transduction, was detected with 14 cysteine sites by Se-probes labeling (Fig. 4I and Figs. S29A–B) [48]. Fatty acid synthase (FASN), an important biosynthetic enzyme, which plays an important role in catalyzing fatty acid synthesis, was detected with 11 cysteine sites (Figs. S29C–D) [49]. We also studied the relationship between the protein length and the number of the modified cysteine residues per protein, but no significant correlation was found (Fig. S28B). It is noteworthy that 60 % of the newly identified proteins by Se-probes are recorded as non-ligandable according to the DrugBank database (Fig. 4J), which provides very important and valuable information for the development of selenium-containing drugs [50]. Further analysis revealed that 81 (11 %) of parent proteins of the identified cysteine sites were involved in various intracellular redox biological processes and they are almost all included in proteins related to intracellular redox in enrichment experiments (Fig. 4K and Fig. S30). Meanwhile, the specific sites of proteins involving in redox homeostasis appeared in previous results of RD-ABPP were also identified, such as TXN_C73, TXNRD1_C577, TXNRD1_C625, GPX1_C78, GPX1_C156, GPX4_93, GPX4_C102, PRDX1_C173, PRDX2_C172, PRDX3_C229, PRDX4_C51, PRDX4_C245, PRDX5_C100, PRDX6_C47, PRDX6_C91 (Figs. S31–32 and Table 1). Besides the target peptides identified, a very small amount of the oxidized Se–S bonds was also formed, demonstrating that the protein cysteines may undergo further oxidation during sample preparation, but the amount is almost negligible (Fig. S33). The results suggests that Se-probes can be used as a powerful tool to study intracellular redox-regulating proteins.

Table 1
MS analysis for the specific sites of representative proteins involving in redox homeostasis and quantitative result in previous results of RD-ABPP.

Uniprot #	Gene name	Site	Peptited sequence	Ratio (NPSP/DMSO)	Ratio (NPSS/DMSO)	Ratio (DPDS/DMSO)
P10599	TXN	73	CMPTFQFFK	13.12	15	15
Q16881	TXNRD1	577	IICNTK	15	15	15
		625	KQLDSTIGIHPVCAEVFTTLSVTK			
P07203	GPX1	78	GLVVLGFPCNQFGHQENAK	15	15	15
		156	LITWSPVCR			
P36969	GPX4	93	YAECGLR	15	15	15
		102	ILAFPCNQFGK			
Q06830	PRDX1	173	HGEVCPAGWKPGSDTIKPDVQK	15	15	15
P32119	PRDX2	172	LVQAFQYTDEHGEVCPAGW-KPGSDTIKPNVDDSK	15	15	15
P30048	PRDX3	229	AFQYVETHGEVCPANW-TPDSPTIKPSPAASK	15	13.81	15
Q13162	PRDX4	51	TREEECHFYAGGQVYPGEASR	15	15	15
		245	HGEVCPAGWKPGSETIIPDPAGK			
P30044	PRDX5	100	GVLFGVPGAFTPGCSK	15	15	15
P30041	PRDX6	47	DFTPVCTTELGR	15	15	15
		91	DINAYNCEEPTEKLPFPIIDDR			

3.5. Identification of cysteine sites using competitive rdTOP-ABPP strategy

Although we had successfully identified cysteine sites labeled by Se-probes using TOP-ABPP strategy, there was a discrepancy between the number of identified sites and the observation of gel-based ABPP, because it appeared that the labeling of cysteines in the proteome by IA-alkyne could be almost completely competed by Se-probes (Fig. 2A and Fig. S7A). Therefore, we decided to use the competitive rdTOP-ABPP strategy to reidentify the cysteine sites labeled by Se-probes [51,52]. According to Fig. 5A, after ReDiMe procedure, the beads were amalgamated and subjected to on-beads elution, following which the desthiobiotin-adducted peptides were released and analyzed through LC-MS/MS. (Fig. S34A and Table S3). The competitive labeling of IA-alkyne-modified proteomes with three Se-probes were confirmed by Western blots (Fig. S34B), which was in coincidence with in-gel fluorescence (Fig. S7D). We designated the peptides with $R \geq 4$ in at least two of the three biological replicates as the target peptides. The proteomic results were displayed in Fig. 5B and Figs. S34C–E and in total, 4434 cysteine residues in 2074 proteins were identified using rdTOP-ABPP strategy. The number of proteins with identified target sites was expanded from 761 to 2074, 95 % of target proteins in rdTOP-ABPP were quantified ($R_{L/H} \geq 10$, sFig. 5C and Fig. S34F) and more than two-thirds of the proteins had amino acid coverages greater than 20 % (Fig. S35), reflecting the quantitative accuracy of RD-ABPP and rd-TOP-ABPP methods.

To further verify the accuracy of mass spectrometry data in rdTOP-ABPP and compare the differences of MS results for the same target peptide between the TOP-ABPP and rdTOP-ABPP, VPTANVSVVDLTCR (z = 3, GAPDH_C247) was chosen again. The MS1 of VPTANVSVVDLTCR (z = 3) in rdTOP-ABPP showed a pair of classical isotopic envelope patterns, with a difference of 2Da resulting from reductive dimethylation, compared to the unique isotopic peaks in TOP-ABPP (Figs. 4C and 5D and Figs. S36A–B). The data in MS/MS spectrum can match the fragment ion of the peptide and the *m/z* difference of 654.3534 Da between y1+ and y2+ supports the expected modification of desthiobiotin-tag on the cysteine residue (Fig. 5E). Meanwhile, the extracted MS1 chromatograms of the peptides and their $R_{L/H}$ on a mean ratio plot were verified by affinity purification and immunoblotting of GAPDH (Fig. 5E). For two proteins, TXNRD1 and P4HB, the labeling by IA-alkyne was competed by Se-SMs with different efficiency (Fig. 5F and Fig. S37). These data indicated the reliability of MS in rdTOP-ABPP. Exceptionally, rdTOP-ABPP strategy did not obtain quantitative information of the corresponding target peptide on the GPI, although the labeling by IA-alkyne for this protein was also competed by Se-SMs with different degrees (Fig. 5F and Fig. S37), which is likely to be related to our experimental operation. As expected, the fragment ions in MS/MS spectrum in rdTOP-ABPP displayed a featureless isotopic envelope pattern in comparison to that in TOP-ABPP (Fig. S20D and Fig. S36C). On all accounts, the results of MS in rdTOP-ABPP can reveal the labeling of cysteine in the proteome by Se-SMs over a wider range.

Besides the 545 commonly identified cysteines, the Se-probes targeted a unique population of 725 cysteines on 544 proteins, which might explain its different fluorescence bands from the IA-alkyne modification in gel-based ABPP (Fig. 5G and Fig. S10). As far as commonly identified cysteines were concerned, the R ($_{L/H}$) in rdTOP-ABPP were diverse according to the different Se-SMs (Fig. 5G), which suggested that even different probes with the similar warhead have different labeling effect on the same cysteine site in proteome. It is worth noting that 11 of the 15 specific sites of redox related proteins found in results of TOP-ABPP were among the commonly identified cysteines (Fig. 5G–H and Fig. S38). This provides strong evidence once again that Se-SMs could disrupt the redox balance in cells by targeting important redox-regulating enzymes, which may provide some new ideas and basic theoretical support for the treatment of redox homeostasis-related diseases, such as tumor and inflammation.

We next sought to identify exact target sites of ebselen-modified proteins using the same rdTOP-ABPP strategy (Table S3). A total of 2677 sites in 1393 proteins were quantified across three biological replicates (Figs. S39A and B). In addition to 180 proteins reported previously with biotinylated ebselen probes [33], there are a good deal of proteins with specific ebselen-modified sites that have not been further studied or reported (Fig. 5I). Nonetheless, a considerable proportion of cysteine sites and proteins that underwent ebselen modification were still present in the roster of Se-probes modified cysteine sites and proteins, suggesting that multiple proteins can be potentially targeted by drugs utilizing Se–N bond or Se–Se bond. (Fig. 5J and Fig. S39C). Therefore, the proteomic information will provide the valuable information for the design and development of novel selenium-containing drugs.

3.6. Bioinformatics analysis of target protein and cysteine sites

Ultimately, the target proteins and active sites were scrutinized through the multi-faceted bioinformatic approaches. Protein class analysis using PANTHER classification system [53] revealed that these proteins are related to various protein types, such as metabolite interconversion enzyme, protein modifying enzyme, RNA metabolism protein, translational protein and cytoskeletal protein (Fig. 6A). Functional survey of these proteins revealed that many of them are correlated with different types of activities and functions, with the dominant groups having binding (42.7 %) and catalytic activity (35.9 %, Fig. 6B). GO analysis revealed that the target proteins were highly enriched in important biological processes such as translation, cytoplasmic

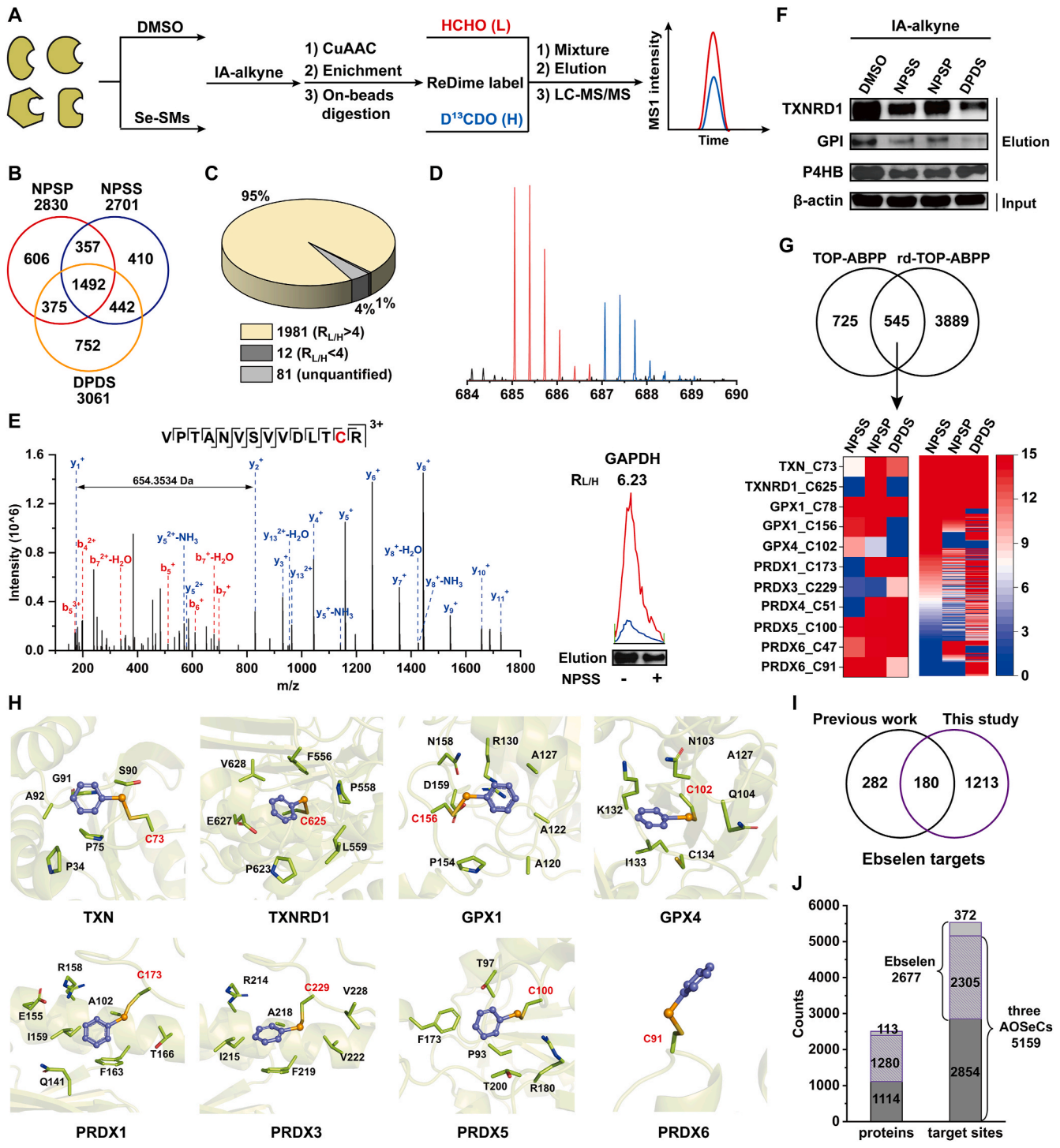

Fig. 5. Identification of cysteine sites from HeLa cell lysates induced with Se-SMs (NPSP, NPSS, DPDS) and IA-alkyne by rdTOP-ABPP. (A) Schematic workflow of rdTOP-ABPP for MS-based Identification of cysteine sites that includes probe vs. control (DMSO). (B) Venn diagram showing the number of cysteine residue sites labeled by Se-SMs and IA-alkyne and quantified in probe vs. DMSO ($R_{L/H} \geq 4$). (C) Distribution of the ratios of cysteine sites identified with Se-SMs and IA-alkyne by rdTOP-ABPP. (D) The MS1 (z = 3) with the corresponding isotope distribution of the final targeted peptide (Light and heavy is shown as red line and blue line, respectively) and mass-to-charge (*m/z*) ratios of light and heavy peaks. (E) MS/MS spectrum of VPTANVSVVDLTC*R (z = 3) from GAPDH. The *m/z* difference of 654.3534 Da between y1+ and y2+ supports the expected modification on the cysteine residue. Extracted MS1 chromatograms of the peptides with ratios ($R_{L/H}$) is shown and Western blots indicates the selected proteins are competed by NPSS with IA-alkyne in HeLa proteomes. (F) Western blots confirming selected proteins, TXNRD1, GPI and P4HB were competed by NPSS with IA-alkyne in HeLa proteomes. (G) Comparison of the number of cysteine sites identified by Se-probes based direct profiling (TOP-ABPP) with that by IA-alkyne-based competitive cysteine profiling (rdTOP-ABPP). The quantified ratios for each of the 545 overlapping sites are characterized on heat map and 11 cysteine sites of representative proteins associated with intracellular redox are highlighted. A value of 0 in the color level indicates no quantitative information obtained. (H) Structures of the proteins TXN (PDB: 1AIU), TXNRD1 (PDB: 2CFY), GPX1 (PDB: 2F8A), GPX4 (PDB: 2GS3), PRDX1 (PDB: 2RII), PRDX3 (PDB: 5GCJ), PRDX5 (PDB: 1H4O) and PRDX6 (PDB: 1PRX) with the reactive cysteine residues modified by Se-SMs. (I) Venn diagram showing 180 proteins found both by biotinylated ebselen probe and by rdTOP-ABPP. (J) The statistics of identification of cysteine sites from HeLa cell lysates induced with ebselen and IA-alkyne by rdTOP-ABPP. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

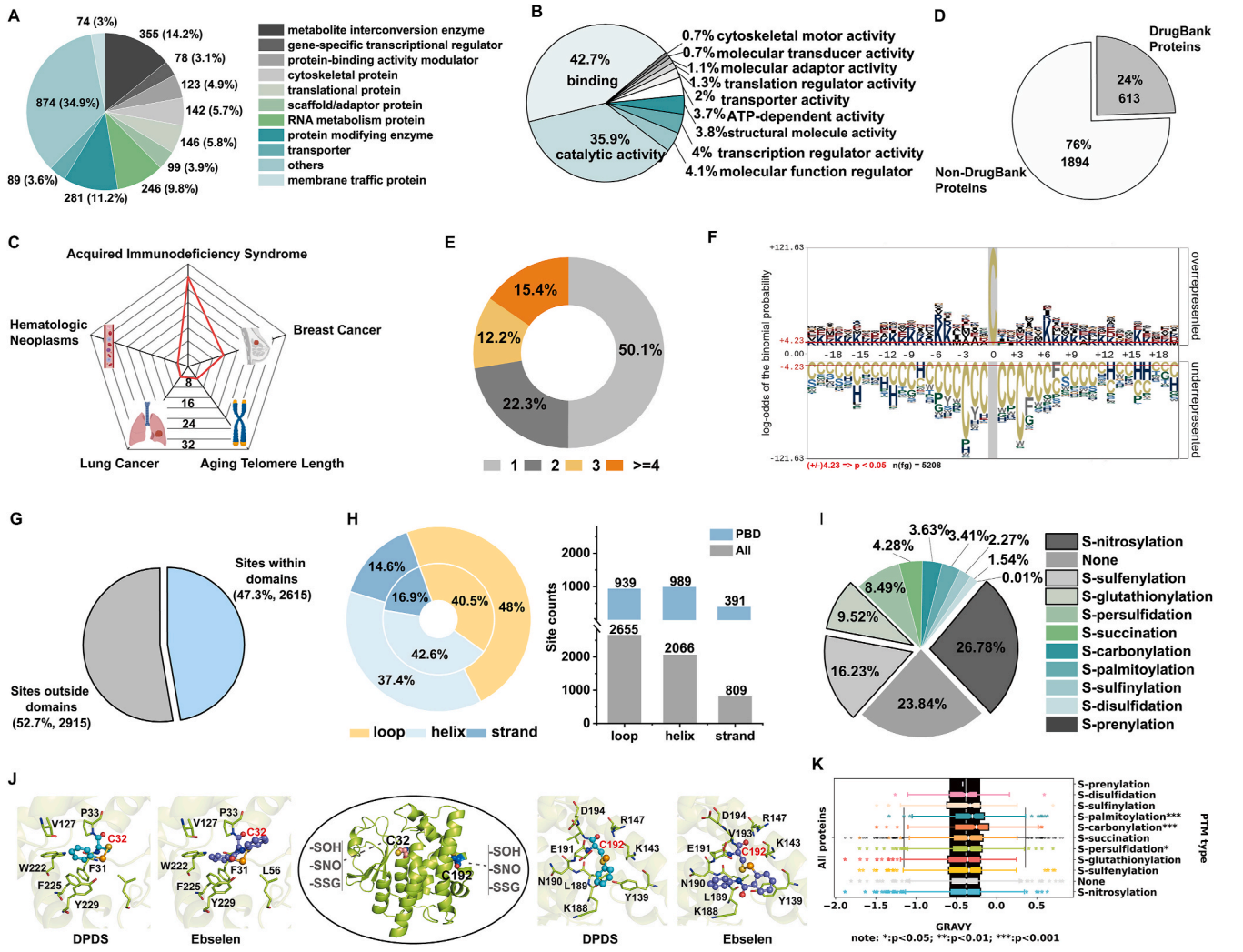

Fig. 6. Bioinformatics analysis of target protein and cysteine sites of Se-probes and Se-SMs. (A) Classification of target proteins. (B) Functional annotation of the proteins containing the cysteine sites. (C) Disease analysis of the target proteins by GO. (D) Fraction of proteins possessing labeled cysteine sites found in DrugBank database. (E) Distribution of the number of the identified cysteine sites per protein. (F) Sequence motif analysis of cysteine sites. (G) Domain analysis of the cysteine sites. (H) The distribution of the cysteine residues in different secondary structures of proteins. (I) The PTMs distribution of cysteine sites. (J) Structures of the proteins GSTO1 (PDB: 1AIU) with the reactive cysteine residues modified by Se-SMs. (K) GRAVY values of proteins with the modified cysteine residues and the values with significant differences in each PTMs.

translation, cell division as well as rRNA-processing (Fig. S40A). Cellular component analysis showed that Se-probes modified proteins are widely distributed in cells including cytosol, nucleoplasm and even extracellular exosome (Fig. S40B). Pathway and process analysis demonstrated that Se-probes modified proteins were overrepresented in nucleocytoplasmic transport, salmonella infection, carbon metabolism, biosynthesis of amino acids and ribosome, among others (Fig. S40C). We also analyzed these target proteins based on annotations recorded in the DAVID disease database and found that target proteins are highly enriched in cancer and aging diseases which are closely related to intracellular redox homeostasis such as breast cancer, lung cancer, hematologic neoplasms and aging telomere length (Fig. 6C). AIDS were found on the top of the list because HIV infection can cause dysregulation of glutathione homeostasis and induce lipid peroxidation [54,55].

For 2507 proteins harboring the identified cysteine residues, only 24 % were found in DrugBank and 76 % of these are considered non-ligandable according to the DrugBank database (Fig. 6D). We further analyzed the identified cysteine sites and found that more than half (50.1 %) of the proteins contained only one modified cysteine, while 15.4 % identified proteins had at least four modified cysteine sites (Fig. 6E and Fig. S41A). There was no significant correlation between the protein length and the number of the modified cysteine residues per protein (Fig. S41B). Notably, besides FLAN and FASN were identified with plentiful cysteine sites (21 and 22 respectively), DNA-dependent protein kinase catalytic subunit (PRKDC) was identified with 26 cysteine sites by using two strategies. Nevertheless, despite the fact that the microtubule-actin cross-linking factor 1 (MACF1) comprises 7388 amino acids, of which 92 are cysteines, solely the cysteine site is susceptible to labeling by Se-SMs. Hence, the Se-probes could in principle be applied for discovery of new ligands with therapeutic potential using identified reactive cysteines as anchors.

Despite the absence of conspicuous conserved motifs encircling the identified cysteine residues, lysine or arginine were observed to occur with greater frequency among overrepresented residues than other amino acids, which implies that Se-SMs are more prone to interact with cysteine containing basic amino acids in their vicinity (Fig. 6F and Fig. S42) [56]. This phenomenon can be attributed to the presence of the amino acid groups in close proximity to the target cysteines, which facilitates the dissociation of thiol and consequently lowers the pKa value of cysteine.

Python was utilized to conduct a comprehensive search and summary of the domain, secondary structure, and post-translational

modification of the target sites across multiple databases (Fig. S43). Of these cysteine sites, 47.3 % were located in different protein domains, including C/N-terminal domain, enzyme, binding domain, RNA domain, kinase domains and others, which indicated that these sites may participate in the mediation of protein functions (Fig. 6G, Fig. S44 and Table S4). We also investigated the relationship between the modified sites and protein secondary structures from PDB, AlphaFoldDB and PSIPRED (Fig. 6H and Table S4). Among all the cysteine residues, the percentage at the loop structure was the highest, while the lowest percentage came from the β-strand structure. Two statistical analyses were conducted, one utilizing all data and the other exclusively experimental data (PDB), with only a minor discrepancy, suggesting the dependability of secondary structure prediction.

Given that the protein cysteine can undergo multiple post-translational modifications (PTMs), acting as a molecular switch to maintain redox homeostasis and regulating a series of biological activities, the PTMs of target cysteine sites obtained from CysModDB and iPTMnet (Fig. 6I, Fig. S45 and Table S4) [57,58]. Based on experimentally verified PTMs data, it is evident that oxidation PTM constitutes the predominant proportion of modifications, including *S*-nitrosylation, *S*-sulfenylation, *S*-glutathionylation, and *S*-persulfidation. Glutathione S-transferase omega-1 (GSTO1) C32 and C192 have been shown to undergo various oxidative post-translational modifications, and we also found that these two cysteine sites can be modified by Se-SMs (Fig. 6I and Fig. S46A). In addition, GAPDH_C247 and PRDX6_C47_C91 mentioned above also have similar oxidative post-translational modification types and modification conditions (Fig. S46B). In comparison to the hydrophobicity of all proteins modified by Se-probes and Se-SMs, the hydrophobicity of proteins modified by multiple oxidation modifications did not exhibit a marked difference. However, the hydrophobicity distribution of S-palmitoylation and S-carbonylation was significantly greater than that of all proteins. The list is, to the best of our knowledge, the first database for OSeCs modified proteins in human proteome, which provides a rich resource for the community to explore the new active organic selenium drug.

4. Discussion

Proteome-wide profiling of sulfhydryl-containing proteins and sites by using specific Se-probes has not been systematically surveyed to date, despite the many advantages of Se–N and Se–S bond containing compounds. Herein, three probes with the alkynyl handle were synthesized and their rapid reaction, high selectivity toward cysteines were investigated on three dimensions: small molecules, peptides and purified proteins (Fig. 1 and Figs. S1–6). The in-gel fluorescence suggested that Se-probes possess a high coverage of cysteines in the proteome like IA-Alkyne (Fig. 2 and Figs. S7–11). While the proteome-wide labeling effects of NPSPyne and NPSSyne were slightly better than that of DPDSyne (Fig. 3D), the fluorescence bands with these probes did not show significant inconsistencies. As for the hydrolysis tendency of NPSPyne and NPSSyne, selenol or selenolate were rapidly produced in aqueous solution, which can be rapidly oxidized to diselenide or be conjugated through Se–S bonds with cysteines in proteome. And diselenide can also be nucleophilic attacked by sulfhydryl to form Se–S bond-like thiol-disulfide exchange [59,60]. Regardless of which type of reaction mechanism, in the case of cysteines in the proteome, the identical alkyne-PhSe-tag was eventually introduced into the target cysteine proteins in the form of Se–S bond (Scheme S2B), which would be favorable for subsequent analysis of mass spectrometry data. However, it is necessary to re-design and synthesize new probes with stable structure containing Se–N bond if further studies about quantitative profiling of Se–N bond-sensitive sites in living cell proteomes, which should certainly deserve consideration.

5034 common proteins were identified by three Se-probes in RD-ABPP indicating that the effectiveness of Se-probes for proteome-wide profiling of cysteines (Fig. 3B). The commonly labeled proteins showed consistencies with the in gel-ABPP fluorescence bands of the Se-probes. Most proteins with identified cysteine sites in TOP-ABPP and rdTOP-ABPP had amino acid coverage greater than 20 % (Fig. S18 and Fig. S35), reflecting the accuracy of quantitative results of RD-ABPP and rdTOP-ABPP strategies. But to our surprise, only 1270 cysteine sites in 761 proteins were identified by TOP-ABPP strategies, which was a far cry from the thousands of proteins identified in RD-ABPP. Possible reasons for this might be as follows: 1) The low concentration of the probes (100 μM) resulted in inadequate labeling of all target proteins in cell lysates. 2) The reducing reagent such as DTT or TCEP, were not utilized during the workflow of TOP-ABPP to avoid reduction of Se–S bond. Therefore, the disulfide bonds in target proteins were not entirely reduced, which resulted in incomplete digestion of the target proteins. 3) The selenium isotope envelope might increase the complexity of selecting parent ions for secondary fragmentation in data dependence acquisition (DDA) pattern. Therefore, in the follow-up study, improvement measures are worth considering and taking, such as increasing the concentration of the probe, enhancing restriction sites with mixed enzyme, and usage of data independence acquisition (DIA) pattern, etc. [61].

The distinctive isotopic envelope pattern observed in the mass spectrum of corresponding peptides can provide the crucial and significant information. In the MS1 data, the isotopic pattern of selenium can be observed in cysteine sites in the target peptides, owning to the introduction of Se-tag. This pattern is notably different from that of common peptide (Figs. 4D and 5D and Figs. S21–S25). Even with an isolation window is set to 1.6 *m/z*, the fragment ions can still display the isotopic pattern since Se-tag can survive in the HCD process (Fig. 4E and Figs. S21–S25). The target peptide exhibits a distinctive peak pattern akin to donning an elaborate coat, enabling its swift identification amidst the mass spectra peaks. Moreover, these two isotopic patterns in MS1 and MS2 can complement each other (Fig. S27), which can enhance the credibility of the cysteine sites identification. We envision that parallel reaction monitoring (PRM) method with the diverse dissociation modes will be an ideal method for target peptide verification. Furthermore, more satisfactory results could be obtained if data acquisition algorithms such as SESTAR and SESTAR++ for selenium-encoded isotopic signature are used, which can bypass the abundance bias for fragmentation to generate MS/MS spectra [38,62]. Using specialized selenium isotope envelop identification programme, selenium probes could be used to explore selenium-containing covalent drugs and identify their targets.

It is important to note that Se-probes also have some weaknesses in the proteome MS-workflow and application for cysteine site identification. For example, Se–S bonds are incompatible with high concentrations of reducing agents such as DTT and TCEP, and subsequent procedures for protein treatment without reducing agents would result in the inability to reductively cleave disulfide bonds and alkylate free sulfydryl within proteins. Complex isotopic envelopes also increase the degree of difficulty in proteomic quantification methods. For example, complex isotope peaks have the potential to cause overlap of spectra of different peptides in MS1-level quantification, thereby confounding quantification. The presence of multiple isotope peaks makes the target peptide unsuitable for labelling with TMT or iTRAQ, which restricts the probe from being used in multiplex LC-MS analysis. In addition, the impact of complex isotopes also needs to be taken into account if accurate quantitative results are to be achieved in the SWATH process, and specific algorithms may need to be introduced into the search programme to distinguish between these isotope peaks and their secondary ion peaks.

When competitive rdTOP-ABPP strategy were explored combining IA-alkyne usage, we identified a total of 4434 Se-probes-modified cysteine sites on 2074 proteins in HeLa cell lysates, most of which were not included in TOP-ABPP results (Fig. 5B and G). In a sense, the Se-probes and IA-alkyne can complement each other for the identification of cysteine sites. In total, 5159 cysteines on 2394 target proteins

were identified with TOP-ABPP and rdTOP-ABPP strategy. This dataset provides valuable resources for probing the functional roles of small molecules containing the Se–N bond or diselenide moiety. Among these targets, we found several key enzymes with redox-regulating functions including TXN, TXNRD1, GPX1, GPX4, PRDX1-6 and the specific cysteine sites of them were identified in both TOP-ABPP and rdTOP-ABPP (Fig. 3E–G, 5G, 5H and Figs. S12, S13, S31, S32). In addition, many proteins involving in intracellular redox were enriched and identified by Se-probes (Fig. 3I, J, 4K and Figs. S15 and S30), suggesting that Se-probes can be used to study redox-regulating proteins. In the future, Se-probes could also be used in conjunction with a variety of quantitative proteomic approaches to explore changes in the expression of redox-regulated proteins under specific conditions or to study the key cysteines on a redox-regulated protein under specific physiological activities.

In summary, we have reported the development and application of specific selenium-containing probes based on Se–N bond and Se–Se bond to chemoselectively identify cysteine residues in the human proteome. Combining bioorthogonal chemistry, three Se-probes with multiplexed proteomics were used for global profiling of thiol-containing proteins and cysteine sites in the human proteome. The unique isotopic envelope patterns in MS1 and MS2 from introduced Se-tag could facilitate peptide sequencing and were valuable in pinpointing the location of cysteine residues of peptides, which relieved further visual biology verification. We anticipate that the selenium isotopic signature will play a persuasive role in cysteine-directed chemical proteomics and design and screening of selenium-based drugs. Going forward, we will fully exploit the superiority of Se–N and Se–Se bonds in identifying sulfhydryl-containing proteins for the development of more novel probes or inhibitors to solve various pathological problems in vivo.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD046847, PXD046852, PXD046856. The data that support the findings of this study are available from the corresponding authors upon request.

Acknowledgements

This work was supported by National Natural Science Foundation of China (21927811, 22134004, 92253304 and 12275163), Natural Science Foundation of Shandong Province (ZR2020MB098, ZR2020MB074 and ZR2021MB068), and Local Science and Technology Development Fund Guided by the Central Government of Shandong Province (YDZX2022012).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.redox.2023.102969.

References

[1] N.M. Giles, G.I. Giles, C. Jacob, Multiple roles of cysteine in biocatalysis, Biochem. Biophys. Res. Commun. 300 (2003) 1–4.
[2] S.M. Marino, V.N. Gladyshev, Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces, J. Mol. Biol. 404 (2010) 902–916.
[3] A. Miseta, P. Csutora, Relationship between the occurrence of cysteine in proteins and the complexity of organisms, Mol. Biol. Evol. 17 (2000) 1232–1239.
[4] S.M. Couvertier, Y. Zhou, E. Weerapana, Chemical-proteomic strategies to investigate cysteine posttranslational modifications, Biochim. Biophys. Acta, Proteins Proteomics 1844 (2014) 2315–2330.
[5] C.E. Paulsen, K.S. Carroll, Cysteine-mediated redox signaling: chemistry, biology, and tools for discovery, Chem. Rev. 113 (2013) 4633–4679.
[6] M. Abo, E. Weerapana, A caged electrophilic probe for global analysis of cysteine reactivity in living cells, J. Am. Chem. Soc. 137 (2015) 7087–7090.
[7] M. Abo, D.W. Bak, E. Weerapana, Optimization of caged electrophiles for improved monitoring of cysteine reactivity in living cells, Chembiochem 18 (2017) 81–84.
[8] D.A. Shannon, et al., Investigating the proteome reactivity and selectivity of aryl halides, J. Am. Chem. Soc. 136 (2014) 3330–3333.
[9] D. Abegg, et al., Proteome-wide profiling of targets of cysteine reactive small molecules by using ethynyl benziodoxolone reagents, Angew. Chem. Int. Ed. 127 (2015) 11002–11007.
[10] D. Abegg, et al., Chemoproteomic profiling by cysteine fluoroalkylation reveals myrocin G as an inhibitor of the nonhomologous end joining DNA repair pathway, J. Am. Chem. Soc. 143 (2021) 20332–20342.
[11] E.W. McConnell, A.L. Smythers, L.M. Hicks, Maleimide-based chemical proteomics for quantitative analysis of cysteine reactivity, J. Am. Soc. Mass Spectrom. 31 (2020) 1697–1705.
[12] R. Wang, et al., Low-toxicity sulfonium-based probes for cysteine-specific profiling in live cells, Anal. Chem. 94 (2022) 4366–4372.
[13] H.F. Motiwala, Y.H. Kuo, B.L. Stinger, B.A. Palfey, B.R. Martin, Tunable heteroaromatic sulfones enhance in-cell cysteine profiling, J. Am. Chem. Soc. 142 (2020) 1801–1810.
[14] R.N. Reddi, et al., Tunable methacrylamides for covalent ligand directed release chemistry, J. Am. Chem. Soc. 143 (2021) 4979–4992.
[15] C. Zambaldo, et al., 2-Sulfonylpyridines as tunable, cysteine-reactive electrophiles, J. Am. Chem. Soc. 142 (2020) 8972–8979.
[16] N. Shindo, et al., Bicyclobutane carboxylic amide as a cysteine-directed strained electrophile for selective targeting of proteins, J. Am. Chem. Soc. 142 (2020) 18522–18531.
[17] T.-Y. Koo, H. Lai, D.K. Nomura, C.Y.-S. Chung, N-Acryloylindole-alkyne (NAIA) enables imaging and profiling new ligandable cysteines and oxidized thiols by chemoproteomics, Nat. Commun. 14 (2023) 3564.
[18] E. Weerapana, et al., Quantitative reactivity profiling predicts functional cysteines in proteomes, Nature 468 (2010) 790–797.
[19] C. Tian, et al., Multiplexed thiol reactivity profiling for target discovery of electrophilic natural products, Cell Chem. Biol. 24 (2017) 1416–1427.e5.
[20] E.A. Grossman, et al., Covalent ligand discovery against druggable hotspots targeted by anti-cancer natural products, Cell Chem. Biol. 24 (2017) 1368–1376.e4.
[21] K.M. Backus, et al., Proteome-wide covalent ligand discovery in native biological systems, Nature 534 (2016) 570–574.
[22] Y. Zhang, W. Qin, C. Wang, Discovery of post-translational modifications in immunometabolism by chemical proteomics, Curr. Opin. Biotechnol. 68 (2021) 37–43.
[23] M.L. Nielsen, M. Vermeulen, T. Bonaldi, J. Cox, L. Moroder, M. Mann, Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry, Nat. Methods 5 (2008) 459.
[24] R.G. Fruchter, A.M. Crestfield, The specific alkylation by Iodoacetamide of histidine-12 in the active site of ribonuclease, J. Biol. Chem. 242 (1967) 5807–5812.
[25] M. Galvani, M. Hamdan, B. Herbert, P.G. Righetti, Alkylation kinetics of proteins in preparation for two-dimensional maps: a matrix assisted laser desorption/ionization-mass spectrometry investigation, Electrophoresis 22 (2001) 2058–2065.
[26] W. Xiao, Y. Chen, C. Wang, Quantitative chemoproteomic methods for reactive cysteinome profiling, Isr. J. Chem. 63 (2023), e20220010.
[27] C. Woo, A. Iavarone, D. Spiciarich, et al., Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N-and O-glycopeptide discovery and analysis, Nat. Methods 12 (2015) 561–567.
[28] K. Xu, et al., Study of highly selective and efficient thiol derivatization using selenium reagents by mass spectrometry, Anal. Chem. 82 (2010) 6926–6932.
[29] B. Tang, et al., A rhodamine-based fluorescent probe containing a Se-N bond for detecting thiols and its application in living cells, J. Am. Chem. Soc. 129 (2007) 11666–11667.
[30] B. Tang, et al., A fast-response, highly sensitive and specific organoselenium fluorescent probe for thiols and its application in bioimaging, Chem. Commun. 2 (2009) 5293–5295.
[31] L. Li, et al., An accurate mass spectrometric approach for the simultaneous comparison of GSH, Cys, and Hcy in L02 cells and HepG2 cells using new NPSP isotope probes, Chem. Commun. 51 (2015) 11317–11320.
[32] Z. Chen, et al., Fabrication of a “selenium signature” chemical probe-modified paper substrate for simultaneous and efficient determination of biothiols by paper spray mass spectrometry, Anal. Chem. 93 (2021) 1749–1756.
[33] Z. Chen, et al., Target discovery of ebselen with a biotinylated probe, Chem. Commun. 54 (2018) 9506–9509.
[34] Y. Zhang, H. Zhang, W. Cui, H. Chen, Tandem MS analysis of selenamide-derivatized peptide ions, J. Am. Soc. Mass Spectrom. 22 (2011) 1610–1621.
[35] Z. Wang, Y. Zhang, H. Zhang, P.B. Harrington, H. Chen, Fast and selective modification of thiol proteins/peptides by N-(phenylseleno) phthalimide, J. Am. Soc. Mass Spectrom. 23 (2012) 520–529.
[36] W.R. Parker, D.D. Holden, V.C. Cotham, H. Xu, J.S. Brodbelt, Cysteine-selective peptide identification: selenium-based chromophore for selective S-Se bond

cleavage with 266 nm ultraviolet photodissociation, Anal. Chem. 88 (2016) 7222–7229.
[37] W.R. Parker, J.S. Brodbelt, Characterization of the cysteine content in proteins utilizing cysteine selenylation with 266 nm ultraviolet photodissociation (UVPD), J. Am. Soc. Mass Spectrom. 27 (2016) 1344–1350.
[38] J. Gao, et al., Selenium-encoded isotopic signature targeted profiling, ACS Cent. Sci. 4 (2018) 960–970.
[39] J. Pei, X. Pan, G. Wei, Y. Hua, Research progress of glutathione peroxidase family (GPX) in redoxidation, Front. Pharmacol. 14 (2023), 1147414.
[40] G. Powis, W.R. Montfort, Annu. Rev. Pharmacol. Toxicol. 41 (2001) 261–295.
[41] H.J. Reich, R.J. Hondal, Why nature chose selenium, ACS Chem. Biol. 11 (2016) 821–841.
[42] N.V. Barbosa, et al., Organoselenium compounds as mimics of selenoproteins and thiol modifier agents, Metallomics 9 (2017) 1703–1734.
[43] P.A. Nogara, M.E. Pereira, C.S. Oliveira, L. Orian, J.B.T. Rocha, Organic selenocompounds: are they the panacea for human illnesses? New J. Chem. 47 (2023) 9959–9988.
[44] H.J. Reich, J.M. Renga, Organoselenium chemistry. Preparation and reactions of benzeneselenenamides, J. Org. Chem. 40 (1975) 3313–3314.
[45] H.J. Reich, C.A. Hoeger, W.W. Willis, Organoselenium chemistry. Characterization of reactive intermediates in the selenoxide syn elimination: selenenic acids and selenolseleninate esters, J. Am. Chem. Soc. 104 (1982) 2936–2937.
[46] N. Chen, et al., Chemical proteomic profiling of protein: N-homocysteinylation with a thioester probe, Chem. Sci. 9 (2018) 2826–2850.
[47] E. Weerapana, A.E. Speers, B.F. Cravatt, Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)-A general method for mapping sites of probe modification in proteomes, Nat. Protoc. 2 (2007) 1414–1425.
[48] A. Van der Flier, A. Sonnenberg, Structural and functional aspects of filamins, Biochim. Biophys. Acta Mol. Cell Res. 1538 (2001) 99–117.
[49] A. Jayakumar, et al., Human fatty acid synthase: properties and molecular cloning, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 8695–8699.
[50] D.S. Wishart, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.
[51] F. Yang, J. Gao, J. Che, G. Jia, C. Wang, A dimethyl-labeling-based strategy for site-specifically quantitative chemical proteomics, Anal. Chem. 90 (2018) 9576–9582.
[52] X.-F. Cheng, et al., Quantitative chemoproteomic profiling of targets of Au(I) complexes by competitive activity-based protein profiling, Bioconjugate Chem. 33 (2022) 1131–1137.
[53] H. Mi, et al., Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0), Nat. Protoc. 14 (2019) 703–721.
[54] L.A. Herzenberg, et al., Glutathione deficiency is associated with impaired survival in HIV disease, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 1967–1972.
[55] S. Thenin-Houssier, et al., Ebselen, a small-molecule capsid inhibitor of HIV-1 replication, Antimicrob. Agents Chemother. 60 (2016) 2195–2208.
[56] J.P. O'Shea, et al., pLogo: a probabilistic approach to visualizing sequence motifs, Nat. Methods 10 (2013) 1211–1212.
[57] Y. Meng, et al., CysModDB: a comprehensive platform with the integration of manually curated resources and analysis tools for cysteine posttranslational modifications, Briefings Bioinf. 23 (2022) bbac460.
[58] H. Huang, et al., iPTMnet: an integrated resource for protein post-translational modification network discovery, Nucleic Acids Res. 46 (2018) D542–D550.
[59] Y. Xue, et al., Selenylsulfide bond-launched reduction-responsive superparamagnetic nanogel combined of acid-responsiveness for achievement of efficient therapy with low side effect, ACS Appl. Mater. Interfaces 9 (2017) 30253–30257.
[60] S.F. Sousa, R.P.P. Neves, S.O. Waheed, P.A. Fernandes, M.J. Ramos, Structural and mechanistic aspects of S-S bonds in the thioredoxin-like family of proteins, Biol. Chem. 400 (2019) 575–587.
[61] F. Yang, G. Jia, J. Guo, Y. Liu, C. Wang, Quantitative chemoproteomic profiling with data-independent acquisition-based mass spectrometry, J. Am. Chem. Soc. 144 (2022) 901–911.
[62] G. Jia, J. Gao, F. Yang, T. Feng, C. Wang, An accelerated and optimized algorithm of selenium-encoded isotopic signature targeted profiling for global selenoproteome analysis, Methods Enzymol. 662 (2022) 241–258.