Global profiling of functional histidines in live cells using small molecule photosensitizer and chemical probe relay labeling

Yansheng Zhai
Shenzhen Bay Laboratory
Xinyu Zhang
Shenzhen Bay Laboratory
Dingyuan Yan
Shenzhen University
Xianghe Wang
Peking University
Kailu Tian
Shenzhen Bay Laboratory
Zhe Zhang
Shenzhen Bay Laboratory
Yan Huang
Shenzhen Bay Laboratory
Wen Sun
State Key Laboratory of Fine Chemicals, Dalian University of Technology https://orcid.org/0000-0003-4316-5350
Xi Yang
Shenzhen Bay Laboratory
Dong Wang
Shenzhen University
Yu-Hsuan Tsai
Shenzhen Bay Laboratory
Gang Li (ligang@szbl.ac.cn)
Shenzhen Bay Laboratory

Article

Keywords:

Posted Date: June 21st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2981935/v1

Additional Declarations: There is **NO** Competing Interest.

Global profiling of functional histidines in live cells using small molecule photosensitizer and chemical probe relay labeling

Yansheng Zhai[1], Xinyu Zhang[1,2], Dingyuan Yan[3], Xianghe Wang[4], Kailu Tian[1], Zhe Zhang[1,5], Yan Huang[1], Wen Sun[6], Xi Yang[1], Dong Wang[3], Yu-Hsuan Tsai [7]& Gang Li[1*]

[1]Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518132, China

[2]State Key Laboratory of Crop Stress Biology for Arid Areas, College of Life Sciences, Northwest A & F University, Yangling 712100 China

[3]Center for AIE Research, Shenzhen Key Laboratory of Polymer Science and Technology, Guangdong Research Center for Interfacial Engineering of Functional Materials, College of Materials Science and Engineering, Shenzhen University, Shenzhen 518060, China

[4]Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

[5]School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China

[6]State Key Laboratory of Fine Chemicals, Frontiers Science Center for Smart Materials Oriented Chemical Engineering, Dalian University of Technology, Dalian 116024, China

[7]Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen 518132, China

*Correspondence: ligang@szbl.ac.cn

ABSTRACT

Recent advances in chemical proteomics have focused on developing chemical probes reacting with nucleophilic amino acid residues throughout the proteome. Among the nucleophilic amino acids, histidine is an attractive candidate due to its presence in enzyme active sites, metal-binding sites, and protein-protein interaction interfaces. However, histidine has moderate nucleophilicity, and its modification is easily influenced by cysteine and lysine, resulting in poor selectivity and narrow proteome coverage. Here, we report a singlet oxygen and chemical probe relay labeling method that achieves exquisite selectivity towards histidine. A small molecular photosensitizer library is screened together with a chemical probe library to optimize the histidine labeling, enabling histidine profiling in live cells with over 7500 unique sites. We applied this method to discover unannotated histidine sites for metalloproteins and characterize the key histidine residues H309 for metabolic enzyme IDH1 and H73 for CRIP1. Finally, we used this method to probe the accessibility of histidine residues in mitophagy, revealing H138 of PARK7 as the key residue influencing the protein subcellular localization and stability. These findings demonstrate the applicability of our method in discovering functional histidine sites in the human proteome.

INTRODUCTION

Chemical proteomics, particularly activity-based protein profiling (ABPP) pioneered by the Cravatt group[1, 2], has proven to be a powerful tool for annotating uncharacterized protein function and expanding the pool of druggable targets. Original activity-based probes target residues with heightened nucleophilicity in active sites within enzyme families, such as serine hydrolase[3], kinase[4], cysteine protease[5], metalloprotease[6], glycosidase[7] and tyrosine phosphatases[8]. These probes have helped uncover novel enzyme roles in human physiology and diseases and have

led to the development of various enzyme inhibitors[9, 10]. To enable a broader exploration of the proteome, recent advances in chemical proteomics leverage new biocompatible chemistries to measure the proteome-wide reactivity of nucleophilic residues, not limited to a specific enzyme family. For instance, thiol-alkylating agents and amine-reactive esters have been developed for cysteine and lysine[11, 12], along with other modalities for tyrosine[13], methionine[14] and aspartic acid/glutamic acid[15, 16]. However, investigation of less nucleophilic residue is still in its infancy.

Among the less-explored residues, histidine represents an attractive candidate. It is commonly found in enzyme active sites, metal-binding sites, and protein-protein interaction interfaces[17]. Specifically, its imidazole side chain serves as both a hydrogen donor and acceptor, making it a useful catalytic component for proton shuttling[18]. As a metal coordination ligand, histidine is often present in the active sites of metalloproteins[19], which account for one third of the proteome[20]. Histidine also mediates pH-dependent protein-protein interaction by regulating electrostatic interactions[21]. In addition, histidine phosphorylation plays a regulatory role in tumorigenesis[22], ion channel activity[23] in eukaryotes, as well as bacterial signal transduction exemplified by numerous two-component systems[24].

Several strategies have been developed for selectively labeling histidine. The first strategy involves direct covalent conjugation of the imidazole side chain (N3) via electrophilic functionalization, such as thiophosphorylation[25], epoxide ring opening[26], acrolein Michael addition[27] and activated thioacetal bioconjugation[28]. The second strategy is based on visible-light promoted Minisci type carbon(C2)-hydrogen activation, which requires a low pH and high organic solvent content[29]. The third method is an umpolung strategy where imidazole is oxidized by singlet oxygen generated by a ruthenium-based photocatalyst, followed by labeling with a nucleophilic probe 1-methyl-4-arylurazole[30]. Despite these efforts, the harsh conditions or instability of

reagents/products have limited the cellular applications of these strategies. Even though acrolein Michael addition was recently employed in proteome-wide histidine labeling, it still requires prior blockage of reactive cysteine residues with a high concentration of *N*-ethylmaleimide, consequently limiting its use in cell lysate[27]. Taken together, a biocompatible and selective histidine labeling method with deep proteome coverage in live cells is ideal but yet to be reported.

Here, we reported a novel method relying on singlet oxygen-driven histidine oxidation and chemical probe relay labeling system to achieve complete selectivity for histidine residues in live cells. Through our screening of 17 different photosensitizers and 16 chemical probes, we observed proteome-wide coverage of histidine with over 7500 unique sites in more than 2500 proteins in HeLa cells. By incorporating histidine labeling into quantitative chemical proteomics, we unraveled the functional significance of H309 of isocitrate dehydrogenase 1 (IDH1) and H73 of cysteine-rich protein 1(CRIP1). Furthermore, we investigated the change in accessibility of solvent-exposed histidine residues using a cellular mitophagy model and discovered the functional importance of H138 in regulating subcellular localization and stability of Parkinson's disease protein 7 (PARK7). Our results illustrate the potential of our method in global profiling of functionally unannotated histidine residues at the molecular level.

RESULTS

Photosensitizer and chemical probe screening. An ideal histidine profiling platform should show three key features: (1) high selectivity towards histidine, (2) deep proteome-wide coverage, and (3) applicability to both cell lysates and live cells. We recently reported a singlet oxygen-dependent proximity labeling approach using protein miniSOG as the photosensitizer and showed exquisite selectivity towards histidine[31]. Inspired by this discovery, we envisaged using small

molecule-based photosensitizers instead of a recombinant protein may improve singlet oxygen yield and result in deeper histidine coverage[32, 33]. Additionally, this small molecule treatment would allow analysis of native cells or tissues without the need for genetic manipulation (Fig. 1a). Mechanistically, oxidation of histidine with singlet oxygen produces intermediate 2-oxo-histidine, which undergoes nucleophilic attack by an alkyne-containing chemical probe. Further modification through copper-catalyzed azide–alkyne cycloaddition (CuAAC), followed by streptavidin enrichment, trypsin digestion, photo-cleavage and LC-MS/MS characterization would reveal reactive histidine sites in proteome[31].

The proteomic coverage likely depends on the efficiency of singlet oxygen generation by the photosensitizer as well as the reactivity of the chemical probe. Thus, we first screened a library of small molecule photosensitizers (Fig. S1A for structures), including commonly used histologic stain dyes Eosin B (**EB**) and Eosin Y (**EY**)[34], naturally occurring pigments Hyprocrellin A and B (**HA, HB**) that are commonly used as photosensitizers in photodynamic therapy[35], endogenous photosensitizer riboflavin (**RF**) [36], traditional photodynamic therapy drugs Rose Bengal (**RB**) and Methylene blue (**MB**)[37], heavy atom-enhanced photosensitizers dibromofluorescein (**DBF**)[38, 39], hemicyanine dye Icy-OH (**IO**)[40], thio-pentamethine cyanine photoinducers TCy5-CHO (**T5C**), TCy5-Btz (**T5B**) and TCy5-Ph-3F (**T5P**)[41], chlorin derivatives methyl pyropheophorbide-a (**MP**) and chlorin e6 trimethyl ester (**CE**)[42], as well as aggregation-induced emission luminogens TTPy-alkyne (**TA**), TTPy-OH (**TO**) and DPA-SCPI (**DS**)[43, 44].

The cytotoxicity of the photosensitizer molecules under the light illumination were evaluated, and no significant cytotoxicity was observed under the short period labeling conditions (Fig. S2). Cells were treated with each photosensitizer, respectively *in situ*, which was excited at their maximum absorbance wavelengths. The generated singlet oxygen oxidized histidine residues to 2-

oxo-histidine, which then reacted with our previous chemical probe 3-ethynylaniline (**3-EA**) prior to click reaction with rhodamine azide. In-gel fluorescence analysis revealed proteome-wide labeling in HeLa cells for most of the photosensitizers (Fig. 1B), and we confirmed that the labeling depended on both light irradiation and the chemical probe (Fig. S3).

In terms of chemical probe, the balance between nucleophilicity and stability towards singlet oxygen should take into account. We then screened four types of chemical probes (Fig. S1B for structures): (1) aniline probes with different substitutions[31, 34], including 2-ethynylainline (**2-EA**), **3-EA**, 4-ethynylaniline (**4-EA**), 3-ethynyl-N-methylaniline (**3E-MA**), 5-ethynylpyridin-3-amine (**5EP-3A**), N-(2-aminophenyl)pent-4-ynamide (**NPA**), 3-ethynyl-4-methylaniline (**3E-4MA**), 3-ethynyl-4-fluoroaniline (**3E-4FA**), 5-ethynyl-2-fluoroaniline (**5E-2FA**), 3-ethynylpyrazin-2-amine (**3EP-2A**); (2) alkylamine probes[45], propylamine (**PA**) and 4-ethynylpiperidine (**4-EPD**); (3) phenol probes[46], 2-ethynylphenol (**2-EP**), 3-ethynylphenol (**3-EP**) and 4-ethynylphenol (**4-EP**); as well as (4) a recently reported thioacetal alkyne (**TAA**) probe[28]. These chemical probes contain reactive amine, hydroxyl and thiol groups. In-gel fluorescent analysis showed intense labelling bands with aniline probes in comparison to alkylamine, phenol and thioacetal probes (Fig. 1C). Among the aniline probes, **5E-2FA** and **3-EA** demonstrated the highest labeling intensities, which may possess the optimized balance between nucleophilicity and stability towards singlet oxygen. In comparison, **2-EA** and **4-EA**, isomers of **3-EA**, showed weaker intensities. Substitutions of **3-EA**, including methyl and fluorine at the ortho or para position of amine group, the N-methylation amine, as well as the pyridine surrogate, significantly impacted the labeling outcomes, highlighting the tunability of the labeling by the substituent. On the other hand, the fluorine atom in **5E-2FA** can be used for ^{19}F NMR detection of the modified proteins. Taken together, the optimized chemical probes can significantly enhance the histidine labeling.

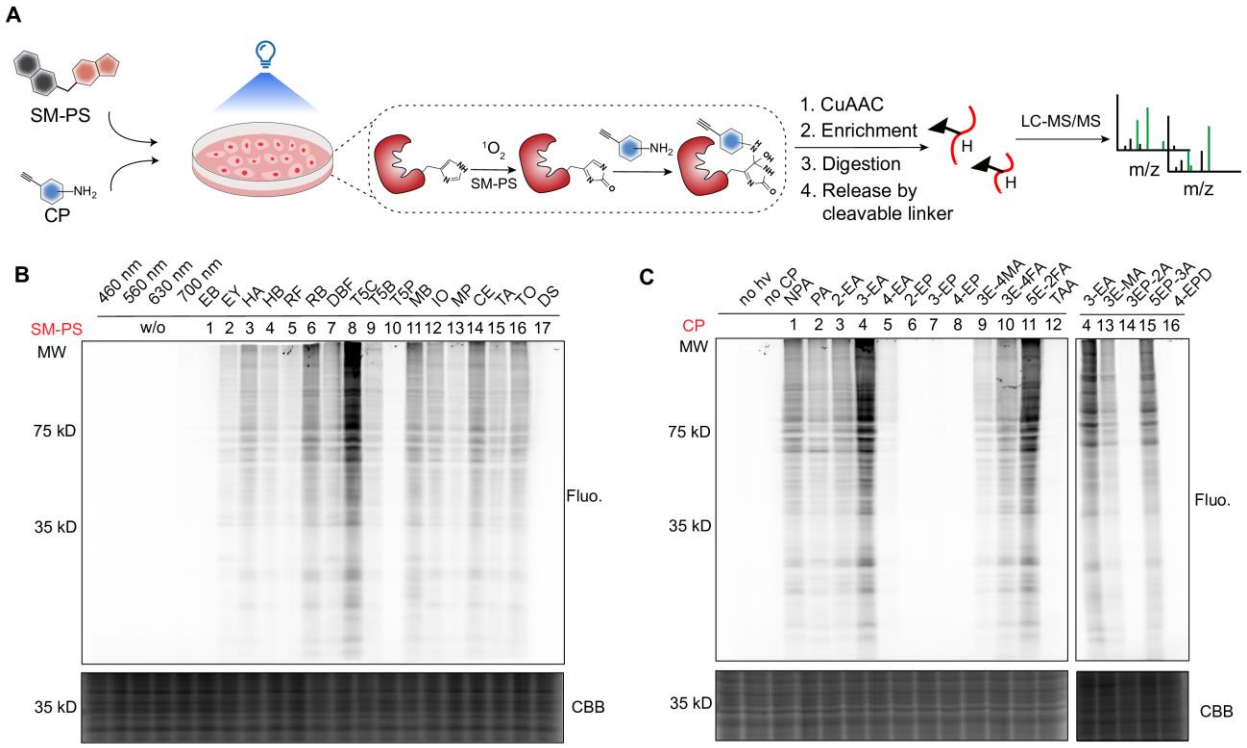

Figure 1. Development of global histidine labeling by a singlet oxygen based chemical proteomic method. **A)** Schematic representation of the workflow for histidine labeling using a small molecule photosensitizer (SM-PS) and an amine-containing chemical probe (CP). The cells were treated with SM-PS and CP, followed by visible light illumination to generate singlet oxygen. The oxidized histidine residues were selectively labeled by CP to form a covalent adduct, which was then enriched by pull-down using click chemistry, followed by trypsin digestion, photo-cleavage and analyzed by LC-MS/MS for quantitative analysis. **B)** In-gel fluorescence screening of 17 SM-PS using 3-ethynylaniline (**3-EA**) as the CP. The photosensitizers were irradiated at their respective absorbance wavelength, and negative control experiments omitting the SM-PS were conducted to evaluate the background labeling. **C)** In gel fluorescence screening of 16 CPs using Rose Bengal (**RB**) as the SM-PS. Negative control experiments omitting irradiation or CP were conducted to evaluate the background labeling. Coomassie brilliant blue (CBB) was used as a protein loading control. Gel imaging were independently repeated twice with similar results.

Histidine profiling in live cells. Next, we sought to assess the scope and selectivity of 17 photosensitizers with the chemical probe **3-EA** in live cells. To achieve this, we used previous

reported chemical proteomics method TOP-ABPP (tandem orthogonal proteolysis activity-based protein profiling) workflow[31, 47]. This workflow involves streptavidin enrichment of the labeled proteins, followed by on-beads trypsin digestion and photo-cleavage to release modified peptides. The modified peptides were then characterized by tandem mass spectrometry (LC-MS/MS) (Fig. 2A).

To determine the modification mass shift and amino acid selectivity, an unbiased open search workflow enabled by the MSfragger-based FragPipe computational platform was used[48], and the mass shifts observed on peptides were summarized using a build-in module called PTM-Shepherd. After removal of common modification artifacts, such as the first isotope peak (delta mass +1) and cysteine oxidation to cysteic acid (delta mass +48), we detected an average of ~4500 peptide spectrum matches (PSMs) for each photosensitizer (Fig. 2B). Among the PSMs, the delta mass +229 Da and its hydrolyzed product +247 Da accounted for 85-90%, along with other redox forms such as +227, +231, +245 Da. This result was consistent with our previous findings using photosensitizer protein miniSOG[31], but we observed more PSMs and redox forms here, likely due to the higher singlet oxygen yield by small molecule photosensitizers.

After determining the mass of the modifications, we performed an offset search using +229 and +247 Da in FragPipe, which allows us to globally search the offset mass at any amino acids or protein termini to evaluate the selectivity. To our delight, we observed 100% specificity for histidine without modification of any other amino acids or termini (Fig. 2C). The high selectivity was likely endowed by aniline probe towards the oxidized intermediate 2-oxo-histidine, as singlet oxygen can oxidize other amino acids including methionine, tryptophan, and tyrosine[49, 50].

We then applied a conventional close search to determine the labeled histidine sites. **RB**, **T5C**, **T5B** yielded over 3000 histidine sites, while seven other photosensitizers yielded over 2000

sites (Fig. 2D). Evaluation of MS2 spectra showed high confidence identification of a large fraction of y ions and b ions (y and b), enabling pinpointing the exact modified histidine residues (Fig. 2E). Only **RF** showed less than 100 sites, likely due to its low singlet oxygen yield. In total, we were able to identify 7586 unique sites in 2544 proteins using a single-shot 100-min LC-MS/MS runs without any fractionation, which represents the deepest histidine coverage in live cells. Unsupervised hierarchical clustering of the histidine sites (excluding **RF**) groups photosensitizers of similar structures together, such as **MP** and **CE**, **HA** and **HB**, **T5C** and **IO**, indicating the high quality of our dataset (Fig. 2F). We then evaluated the reproducibility using **RB** and obtained correlation coefficients around 0.9 in three biological replicates, demonstrating the reliability of our method (Fig. S4). While some photosensitizers were previously reported to localize in specific organelles, such as **IO** and **T5C** in mitochondria, **T5B** in ER, and **MB** in nucleus[40, 41, 51], our gene ontology (GO) analysis showed that the labeled proteins were distributed throughout the cells (Fig. S5). We speculate that the high singlet oxygen yield as well as the sub-organelle localization of the photosensitizer may account for labeling outside of the expected organelles. To confirm the subcellular localization, we performed confocal microscopy analysis of cells underwent histidine labeling followed by click reaction with a Cy3 dye, along with mitochondria and ER trackers (Fig. S6). Consistent with the GO analysis, most photosensitizers showed poor correlation with ER and mitochondria for both 2 and 20 minutes illumination as the labeling was distributed throughout the cells. **MB** labeled proteins predominantly located in the nucleus, likely due to the electrostatic interaction between positively charged **MB** and negatively charged nucleic acids. As expected, **MB** labeled proteome was mainly in the nucleus (~60%), demonstrating the advantages of **MB** in nuclear histidine analysis.

Additionally, we performed side-by-side comparison of the optimized aniline probes **5E-2FA** and **3-EA** in gel imaging experiment using photosensitizer **RB**. The open search showed the delta mass of +247 and +265 Da for **5E-2FA**, along with other redox form +245 and +249 Da, indicating the fluorine substitution was incorporated in the proteome through our labeling method. The closed search showed approximate 3372 and 3427 sites for **3-EA** and **5E-2FA**, respectively (Fig. S7), in consistent with the in-gel fluorescence result.

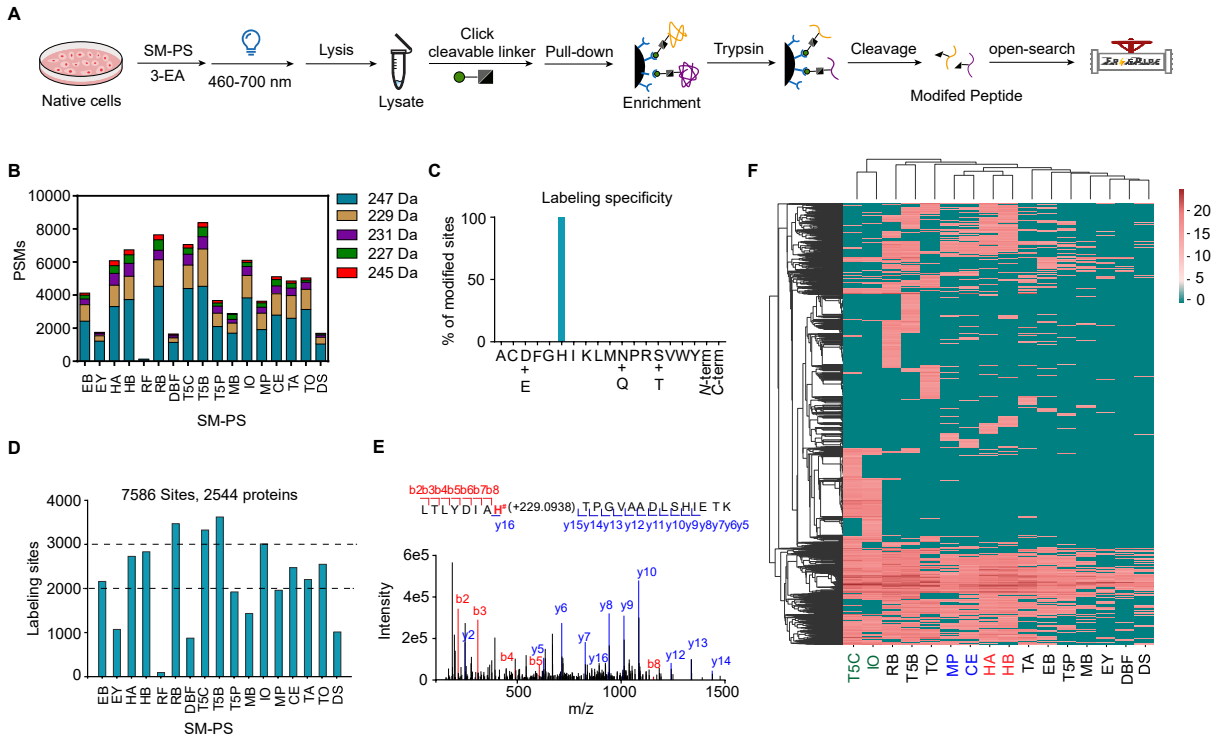

Figure 2. Characterization of histidine sites in live cells. **A)** An unbiased workflow using the MSFragger-based FragPipe computational platform was employed to study residue selectivity. Click chemistry with a cleavable linker was used to modify the histidine sites, followed by photo-cleavage of the modified peptide from the streptavidin resin for open-search to identify the masses of modification and the residue. **B**) The masses of modifications in open search for each photosensitizer were assigned, and the top five unannotated masses were listed. PSMs = peptide spectrum matches. **C**) The results of offset search using +229 and +247 Da as the shift masses were obtained by combining all photosensitizers' dataset. **D**) Closed search for each photosensitizer was performed using +229 and +247 Da as the differential modifications. **E**) An MS2 spectrum annotation of a **3-EA**-modified histidine site is shown. Covalent reaction with **3-EA** adds +229

Da to the modified amino acid as a representative example. **F**) Unsupervised hierarchical clustering of the histidine sites for photosensitizers (excluding **RF**) was performed. These experiments were independently repeated at least twice with similar results. Source data are provided as a Source Data file.

Bioinformatics analysis of histidine functions. We next asked whether the identified histidine sites are functional. Of the total identified proteins, 240 proteins have annotated active sites in UniProt database. We downloaded the protein structures in PDB and measured the distance between identified histidines and active site residues using a python script (Fig. 3A). Our method directly labeled the active histidine residues in eight enzymes reflected by 0 Å distance (Fig. 3B), including H118 of nucleoside diphosphate kinase B (NME2), H187 of malate dehydrogenase (MDH1), H487 of dihydrolipoyl dehydrogenase (DLD), H96 of triosephosphate isomerase (TPI1), H235 of fumarate hydratase (FH), H193 of L-lactate dehydrogenase (LDHA), H347 of citrate synthase (CS), and H178 of ATP-dependent clp protease (CLPP). Mechanistically, the active site labeling will abolish the enzymatic activities. Additionally, 32 and 86 labeled histidine sites are within 5 Å and 10 Å, respectively, to the annotated protein active site residues (Fig. 3B). These histidine sites will possess significant influence on the enzyme functions. Combined, the direct or adjacent labeling of the active sites implies the applicability of our method for inhibitor discovery when applied in a competitive manner, similar to other ABPP strategies reported to other nucleophilic amino acids[9].

To obtain more structural features of the labeled histidine residues, we conducted a solvent-accessible surface area (SASA) analysis[52] and found that ~85% of labeled sites were solvent exposed with the remaining ~15% of sites located within the hydrophobic environment as determined by relative solvent accessibility (RSA) value (Fig. 3C). This result indicates that the labeled histidine residues are predominantly located on the protein surface. Thus, our method could

profile the difference in histidine accessibility mediated by protein-protein interaction, subcellular translocation, or protein conformational change.

To conduct a deeper analysis of the function of the labeled histidine sites, we aligned the identified histidines with the protein domains annotated by PROSITE to determine the enriched domains[13]. We found several highly enriched domains among the photosensitizer-labeled domains ($Q < 0.01$) from the whole dataset, including proteins involved in RNA recognition (RNA recognition motif (RRM) domain; K homology (KH) domain), small GTPase protein family (Rab, Ras and Rho domain), guanine nucleotide-binding domain (translational-type guanine nucleotide-binding domain: G_TR_2; septin-type guanine nucleotide-binding domain: G_SEPTIN), metal binding domain (zinc binding: DnaJ_2; calcium-binding: Annexin_2), disulfide bond functional domain (coiled coil-helix-coiled coil-helix (CHCH) domain, thioredoxin_2 domain) (Fig. 3D). In addition, the photosensitizer labeled proteins from live cell profiling were largely absent from the DrugBank database (75%, Fig. 3E). The DrugBank protein group contains mainly enzymes (e.g., aminoacyl-tRNA synthetases), thioredoxin, proteasome, and annexin domain. In contrast, non-DrugBank protein group largely recapitulates the domains in the whole dataset, which are mainly the non-enzyme and structural-related domains (Fig. S8).

We evaluated the locations of the histidine residues in secondary protein structures using the structures generated by Alphafold 2.0 and observed that 35.3% and 26.0% of the sites were located in the α-helix and β-sheet structures[52, 53], respectively, while 38.7% were located in the loop region (Fig. 3F). Analysis of the local sequence context of labeled histidine revealed that no conserved motif was observed in our method (Fig. 3G), which further indicates that our method could globally profile histidine labeling.

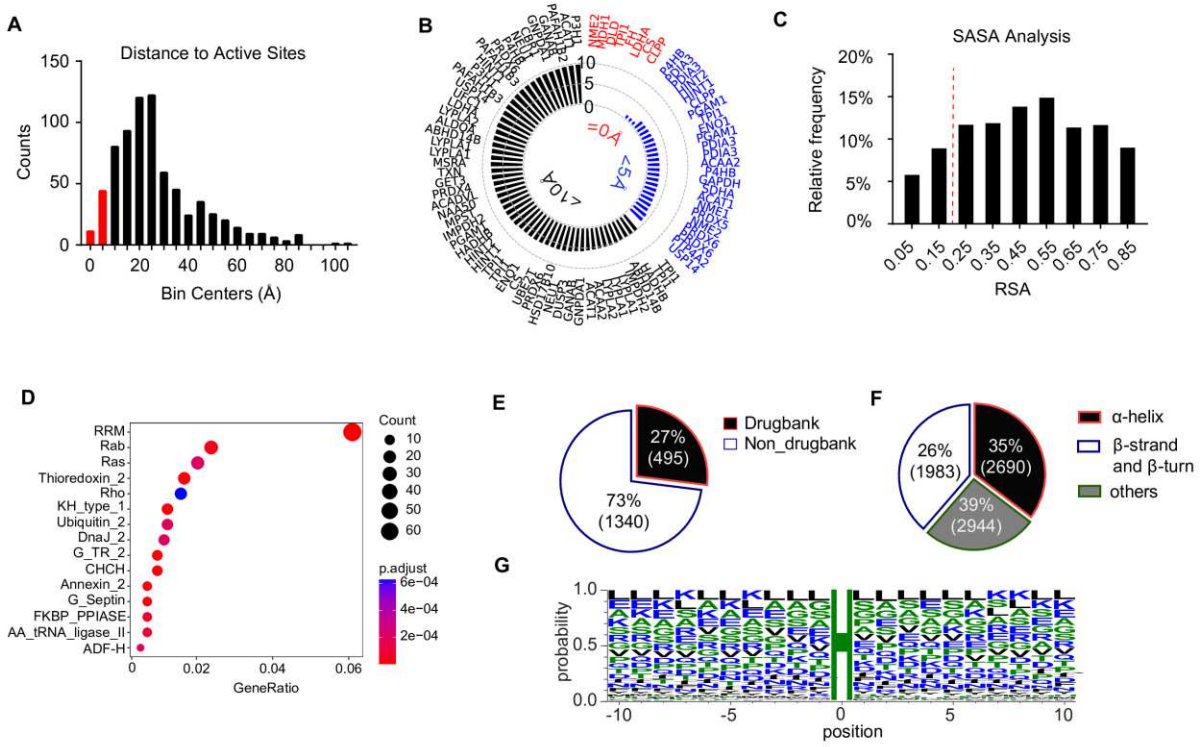

Figure 3. Bioinformatics analysis of the functions of labeled histidines. **A**) Histogram plot showing the distribution of distances between labeled histidines and active sites. **B**) Circular bar plot showing the distances equal to 0 Å and within 5 Å or 10 Å. **C**) Relative solvent accessibility (RSA) analysis of labeled histidines, with a threshold of 20% labeled. **D**) Enriched domains annotations having a Q < 0.01 after Benjamini–Hochberg correction of a two-sided binomial test. **E**) Comparison of photosensitizer-labeled proteins with DrugBank proteins, with the non-DrugBank group consisting of proteins that did not match a DrugBank entry. **F**) Percentage of histidine residues in different types of secondary structures. **G**) Weblogo analysis of the local sequence context of modified histidines. Source data are provided as a Source Data file.

Discovery of novel histidine functions in metalloproteins. Considering the important role of histidine in metal-binding sites, we investigated whether our methodology could discover uncharacterized functional histidine sites in metalloproteins. Our rationale was that histidine binding to a metal reduces electron density on the imidazole ring, thus decreasing its reactivity to singlet oxygen and labeling by aniline probe. To test this hypothesis, we treated a model protein,

small ubiquitin-like modifier (SUMO) with an N-terminal 6xHis tag, with or without nickel sulfate and performed the histidine labeling workflow. The ion intensity of N-terminal peptide with the 6xHis tag decreased by 74-fold after nickel treatment, and the corresponding spectral count decreased from 20 to 2 (Fig. 4A). These results validate the hypothesis and demonstrate the potential of the method to discover critical metal-binding histidine sites.

To quantitatively identify differential histidine labeling, human cell proteomes derived from isotopically light and heavy amino acid-labeled HeLa cells (i.e., stable isotopy labeling with amino acids in cell culture (SILAC)) were treated with or without the metal chelator ethylenediaminetetraacetic acid (EDTA) (Fig. S9A). To ensure the data reliability, forward ('heavy': + EDTA; 'light': - EDTA) and reverse ('heavy': - EDTA; 'light': + EDTA) experiments were performed in parallel. Proteins with SILAC ratio heavy/light (H/L) > 2 in the forward experiment and L/H > 2 in the reverse experiment were designated as the metal-binding dependent histidine sites (Fig. 4B). Ribosomal proteins, such as RPS19 and RPS21 were identified as more reactive after EDTA treatment, in line with the domain analysis showing high enrichment of histidine residues in RNA recognition motifs. Upon close inspection of the ribosomal proteins, we found that both monovalent cations and magnesium ions are essential for ribosomal complex[54]. Additionally, we identified isocitrate dehydrogenase (IDH1), a cancer-related metabolic enzyme that converts isocitrate to α-ketoglutarate[55], and cysteine-rich protein 1 (CRIP1), a member of the LIM/double zinc finger protein family[56], with significantly changed histidine labeling (Fig. 4B). To validate the proteomic result, we performed western blot analysis of the samples treated with or without EDTA, which showed consistent results to LC-MS/MS analysis (Fig. 4C).

To uncover novel functions of the identified histidine sites, we examined the crystal structures of the proteins. While in the available PDB structure H309 in IDH1 is about 9 Å from

the calcium ion (Fig. 4D), our SILAC result showed 3.6-fold difference in samples treated with or without EDTA, indicating the involvement of this residue in metal binding. In addition, the co-evolution-based and machine learning-enable pipeline "MetalNet"[57] predicts that H309 co-evolutes with D375 and showed the metal binding probability 0.67, in consistent with our result. To further investigate the importance of this residue in enzyme function, we performed site directed mutagenesis (H309M, H309A, H309G, and H309R) and assessed IDH activity using a commercial assay kit in both cell lysate and purified protein. The mutations almost completely abolished IDH1 activity in both cell lysate and purified protein, indicating the key role of H309 in enzyme function (Fig. 4E).

Regarding CRIP1, there is only a NMR structure of the mouse homolog[56]. CRIP1 consists of 76 amino acid residues and contains both N-terminal (C3, C6, H24, C27) and C-terminal (C30, C33, C51, C55) LIM domains as well as a C-terminal disordered region (P62 to K76). Although the identified H73 was located in the C terminal disordered region, we speculated its involvement in Zn^{2+} binding as 3.8-fold labeling increase after EDTA treatment. During the protein purification process, enterokinase was used to remove the 6xHis tag. In the absence of the 6xHis tag, the wild-type protein had a stronger binding affinity to the nickel column than the H73A mutant (Fig. 4F). This observation indicates the involvement of H73 in metal binding. To further investigate this, inductively coupled plasma mass spectrometry (ICP-MS) was employed to determine the zinc binding capacity of the protein. The results showed 30% higher capacity of the wild-type CRIP1 than H73A mutant, indicating H73 has the zinc-binding ability as exemplified by the corresponding simulated structure (Fig. 4G, H). MetalNet failed to provide the information because the MSA (multiple sequence alignment) information for the disordered region is lacking, highlighting our experimental discovery of functional sites. Taken together, these findings

demonstrate the ability of our histidine labeling technique to uncover new functional sites of metalloproteins.

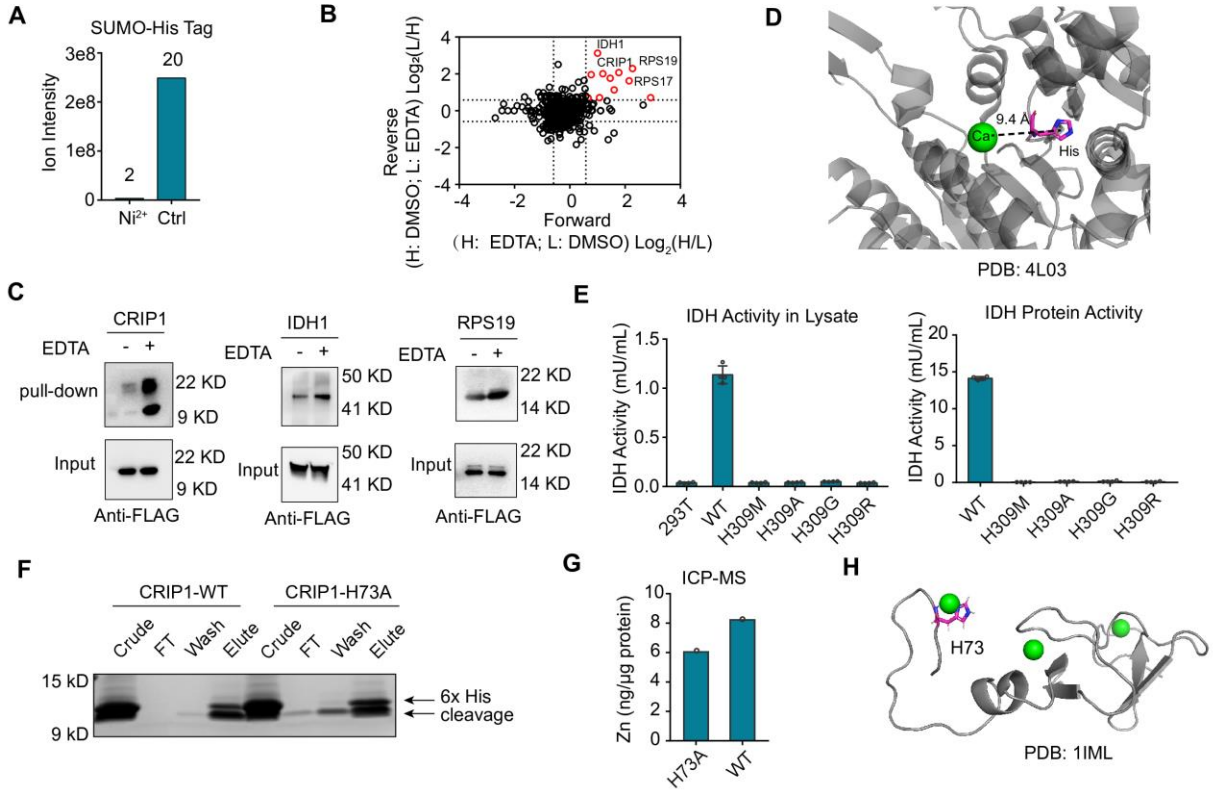

Figure 4. Characterization of unannotated, functional histidines in metalloproteins. **A**) Histidine labeling of 6xHis-tagged SUMO protein before and after treatment with nickel sulfate, with quantification of ion intensity and spectrum counts of the labeling peptide containing 6x His. **B**) Forward and reverse SILAC experiments were used to discover the metal-binding dependent histidine sites. **C**) Histidine labeling and subsequent Western blot analysis of Flag-tagged CRIP1, IDH1, and RSP19 in the presence or absence of EDTA treatment. **D**) Protein structure of IDH1 highlighting the distance between H309 with the calcium ion (PDB: 4L03). **E**) Measurement of IDH activity in cell lysate and purified protein of IDH1 WT, H309M, H309A, H309G, H309R. **F**) Nickel column purification of wild-type and H73A CRIP1 after cleavage of the 6xHis tag by enterokinase. Crude: protein mixture after cleavage; FT: flowthrough after nickel column binding; wash: 50 mM imidazole solution wash; elute: 200 mM imidazole solution to elute the binding proteins. **G**) ICP-MS measurement the zinc in WT and H73A CRIP1. **H**) Proposed structure of CRIP1 with

zinc binding at H73. These experiments were independently repeated at least twice with similar results. Source data are provided as a Source Data file.

Discovery of functional histidine in mitophagy. Given the labeled histidine residues are mainly located on the protein surface as determined by SASA analysis and those solvent-exposed histidine residues may undergo accessibility change in response to various biological processes such as protein-protein interaction, subcellular translocation, or protein conformational change, interfering histidine oxidation by singlet oxygen and/or subsequent labeling. Therefore, our method could potentially capture the accessibility change for histidine. To test this hypothesis, we again treated the N-terminally 6xHis tagged SUMO with or without anti-His monoclonal antibody to construct an artificial protein-protein interaction. Indeed, the ion intensity of the peptides containing 6xHis tag after antibody binding decreased by 6.5-fold (Fig. 5A), validating the hypothesis and demonstrating the potential of the method to characterize the functional histidine sites.

We then used mitophagy cell model induced by the pharmacological agent, carbonyl cyanide *m*-chlorophenylhydrazone (CCCP), where depolarized mitochondria are removed from cells through the autophagy pathway involving a large number of new protein-protein interactions, organelle interactions, and protein translocations. To quantitatively identify differential histidine labeling during mitophagy, we treated SILAC cells with CCCP or vehicle, followed by our histidine workflow (Fig. S9B). To ensure data reliability, forward ('heavy': + CCCP; 'light': + vehicle) and reverse ('heavy': + vehicle; 'light': + CCCP) experiments were performed in parallel. To increase the stringency, proteins with SILAC ratio heavy/light (H/L) > 4 in the forward experiment and L/H > 4 in the reverse experiment were designated as significantly changed sites in mitophagy. Surprisingly, we exclusively identified 18 proteins with decreased histidine labeling after CCCP treatment, while no increased labeling was detected. The protein-protein interaction

enrichment analysis of these proteins, powered by Metascape[58], revealed that cristae formation, mitochondrial biogenesis and respiratory electron transport were the top three enriched pathways or processes (Fig. S10A), in accordance with the CCCP-induced mitophagy. Of the 18 proteins, six proteins including PARK7, TRAP1, HSP70, PRDX3, ATP5F1B and ATP5F1D were located in the mitochondria (Fig. 5B). To validate that the decreased labeling was due to histidine accessibility difference rather than protein abundance, we ruled out the protein abundance change using western blots and SILAC bulk proteomics for PARK7, PRDX3 and HSP70 after 10-min CCCP treatment (Fig. S10B, C). Consistent with the proteomic results, histidine labeling pull-down and subsequent western blot analysis showed decreased labeling after CCCP treatment (Fig. 5C), indicating that accessibility change prevents histidine labeling.

Among the identified histidine sites, PARK7, also called DJ1, is a protein that plays a role in protecting cells from oxidative stress and acts as an essential downstream mediator in PINK1/parkin-dependent mitophagy[59]. Mutations in PARK7 has been reported to disrupt the mitophagy, leading to the accumulation of damaged mitochondria and the eventual death of dopaminergic neurons[60]. In our study, we identified only one histidine site, H138, in PARK7 with decreased labeling in mitophagy. To investigate the function of H138, we performed FLAG-tag co-immunoprecipitation experiments comparing the difference in interacting partners between the wild-type (WT) and H138A mutant. SILAC-based quantitative proteomics with forward ('heavy': WT; 'light': H138A) and reverse ('heavy': H138A; 'light': WT) experiments were conducted in parallel (Fig. 5D). We found the H138A mutation increased the interaction with chaperon proteins, including GRP75 (HSPA9, mitochondrial HSP70), HSPA8 (HSP70 protein 8) and HSPA1A (HSP70 protein 1A). Previous studies have reported that the translocation of DJ1 to mitochondria requires the assistance of the molecular chaperones HSP70 and GRP75[61], indicating that H138A

mutation has increased mitochondrial location. In CCCP-induced mitophagy, our proteomic results showed similar results where H138A had increased interaction with GRP75, HSPA8 and HSPA1A (Fig. 5E). These results indicate that H138A mutation of PARK7 leads to more mitochondrial localization. To confirm the mitochondria localization, we extracted the mitochondria and cytosolic fraction and detected PARK7 abundance by western blot. In line with the proteomic result, H138A mutation was more abundant in mitochondria than WT (Fig. 5F and Fig. S10D). Taken together, our findings discovered a novel function of PARK7 H138A for subcellular translocation (Fig. 5G).

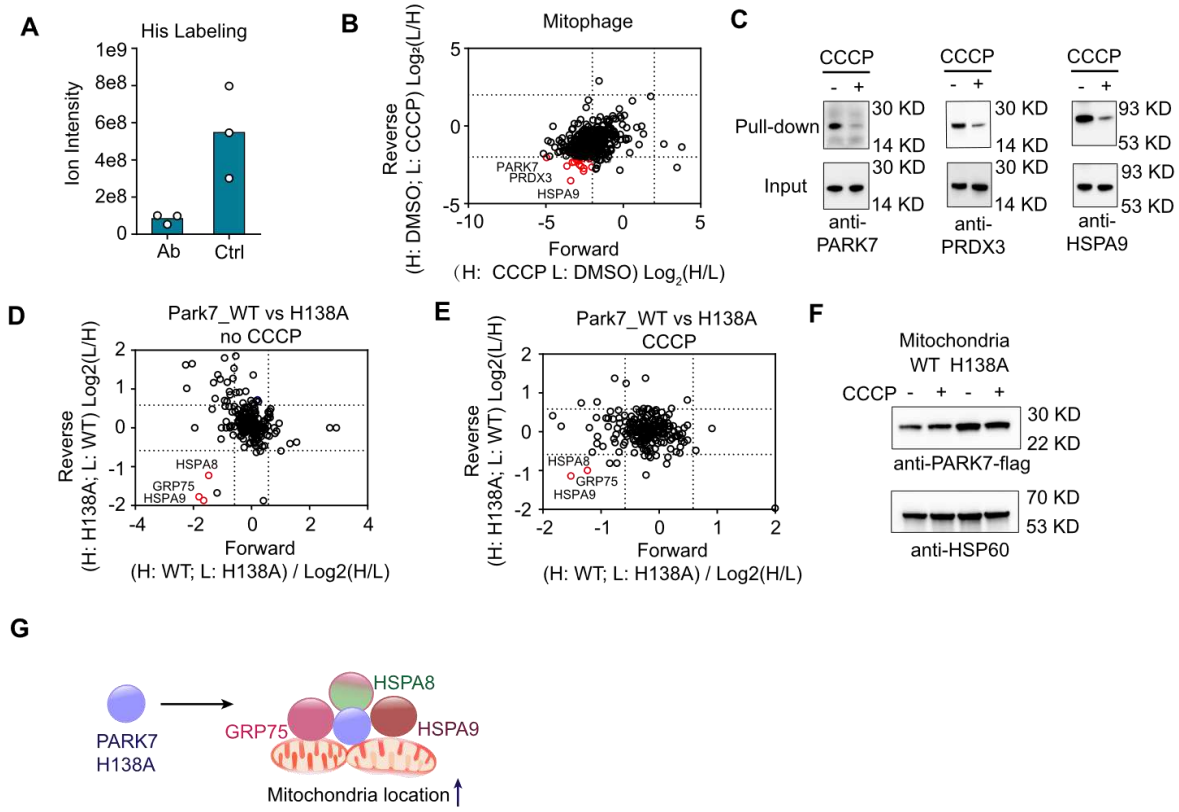

Figure 5. Discovery of functional histidine in mitophagy. **A**) Histidine labeling of 6xHis-tagged SUMO protein before and after binding with anti-His monoclonal antibody, with quantification of ion intensity of the labeling peptide containing 6xHis. **B**) Forward and reverse SILAC experiments were performed to identify differentially labeled histidines in mitophagy. **C**) Verification of protein abundance change and histidine labeling by Western blot for PARK7, PRDX3, and HSPA9 in the presence or absence of CCCP.

D-E) SILAC-based quantitative proteomics were used to investigate the differential interacting partners of PARK7 WT and H138A using FLAG-tag co-immunoprecipitation without CCCP (**D**) and with CCCP (**E**). **F**) Western blot analysis of PARK7 WT and H138A in mitochondria fraction. Mitochondria specific protein HSP60 was used as a loading control. **G**) Working model of PARK7 H138A enabled translocation to mitochondria. These experiments were independently repeated at least twice with similar results. Source data are provided as a Source Data file.

DISCUSSION

We have described a novel approach for global histidine profiling in live cells and whole proteomes, using small molecule-based photosensitizers and aniline relay reactions. This method builds upon the previous advances in photosensitizers[62], bridging photodynamic therapy with chemical proteomics for histidine labeling. While several methods have been developed for histidine bioconjugation, the harsh conditions as well as the weak nucleophilicity and electrophilicity of histidine have hindered previous efforts for global profiling, especially in live cells. Recently, an acrolein based method was developed to map the histidine reactivity[27], but this method needs pre-blockage of the cysteine residues and thus can only be conducted with cell lysates. Our method, utilizing the redox reactivity of imidazole with singlet oxygen through a Diels-Alder type reaction, enables 100% selectivity towards histidine with deep proteome coverage in live cells. We were able to identify over 3200 histidine sites in a single run without any fractionation, using commercially available photosensitizers Rose Bengal or home-made **T5C / T5B**. In total, we have identified over 7500 unique sites in over 2500 proteins.

Our functional analysis of histidine sites revealed high enrichment in RNA-recognition motifs, small GTPase protein family and guanine nucleotide-binding domain, which may attribute to the interaction between histidine and negatively charged nucleotide. Moreover, metal binding

domains including zinc and calcium were also enriched, which is consistent with our discovery in the LIM-containing protein CRIP1. Our method also proved useful in discovering key metal binding residues in proteins, where coordination with metal decreases the labeling, providing an opportunity to discover novel functional sites. Biochemical analysis confirmed that H309 mutation in IDH1 completely abolished its enzyme activities. H73 is a newly identified metal binding site in CRIP1 as determined by ICP-MS, which is occluded in its crystal structure. These two examples showcase the potential of our methods in discovering uncharacterized metalloprotein functional sites. Our work paves the way for further studies on high level metal-demanding or tightly metal controlled species such as glia cells or bacteria.

In addition, we applied histidine labeling in profiling histidine accessibility change caused by protein-protein interaction, translocation, conformational change. In CCCP-induced mitophagy, six mitochondria proteins with decreased histidine labeling were identified, while the protein abundances remained unchanged. PARK7, a regulator of mitophagy, caught our attention to further study the function of the identified site H138. Co-immunoprecipitation coupled with LC-MS/MS revealed that the H138A mutation increased the interaction with mitochondrial chaperon proteins to promote its translocation to mitochondria, validating the novel function of H138 of PARK7. This result showed the general applicability of our histidine profiling in discovering new functional histidines in other cellular models or disease processes.

In summary, our small molecule-based photosensitizer and aniline relay reaction offer a powerful method for global histidine profiling and functional analysis in live cells and whole proteomes, opening up new opportunities for discovering unexplored functions of histidine and other metalloproteins.

Methods

Cell culture. HEK293T and HeLa cell lines were obtained from ATCC (CRL-3216 and CRM-CCL-2). These cell lines had been tested negative for mycoplasma contamination and cultured in DMEM (Thermo, #C11995500BT) supplemented with 10% fetal bovine serum (FBS, Vistech, #SE100-B) and 1% Penicillin/Streptomycin (Hyclone, #SV30010).

Chemical probes. The following chemical probes were purchased from commercial sources, including 2-ethynylainline (**2-EA,** Bidepharm, #52670-38-9), 3-ethynylainline (**3-EA,** Bidepharm, #54060-30-9**)**, 4-ethynylaniline (**4-EA,** Bidepharm, #197844-23-8), Propyl amine (**PA,** Energy-chemicals, #107-10-8), 2-Ethynylphenol (**2-EP**, Bidepharm, #5101-44-0), 3-Ethynylphenol (**3-EP,** Bidepharm, #10401-11-3), 4-Ethynylphenol (**4-EP**, Bidepharm, #2200-91-1), 3-Ethynyl-4-methylaniline (**3E-4MA,** Leyan, #77123-60-5) , 3-Ethynyl-4-fluoroaniline (**3E-4FA,** Leyan, #134690-40-7), 5-Ethynyl-2-fluoroaniline (**5E-2FA**, Bidepharm, #1010422-58-8), 3-Ethynyl-N-methylaniline (**3E-MA**, AMETEK Scientific, #132056-23-6),_3-Ethynylpyrazin-2-amine (**3EP-2A**, Bidepharm, #1005349-13-2), 5-Ethynylpyridin-3-amine (**5EP-3A**, Bidepharm, # 667932-40-3), and 4-Ethynylpiperidine hydrochloride (**4-EPD**, Bidepharm, # 287192-97-6). N-(2-aminophenyl)pent-4-ynamide (**NPA**) was synthesized according to a published procedure[63]. Thioacetal alkyne (**TAA**) probe was a kind gift from Dr. Zigang Li.

Small molecule photosensitizers. The following photosensitizers were purchased from commercial sources, including Eosin B (**EB**, Aladdin, #548-24-3), Eosin Y (**EY**, Aladdin, #15086-94-9), Hypocrellin A (**HA**, MCE, #77029-83-5), Hypocrellin B (**HB**, MCE, #123940-54-5), 4',5'-Dibromofluorescein (**DBF**, Aladdin, #596-03-2), Rose Bengal (**RB**, Sigma-Aldrich, #632-69-9), Methylene blue (**MB**, Sigma-Aldrich, #66720), Methyl pyropheophorbide-a (**MP**, MCE, #6453-67-4), Chlorin e6 trimethyl ester (**CE**, MCE, #35038-32-5), Riboflavin (**RF**, Sigma-Aldrich, #83-88-5), ICy-OH (**IO**, MCE, HY-150970). TCy5-

CHO (**T5C**), TCy5-Btz (**T5B**) and TCy5-Ph-3F (**T5P**) were kind gifts from Dr. Wen Sun group and TTPy-alkyne (**TA**), TTPy-OH (**TO**) and DPA-SCPI (**DS**) were kind gifts from Dr. Dong Wang group.

Plasmid construction. The genetic constructs used in this study were listed in Supplementary Table 1. Human *IDH1* (NM_005896.4), *CRIP1* (NM_001311.5), *RPS19* (NM_001321483.2) DNA amplified via PCR (NEB, #M0491L) from HEK293T derived cDNA library. These DNA fragments were cloned into lentiviral vector for over-expression in mammalian cells. IDH1 point mutation plasmids were constructed by Gibson assembly (Beyotime, #D7010S). Human *PARK7* (NM_007262.5) and *PARK7(H138A)* DNA were synthesized by General Biol and cloned into pLX304 vector. For bacterial expression, CRIP1 and CRIP1(H37A) were cloned into a pET28a vector with 6x His tag at the N-terminus. Enterokinase cleaved site was in the middle of 6x His tag and CRIP1or CRIP1(H37A).

Stable cell line generation. HEK293T cells were seeded in 6-well plate at 2.0 x 10^5 cells per well. After 24 h, recombinant lentiviral plasmid (2.4 μg pLX304) and virus packaging plasmids (1.5 μg psPAX2 and 1.2 μg pMD2.G) were transfected using Lipo8000 (Beyotime, #C0533) at ~80% confluency. Following overnight transfection, media was exchanged and allowed to incubate for an additional 24 h. Viral collection was performed at 24, 48, and 72 h. Viral media was filtered with a 0.45 μm filter (Merck, #millex-GP) and Polybrene (Solarbio, #H8761) was added to a concentration of 5 μg/mL before infection of target cell lines. After 24 h, cells were allowed to recover by exchanging the media. Cells were selected with Blasticidin (Solarbio, #3513-03-9) at 5 μg/mL for the first three passages as a lower stringency selection. Then, 20 μg/mL was employed as a higher stringency for the following three passages.

Histidine labeling in living cells for gel analysis. Hela cells were seeded in a 6-cm dish. When reached

~90% confluency, cells were washed once with PBS, followed by incubation with 2 mM **3-EA** probe and 10 μM photosensitizers in 3 ml Hanks Balanced Salt Solution HBSS (Gibco, #14025092) at 37 °C for 1 h. Cells were then illuminated with a 10 W LED for 10 min at room temperature. The maximum absorbance wavelengths for each photosensitizer were applied (**MP**, **CE**: 700 nm; **T5C**, **T5B**, **T5P**, **MB**, **IO**: 630 nm; **RB**, **HA**, **HB**, **DBF**: 560 nm; **EB**, **EY**, **RF**, **TO**, **TA**, **DS**: 460 nm). After illumination, cells were washed with cold PBS twice, scraped, and resuspended in 200 μL ice-cold PBS buffer containing EDTA-free protease inhibitor (MCE, # HY-K0011). The cells were lysed in by tip sonication for 1min (1 s on and 2 s off, 35% amplitude). The resulting mixture was centrifuged at 15,871×g for 10min at 4 °C to remove the debris and the concentration of the supernatant was adjusted to 1 mg/mL using a BCA protein assay kit (Beyotime, #P0009). 50 μL of the above lysate was incubated with 0.1 mM rhodamine-azide (Aladdin, #T131368), 1 mM TCEP (Sangon, #A600974), 0.1 mM TBTA (Aladdin, #T162437) ligand and 1 mM $CuSO_4$ for 1 h with bottom-up rotation at room temperature. After click reaction, 4x Laemmli buffer was added to the mixture directly and boiled for 10 min at 95 °C. The samples were run on SDS-PAGE long gel and visualized by Bio-rad ChemiDoc MP Touch imaging system with Image Lab Touch Software.

Histidine labeling in live cells for fluorescence imaging analysis. Cells were seeded in 35-mm confocal dish (Biosharp, #BS-15-GJM) at a density of ~$3x10^5$ cells per well. When reached to ~60% confluency after 24 h, cells were washed once with PBS, followed by incubation with 2 mM **3-EA** probe and 10 μM photosensitizers in HBSS buffer at 37 °C for 1 h. Cells were then illumination with a 10 W LED for 10 min at room temperature. The maximum absorbance wavelengths for each photosensitizer were applied as mentioned above. Thereafter, cells were washed with PBS twice and fixed with 4% formaldehyde (Sangon, #E672002) in PBS at room temperature for 15 min. Excess formaldehyde was removed from fixed cells through washing with PBS three times. Cells were then permeabilized with 0.5% Triton X-100 (Sangon, #A600198) in PBS for 30 min and then washed three more times with PBS. Next, 100 μL mixture of click reaction reagents was added to each sample, containing 50 μM cy3-azide (Aladdin, #C196720), 2 mM

CuSO4, 1 mM BTTAA (Confluore, #BDJ-4) and 0.5 mg/ml sodium ascorbate (Aladdin, #S105024) in PBS, and incubated at room temperature for 30 min. After the click reaction, cells were washed with PBS containing 0.1% Tween-20 (Sangon, # A600560) (PBST) six times and then blocked with 5% BSA (Abcone, #B24726) in PBST for 2 h at room temperature.

For confocal analysis, cells were incubated with primary antibodies according to indicated conditions: anti-488-conjugated HSP60 Monoclonal antibody (1:500, Proteintech, #CL488-66041), anti-647-conjugated Calnexin Polyclonal antibody (1:500, Proteintech, #CL647-10427) overnight at 4 ℃. After washed with PBST three times, cells were counterstained with 1 μg/ml DAPI (Thermo, #D1306) in PBS for 15 min at room temperature and washed three times again with PBS. Immunofluorescence images were collected with ZEISS LSM 980 confocal microscope with software ZNE 3.5.

Histidine sites identification by LC-MS/MS. Hela cells were seeded in a 15-cm dish. When reached to ~90% confluency, cells were treated with aniline probe and photosensitizers as mentioned above. After illumination, cells were washed twice with cold PBS, scraped, and resuspended in an ice-code PBS buffer containing EDTA-free protease inhibitor. The cells were lysed by tip sonication for 1 min (1s on and 2s off, 35% amplitude). The resulting mixture was centrifuged at 15,871×g for 10min at 4 °C to remove the debris and the concentration of the supernatant was adjusted to 4 mg/mL using a BCA protein assay kit (Beyotime, #P0009). 1 mL of the above lysate was incubated with 0.1 mM photo-cleavable biotin-azide (Confluore, #BBBD-14), 1 mM TCEP (Sangon, #A600974), 0.1 mM (Aladdin, #T162437) TBTA ligand, and 1 mM $CuSO_4$ for 1 h with bottom-up rotation at room temperature. After click reaction, the mixture was added to a pre-mixed solution (MeOH: CHCl3: H_2O = 4 mL: 1 mL: 3 mL) in a 10 mL glass bottle. Samples were mixed and centrifuged at 4500×g for 10 min at room temperature. The bottom and upper layer solution was discarded sequentially, and the pellet was subsequently washed twice with 1 mL methanol followed by centrifuging at 15871×g for 5 min at 4 °C. 1 mL of 8 M urea (Aladdin, #U111902) in 25 mM ammonium bicarbonate (ABC, Aladdin, #A110539) was added to dissolve the pellet. The samples were reduced with

10 mM dithiothreitol (Sangon, #A100281, in 25 mM ABC) at 55 °C for 40 min and then alkylated by adding 15 mM fresh iodoacetamide (Sangon, #A600539) in dark at room temperature for 30 min. Additional 5 mM of dithiothreitol was added to stop the reaction. About 200 μL NeutrAvidin agarose resin beads (Thermo, #29202) for each sample were prepared by washing three times with 1 mL PBS. The above proteome solution was diluted with 5 mL PBS and incubated with pre-washed NeutrAvidin agarose resin beads for 4 h at room temperature. Next, the beads were washed with 5 mL PBS containing 0.2% SDS (Sangon, #A600485) three times, 5 mL PBS containing 1 M urea three times, and 5 mL PBS three times. The beads were then collected by centrifugation and resuspended in 300 μL 25 mM ABC containing 1 M urea, 1 mM $CaCl_2$ (Macklin, #C805228) and 20 ng/μL trypsin (Beijing Life Proteomic, #V5280). Trypsin digestion was performed at 37 °C with rotation overnight. The beads were washed with 1 ml PBS three times, 1 ml distilled water three times. Release of modified peptides by photo (365 nm) cleavage for 60 min with 800 μL 70% MeOH in 24-well plate on ice. The supernatant was collected in 1.5mL low binding tube. Then the beads were washed twice with 200 μL 70% MeOH, and the supernatant was combined. The samples were dried in a speedvac vacuum concentrator.

For identification and quantification of the modified peptides, the samples were redissolved in 0.1% formic acid and 1 μg peptides were analyzed with a QE Plus mass spectrometer equipped with a nano-ESI source with the vendor-provided Tune and Xcalibar 4.3 software. The samples were separated on an in-house packed 150 μm × 20 cm capillary column with 1.9 μm C18 material (ReproSil-pur, #r13.b9.) and connected to an EASY-nLC 1200 UHPLC system (Thermo). Peptides were chromatographically separated by a linear 95 min gradient from 8 to 50% solvent B (A = 0.1% formic acid in water, B = 0.1% formic acid in 80% acetonitrile) and followed by a linear increase to 98% B in 6 min at a flow rate of 300 nL/min. The QE Plus acquired data in a data-dependent manner alternating between full-scan MS and MS2 scans. The spray voltage was set at 2.0 kV and the temperature of ion transfer capillary was 320 °C. The MS spectra (350−2000 m/z) were collected with 120,000-resolution, AGC of 4 × 10^5 and 150 ms maximal injection time. The top ten most abundant multiply charged precursors from each full scan were fragmented by HCD

with 30% normalized collision energy, quadrupole isolation windows of 1.6 m/z, and resolution setting of 30,000. AGC target for tandem mass spectrum of 5 × 10^4 and 150 ms maximal injection time were used. Dynamic exclusion was set to 30 s. Unassigned ions or those with a charge of 1+ and >7+ were rejected for MS/MS.

The raw data were processed using the MSFragger-based FragPipe computational platform. Open search algorithm with precursor mass tolerance -150 to 500 Da were used to determine the mass shift and corresponding amino acids. Then the modifications on histidine with delta mass +229.0964 and +247.1069 Da were used in FragPipe to identify the modified peptides.

EDTA-treated proteome for functional histidine identification. Hela cells were cultured in SILAC DMEM (Thermo Scientific, #88364) with 10% dialyzed FBS (Viva Cell, C04001-050) and 1% Penicillin/Streptomycin. The medium was supplemented with either 'light' unlabeled lysine (K0, Aladdin, #L113006) and arginine (R0, Aladdin, #A118651) or 'heavy' isotope-labeled 13C6,15N2–lysine (K8, Reertech, #CNLM-291-H-1) and 13C6,15N4–arginine (R10, Reertech, CNLM-539-H-1). SILAC cells were first cultured for more than 5 passages to stably incorporate the isotope. Then, heavy and light labeled Hela cells were cultured separately in 15-cm dish. When reached to 90% confluency, cells were washed with cold PBS twice, scraped and resuspended in an ice-cold PBS buffer containing EDTA-free protease inhibitor. The cells were lysed by tip sonication for 1 min (1 s on and 2 s off, 35% amplitude). The resulting mixture was centrifuged at 15,871×g for 10 min at 4 °C to remove the debris and the concentration of the supernatant was adjusted to 4 mg/mL using a BCA protein assay kit. Forward ('heavy': + EDTA; 'light': - EDTA) and reverse ('heavy': - EDTA; 'light': + EDTA) experiments were performed in parallel. For example, 1 mL of the above heavy proteome was incubated with 4 mM EDTA with bottom-up rotation for 1 h at room temperature, while light proteome was left untreated. Then 10 μM **RB** and 2 mM **3-EA** were added to both light and heavy proteome. The samples were illuminated with 560 nm LED for 1 h at room temperature with bottom-up rotation every 10 min. After illumination, heavy and light proteome were 1:1

mixed, precipitated through adding the solution to a pre-mixed solution (MeOH: $CHCl_3$: H_2O = 4 mL: 1 mL: 3 mL) in a 10 mL glass bottle to remove the EDTA. The pellet was dissolved in 1 mL of 2 M urea in 25 mM ammonium bicarbonate solution for click reaction as mentioned above. The resulting solution was added to a pre-mixed solution (MeOH: $CHCl_3$: H_2O = 4 mL: 1 mL: 3 mL) in a 10 mL glass bottle once again to remove the excess small molecules. The following proteomics steps were the same as above. The raw data were processed using the MSFragger-based Fragpipe software and the built-in SILAC workflow. The modifications on histidine with delta mass +229.0964 and +247.1069 Da were applied for the quantification.

IDH1 activity assay. HEK293T cells were seeded in a 10-cm dish. When reached to ~60% confluency, cells were transfected with 12.5 μg wild type and point mutation IDH1 plasmids separately by Lipo8000 Reagent (Beyotime, #C0533). For the activity measurement in HEK293T cell lysate, cells were homogenized in 200 μL of ice-cold IDH assay buffer after 48 h of transfection. The samples were centrifuged at 13,000×g for 10 min to remove insoluble material. Subsequent experimental procedures were followed the instructions of isocitrate dehydrogenase activity assay kit (Sigma-Aldrich, #MAK062). IDH1 activity was determined using isocitrate as the substrate in an enzyme reaction, which results in a colorimetric (450 nm) product proportional to the enzymatic activity present.

For the activity measurement of purified wild type and point mutation IDH1, HEK293T cells were transiently transfected with the corresponding plasmids. After 48 hours, cells were washed with cold PBS twice and lysed in 1 ml Pierce IP Lysis Buffer (Thermo Fisher, #87787) with 1x EDTA-free protease inhibitor for 30 min at 4 ℃. Thereafter, lysates were collected in 1.5 mL centrifuge tube and centrifuged at 15,871×g for 10 min at 4 ℃. The supernatants were collected and incubated with 50 μL Anti-Flag Magnetic Beads (Sigma-Aldrich, #M8823) by bottom-up rotation for 40 min at room temperature. The beads were washed with 1ml PBS three times. 100 μL 0.1 mM 3x flag peptide (GLPBIO, #402750-12-3) in PBS was added and rotated on shaker for 1h at room temperature. The eluted were collected and the concentrations

were measured using the PAGE gel using BSA as the external standard. The enzymatic activities were determined as mentioned above.

CRIP1 protein purification and inductively coupled plasma mass spectroscopy (ICP-MS) Analysis. The plasmids pET-28a-6xHis-EK-CRIP1/CRIP1(H37A) were transformed into competent E. coli BL21 (DE3) cells. Cells were grown in LB medium containing 50 μg/ml of kanamycin antibiotics, and incubated at 37 ℃ and 220×rpm. Protein expression was induced by 1 mM of isopropyl-beta-d-thiogalactopyranoside (IPTG, Sangon, #A600168) when OD600 reached to 0.6. After 4 hours, the cells were harvested by centrifugation (5000×rpm, 10 min) and resuspended in buffer A (20 mM Tris, 150 mM NaCl, 1 mM DTT, pH=7.5) plus 1 mM phenylmethylsulfonyl fluoride (PMSF, ACMEC biochemical, #P35750). The cells were then homogenized by high pressure homogenizer, and the protein was purified via Ni-NTA agarose beads (MCE, #70666), followed by dialysis overnight. To remove the 6x his tag, enterokinase (Beyotime, #P4237) was added to the protein solution for 16 hours at room temperature, and then purified by ÄKTA™ pure. For ICP-MS analysis, 100 μg CRIP1 and CRIP1(H37A) proteins in buffer B (20 mM HEPES, 200 mM NaCl, 1 mM DTT, pH=7. 4) were incubated with 0.2 mM $ZnSO_4$ separately for 1 hour at room temperature. To remove the $ZnSO_4$ in solution, the protein was ultracentrifuged with buffer B three times using 3 kD ultrafiltration tubes (Millipore, #UFC500324). The protein concentrations were determined by BCA assay. Finally, 100 μL protein was transferred to 1.5 mL Eppendorf tube. 200 μL concentrated nitric acid and 20 μL concentrated hydrogen peroxide were added. The sample was incubated at 70-80 °C for 1.5 hours and then 5 mL ddH_2O was added to dilute the sample for ICP-MS analysis (PerkinElmer, NexlON 350X).

Functional histidine site identification in CCCP-regulated autophagy. SILAC labeled heavy (K8, R10) and light (K0, R0) Hela cells were cultured separately in 10-cm dishes. Forward ('heavy': + CCCP; 'light':

+ DMSO) and reverse ('heavy': + DMSO; 'light': + CCCP) experiments were performed in parallel. For example, when heavy cells reach to ~90% confluency, cells are washed with cold PBS twice, followed by incubation with 20 μM protonophore carbonyl cyanide m-chlorophenyl hydrazone CCCP (Solarbio, #C6700), 5 μM photosensitizer **T5C** and 2 mM **3-EA** in HBSS at 37 °C for 1 hour, while light cells were treated with DMSO. Thereafter, cells were illuminated for 10 min, scraped and resuspended in an ice-cold PBS buffer containing EDTA-free protease inhibitor, then lysed by tip sonication for 1 min (1 s on and 2 s off, 35% amplitude). The resulting mixture was centrifuged at 15,871×g for 10 min at 4 °C to remove the debris and the concentration of the supernatant was adjusted to 2 mg/ml. Heavy and light proteome were 1:1 mixed and the resulting solution underwent click reaction with 0.1 mM photo-cleavable biotin azide, 1 mM TCEP, 0.1 mM TBTA ligand and 1 mM $CuSO_4$ for 1 hour with bottom-up rotation at room temperature. The mixture was added to a pre-mixed solution (MeOH: $CHCl_3$: H_2O = 4 mL: 1 mL: 3 mL) in a 10 mL glass bottle. The next steps were the same as above for histidine sites profiling proteomic workflow. The raw data were processed using the MSFragger-based Fragpipe software and the built-in SILAC workflow. The modifications on histidine with delta mass +229.0964 and +247.1069 Da were applied for the quantification.

Pull-down and western blot analysis. Pull-down and western blot analysis were deployed to verify the proteomics results. For EDTA-regulated functional histidine sites in IDH1, CRIP1 and RPS19, their corresponding overexpressing pLenti plasmids were transfected to Hela cells separately for 48 hours. Cells were washed with cold PBS twice and lysed in 1ml PBS buffer containing EDTA-free protease inhibitor by tip sonication for 1 min. The resulting mixture was centrifuged at 15,871×g for 10 min at 4 °C to remove the debris and the concentration of the supernatant was adjusted to 2 mg/ml. 1 mL of the above lysate was incubated with 4 mM EDTA for 1 hour at room temperature, while 1 mL of lysate was left untreated as the control experiment. Then, 10 μM **RB** and 2 mM **3-EA** were added, followed by illumination with 560 nm LED for 1 hour with bottom-up rotation every 10 min. After illumination, the mixture was added to a pre-

mixed solution (MeOH: $CHCl_3$: H_2O = 4 mL: 1 mL: 3 mL) in a 10 mL glass bottle. The pellet was dissolved in 1 mL of 2 M urea in 25 mM ammonium bicarbonate for Click reaction as mentioned above. The mixture was added to a pre-mixed solution (MeOH: $CHCl_3$: H_2O = 4 mL: 1 mL: 3 mL) solution again. 1 mL of 8 M urea in PBS was added to dissolve the pellet and then diluted with 4 mL PBS. The samples were incubated with pre-washed 200 μL NeutrAvidin agarose resin beads for 4 h at room temperature. Next, the beads were washed with 5 mL PBS containing 0.2% SDS three times, 5 mL PBS containing 1 M urea three times, and 5 mL PBS three times. After washing, 50 μL Laemmli buffer was added to the beads and boiling for 5 min at 95 °C. Samples were analyzed by SDS-PAGE and transferred to PVDF membranes (Millipore, #ISEQ00010) using standard western blotting methods. Membranes were blocked in 5% non-fat milk (Sangon, #A600669) in TBS containing 0.1% tween-20 (TBST) and incubated with primary and secondary antibodies sequentially. Primary antibody was used at 1:1000 dilution in 5% non-fat milk in TBST and incubated overnight at 4 °C. Secondary antibodies were used at 1:5000 and incubated for 1 h at room temperature. The membranes were visualized using chemiluminescence by Chemidoc MP imaging system.

As for functional histidine sites of PARK7, PRDX3 and HSPA9 identified in CCCP-induced mitophagy, Hela cells were seeded in 10-cm dishes, and treatment with 20 μM CCCP, 5 μM photosensitizer **T5C** and 2 mM **3-EA** for 1 hour. Thereafter, cells were illuminated for 10 min, scraped and lysed. The following steps were the same as above.

Primary antibodies used in this study include anti-PARK7 (ABclonal, #A0987), anti-PRDX3 (ABclonal, #A2398), anti-HSPA9 (ABclonal, #A0558), anti-FLAG (Proteintech, #66008-3-Ig), anti-GAPDH (Biosharp, #BL006B) and anti-HSP60 (ABclonal, #A0564). Secondary antibodies used in this study include anti-rabbit IgG (TransGen, #HT101), anti-mouse IgG (TransGen, #HS201).

FLAG tag immunoprecipitation mass spectrometry analysis. PARK7 and PARK7(H138A) overexpressing stable Hela cell lines were generated as mentioned above. Then each cell line was cultured in SILAC heavy and light medium for over 5 passages. Thereafter, cells were seeded in 10-cm dishes with

vehicle or CCCP treatment. Cells were then lysed in 500 µL Pierce IP Lysis Buffer with 1x EDTA free protease inhibitor for 30 min at 4 °C. The supernatant was collected in 1.5 mL centrifuge tube and centrifuged at 15,871×g for 10 min at 4 ℃. The concentration of the supernatant was adjusted to 2 mg/mL using a BCA protein assay kit. The heavy and light proteomes were mixed 1:1 according to different experimental designs. The mixture was incubated with 50 µL anti-Flag magnetic beads with bottom-up rotation for 40 min at room temperature. The beads were washed with 1 mL PBS for three times. Thereafter, 300 µL 25 mM ABC containing 1 M urea was added, and then the proteome was reduced with 10 mM DTT at 55 °C for 40 min and alkylated by 15 mM IAA in dark at room temperature for 30 min. Additional 5 mM of DTT was added for 10 min to quench the alkylation process. 8 µg of trypsin was added and performed at 37 °C with rotation overnight. The reaction was quenched by adding formic acid till pH reached to 2-3. The peptide solution was desalted using SOLAµ HRP column (Thermo, #60209-001) and dried in speedvac vacuum concentrator. Peptides were redissolved in 0.1% formic acid and 1 µg peptides were analyzed by LC-MS/MS. The raw data were processed using the MSFragger-based Fragpipe software and the built-in SILAC workflow.

Isolation of crude mitochondria. PARK7 and PARK7(H138A) overexpressing stable Hela cell lines were seeded in a 10-cm dish. When reached to ~90% confluency, cells were treated with DMSO or 20 µM CCCP in HBSS buffer for 1h. Cells were washed twice with PBS, and 1 mL of trypsin solution was added to detach cells. 6 mL of DMEM was added to stop trypsinization and detach cells, and cells were centrifuged at 600×g for 5 min at 4 °C in 15-mL polypropylene tube. After centrifugation, the supernatant was discarded, and the cell pellet was washed twice with PBS. Cells were centrifuged at 600×g for 5 min at 4 °C. After centrifugation, the supernatant was discarded, and cell pellet was resuspended in 2 mL of ice-cold IB cells buffer (225 mM mannitol (Energy-chemicals, #69-65-8), 75 mM sucrose (Sangon, #A100335-0250), 0.07 mM EDTA and 30 mM Tris–HCl, pH=7.4) containing EDTA-free protease inhibitor. Cells were homogenized using a Dounce Glass Tissue Grinder for 30 times on ice. The homogenate was centrifuged

at 600×g for 5 min at 4 °C, and the supernatant was collected. The pellet (containing unbroken cells and nucleus) was discarded. Then, the supernatant was centrifuged at 7000×g for 10 min at 4 °C. After centrifugation, the supernatant was cytosolic fraction, and the pellet was crude mitochondria. Subsequently, these samples were used for Western blot analysis.

Solvent accessible surface area (SASA) analysis. All the structures used for solvent accessible surface area (SASA) analysis were obtained from the Protein Data Bank (PDB) [64] or the AlphaFold Protein Structure Database[52]. The absolute SASA of each residue was computed using the FreeSASA program[65]. Only SASA data of both the labeled histidines and their neighbors that were complete and unambiguous were used to obtain the average SASA for each structure. The Relative Solvent Accessibility (RSA) for each histidine was calculated by dividing absolute SASA values by the empirical maximum possible solvent accessible surface area of the residue[66]. Then all the histidines were classified as buried if the average RSA were lower than 20%, and exposed otherwise[67].

Domain enrichment analysis. Labeled histidine sites were mapped to Prosite domain database (https://prosite.expasy.org/) annotated in human UniProt proteome (https://www.uniprot.org/). The matched domain containing the labeled histidine(s) was considered a 'hit' and counted as enrichment of a domain by our method. Following Hsu *et al*'s analysis rules[13]: 1) several sites within the same domain annotation were considered as a single hit; 2) If a site had several annotations each one was considered a hit.

In the reference UniProt human database, the total number of domains is N. The total number of the domain hits in our experimental data that could match with human UniProt proteome is K. For each domain, the total number of its occurrence in the UniProt reference database is n, whereas its occurrence in our experimental data is k. The P values for each domain were calculated using a binomial test that is widely used in gene ontology (GO) statistical analysis[68]:

744 $$P = \sum \binom{K}{k} (n/N)^k (1 - n/N)^{K-k}$$

745 The *P* values was then corrected for a 1% false discovery rate using the Benjamini-Hochberg
746 correction for multiple hypothesis testing. Prosite domains that showed statistically significant
747 overrepresentation (Q <0.01) were used to generate the bubble plot in R. Similar analysis was applied in
748 DrugBank and non-DrugBank proteins that contain the labeled histidines.

749

750 **Data analysis.** Raw files acquired in DDA mode were searched against corresponding SwissProt-reviewed
751 protein databases containing common contaminants using Proteome Discoverer (v2.5) or MSfragger
752 (Fragpipe v15.0). Peptides were required to be fully tryptic with a maximum of two missed cleavage sites,
753 carbamidomethylation as fixed modification, and methionine oxidation as a dynamic modification. The
754 precursor and fragment mass tolerance were set to 10 ppm and 0.02 Da (MS2 orbitrap), respectively.
755 Contaminant hits were removed, and proteins were filtered to obtain a false discovery rate of <1%. Protein
756 subcellular localization analysis was enabled by Gene Ontology (GO) analysis from DAVID Bioinformatics
757 Resources.

758

759 **Statistics and reproducibility.** Three biological replicates were performed with similar results. Statistical
760 analysis was performed on GraphPad Prism (GraphPad Software) and data processing were performed
761 using Python (3.10.11). For comparison between two groups, p values were determined using two-sided
762 Student's t test. Error bars represent means ± SD. No statistical method was used to predetermine sample
763 size. The experiments were not randomized. The Investigators were not blinded to allocation during
764 experiments and outcome assessment.

765

766 **Reporting summary.** Further information on research design is available in the Nature Research Reporting

Summary linked to this article.

Data availability

The mass spectrometry data generated in this study have been deposited to the ProteomeXchange Consortium via the iProX[69] partner repository with the dataset identifier PXD042377 (Histidine_Profiling_MS dataset). Source data are provided as a Source Data file.

References

1. Liu Y, Patricelli MP, Cravatt BF. Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* **96**, 14694-14699 (1999).

2. Cravatt BF, Wright AT, Kozarich JW. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383-414 (2008).

3. Long JZ, Cravatt BF. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem Rev* **111**, 6022-6063 (2011).

4. Patricelli MP*, et al.* Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **46**, 350-358 (2007).

5. Kato D*, et al.* Activity-based probes that target diverse cysteine protease families. *Nature Chemical Biology* **1**, 33-38 (2005).

6. Saghatelian A, Jessani N, Joseph A, Humphrey M, Cravatt BF. Activity-based probes for the proteomic profiling of metalloproteases. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10000-10005 (2004).

7. Vocadlo DJ, Bertozzi CR. A strategy for functional proteomic analysis of glycosidase activity from cell lysates. *Angew Chem Int Ed Engl* **43**, 5338-5342 (2004).

8. Kumar S, Zhou B, Liang FB, Wang WQ, Huang ZH, Zhang ZY. Activity-based probes for protein tyrosine phosphatases. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 7943-7948 (2004).

9. Bachovchin DA, Brown SJ, Rosen H, Cravatt BF. Identification of selective inhibitors of uncharacterized enzymes by high-throughput screening with fluorescent activity-based probes. *Nat Biotechnol* **27**, 387-394 (2009).

10. Nomura DK, Long JZ, Niessen S, Hoover HS, Ng SW, Cravatt BF. Monoacylglycerol Lipase Regulates a Fatty Acid Network that Promotes Cancer Pathogenesis. *Cell* **140**, 49-61 (2010).

11. Weerapana E*, et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790-795 (2010).

12. Hacker SM, Backus KM, Lazear MR, Forli S, Correia BE, Cravatt BF. Global profiling of lysine reactivity and ligandability in the human proteome. *Nature Chemistry* **9**, 1181-1190 (2017).

13. Hahm HS*, et al.* Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nature Chemical Biology* **16**, 150-160 (2020).

14. Lin SX*, et al.* Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597-602 (2017).

15. Ma N*, et al.* 2H-Azirine-Based Reagents for Chemoselective Bioconjugation at Carboxyl Residues Inside Live Cells. *Journal of the American Chemical Society*, (2020).

16. Bach K, Beerkens BLH, Zanon PRA, Hacker SM. Light-Activatable, 2,5-Disubstituted Tetrazoles for the Proteome-wide Profiling of Aspartates and Glutamates in Living Bacteria. *ACS Cent Sci* **6**, 546-554 (2020).

17. Gutteridge A, Thornton JM. Understanding nature's catalytic toolkit. *Trends Biochem Sci* **30**, 622-629 (2005).

18. Parsons WH*, et al.* AIG1 and ADTRP are atypical integral membrane hydrolases that degrade bioactive FAHFAs. *Nature Chemical Biology* **12**, 367-+ (2016).

19. Dokmanic I, Sikic M, Tomic S. Metals in proteins: correlation between the metal-ion type, coordination number and the amino-acid residues involved in the coordination. *Acta Crystallogr D Biol Crystallogr* **64**, 257-263 (2008).

20. Martinez-Fabregas J, Rubio S, Diaz-Quintana A, Diaz-Moreno I, De la Rosa MA. Proteomic tools for the analysis of transient interactions between metalloproteins. *Febs J* **278**, 1401-1410 (2011).

21. Watanabe H*, et al.* Histidine-Mediated Intramolecular Electrostatic Repulsion for Controlling pH-Dependent Protein-Protein Interaction. *Acs Chem Biol* **14**, 2729-2736 (2019).

22. Hindupur SK*, et al.* The protein histidine phosphatase LHPP is a tumour suppressor. *Nature* **555**, 678-682 (2018).

23. Srivastava S*, et al.* Histidine phosphorylation relieves copper inhibition in the mammalian potassium channel KCa3.1. *Elife* **5**, (2016).

24. Wilke KE, Francis S, Carlson EE. Activity-Based Probe for Histidine Kinase Signaling. *Journal of the American Chemical Society* **134**, 9150-9153 (2012).

25. Jia S, He D, Chang CJ. Bioinspired Thiophosphorodichloridate Reagents for Chemoselective Histidine Bioconjugation. *J Am Chem Soc* **141**, 7294-7301 (2019).

26. Takaoka Y, Tsutsumi H, Kasagi N, Nakata E, Hamachi I. One-pot and sequential organic chemistry on an enzyme surface to tether a fluorescent probe at the proximity of the active site with restoring enzyme activity. *J Am Chem Soc* **128**, 3273-3280 (2006).

27. Li J*, et al.* ACR-Based Probe for the Quantitative Profiling of Histidine Reactivity in the Human Proteome. *J Am Chem Soc*, (2023).

28. Wan C*, et al.* Histidine-specific bioconjugation via visible-light-promoted thioacetal activation. *Chem Sci* **13**, 8289-8296 (2022).

29. Chen X*, et al.* Histidine-Specific Peptide Modification via Visible-Light-Promoted C-H Alkylation. *J Am Chem Soc* **141**, 18230-18237 (2019).

30. Nakane K*, et al.* Proximity Histidine Labeling by Umpolung Strategy Using Singlet Oxygen. *J Am Chem Soc*, (2021).

31. Zhai Y*, et al.* Spatiotemporal-resolved protein networks profiling with photoactivation dependent proximity labeling. *Nat Commun* **13**, 4906 (2022).

32. Zhao X, Liu J, Fan J, Chao H, Peng X. Recent progress in photosensitizers for overcoming the challenges of photodynamic therapy: from molecular design to application. *Chem Soc Rev* **50**, 4185-4219 (2021).

33. Xu W*, et al.* Three-Pronged Attack by Homologous Far-red/NIR AIEgens to Achieve 1+1+1>3 Synergistic Enhanced Photodynamic Therapy. *Angew Chem Int Ed Engl* **59**, 9610-9616 (2020).

34. Luo H*, et al.* Photocatalytic Chemical Crosslinking for Profiling RNA-Protein Interactions in Living Cells. *Angew Chem Int Ed Engl* **61**, e202202008 (2022).

35. Kitamura T, Nakata H, Takahashi D, Toshima K. Hypocrellin B-based activatable photosensitizers for specific photodynamic effects against high H(2)O(2)-expressing cancer cells. *Chem Commun (Camb)* **58**, 242-245 (2021).

36. Baier J, Maisch T, Maier M, Engel E, Landthaler M, Baumler W. Singlet oxygen generation by UVA light exposure of endogenous photosensitizers. *Biophys J* **91**, 1452-1459 (2006).

37. Fang Y, Zou P. Photocatalytic Proximity Labeling for Profiling the Subcellular Organization of Biomolecules. *Chembiochem* **24**, e202200745 (2023).

38. Tamura T, Takato M, Shiono K, Hamachi I. Development of a Photoactivatable Proximity Labeling Method for the Identification of Nuclear Proteins. *Chem Lett* **49**, 145-148 (2020).

39. Liu H*, et al.* Antigen-Specific T Cell Detection via Photocatalytic Proximity Cell Labeling (PhoXCELL). *J Am Chem Soc* **144**, 5517-5526 (2022).

40. Xu F*, et al.* Hypoxia-activated NIR photosensitizer anchoring in the mitochondria for photodynamic therapy. *Chem Sci* **10**, 10586-10594 (2019).

41. Ma H*, et al.* New Cy5 photosensitizers for cancer phototherapy: a low singlet-triplet gap provides high quantum yield of singlet oxygen. *Chem Sci* **12**, 13809-13816 (2021).

42. Bauer D, Montforts FP, Losi A, Gorner H. Photoprocesses of chlorin e6 glucose derivatives. *Photochem Photobiol Sci* **11**, 925-930 (2012).

43. Liu S, Feng G, Tang BZ, Liu B. Recent advances of AIE light-up probes for photodynamic therapy. *Chem Sci* **12**, 6488-6506 (2021).

44. Hu F, Xu S, Liu B. Photosensitizers with Aggregation-Induced Emission: Materials and Biomedical Applications. *Adv Mater* **30**, e1801350 (2018).

45. Wang P*, et al.* Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nat Chem Biol* **15**, 1110-1119 (2019).

46. Hananya N, Ye X, Koren S, Muir TW. A genetically encoded photoproximity labeling approach for mapping protein territories. *Proc Natl Acad Sci U S A* **120**, e2219339120 (2023).

47. Weerapana E, Speers AE, Cravatt BF. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes. *Nat Protoc* **2**, 1414-1425 (2007).

48. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **14**, 513-520 (2017).

49. Pattison DI, Rahmanto AS, Davies MJ. Photo-oxidation of proteins. *Photochem Photobiol Sci* **11**, 38-53 (2012).

50. Grassi L, Cabrele C. Susceptibility of protein therapeutics to spontaneous chemical modifications by oxidation, cyclization, and elimination reactions. *Amino Acids* **51**, 1409-1431 (2019).

51. Marconi G, Quintana R. Methylene blue dyeing of cellular nuclei during salpingoscopy, a new in-vivo method to evaluate vitality of tubal epithelium. *Hum Reprod* **13**, 3414-3417 (1998).

52. Varadi M*, et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022).

53. Jumper J*, et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

54. Klein DJ, Moore PB, Steitz TA. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10**, 1366-1379 (2004).

55. Dang L*, et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739-744 (2009).

56. Perez-Alvarado GC, Kosa JL, Louis HA, Beckerle MC, Winge DR, Summers MF. Structure of the cysteine-rich intestinal protein, CRIP. *J Mol Biol* **257**, 153-174 (1996).

57. Cheng Y*, et al.* Co-evolution-based prediction of metal-binding sites in proteomes by machine learning. *Nat Chem Biol* **19**, 548-555 (2023).

58. Zhou Y*, et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).

59. Imberechts D, Vandenberghe W. Defects in PINK-PRKN-PARK7/DJ-1-dependent mitophagy and autosomal recessive Parkinson disease. *Autophagy*, 1-2 (2022).

60. Dolgacheva LP, Berezhnov AV, Fedotova EI, Zinchenko VP, Abramov AY. Role of DJ-1 in the mechanism of pathogenesis of Parkinson's disease. *J Bioenerg Biomembr* **51**, 175-188 (2019).

61. Hu S*, et al.* Molecular chaperones and Parkinson's disease. *Neurobiol Dis* **160**, 105527 (2021).

62. Ryu KA, Kaszuba CM, Bissonnette NB, Oslund RC, Fadeyi OO. Interrogating biological systems using visible-light-powered catalysis. *Nature Reviews Chemistry* **5**, 322-337 (2021).

63. Zhang X*, et al.* An Effective Synthetic Entry to Fused Benzimidazoles via Iodocyclization. *Adv Synth Catal* **353**, 1429-1437 (2011).

64. Rose PW*, et al.* The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res* **41**, D475-482 (2013).

65. Mitternacht S. FreeSASA: An open source C library for solvent accessible surface area calculations. *F1000Res* **5**, 189 (2016).

66. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. Maximum allowed solvent accessibilites of residues in proteins. *Plos One* **8**, e80635 (2013).

67. Savojardo C, Manfredi M, Martelli PL, Casadio R. Solvent Accessibility of Residues Undergoing Pathogenic Variations in Humans: From Protein Structures to Protein Sequences. *Front Mol Biosci* **7**, 626363 (2020).

68. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **8**, 1551-1566 (2013).

69. Ma J*, et al.* iProX: an integrated proteome resource. *Nucleic Acids Res* **47**, D1211-D1217 (2019).

Acknowledgements

We thank the mass spectrometry core facility, imaging core facility and sequencing core facility in Shenzhen Bay Laboratory for their assistance in running samples. We thank Dr. Xiaoyu Li, Dr. Yuan Liu, Dr. Chu Wang and Ms. Weiye Zhong for helpful discussion and proofreading assistance. We are grateful for financial support of this work from the following: Grant from the National Natural Science Foundation (32101200), Grant from Guangdong-Shenzhen Regional Joint Fund (2020A1515110903), and Grant from Shenzhen Bay Laboratory Open Fund (SZBL2020090501008).

Author contributions

All authors reviewed the manuscript. G.L. conceived of the study and supervised research. Y.Z., X.Z., Y.H., Y.T., and G.L. designed and analyzed experiments. Y.Z., X.Z., Y.H. performed experiments. X.Y. wrote the python program for data processing and generated the figures. K.T. did the domain enrichment analysis. Z.Z. performed the SASA analysis, secondary structure distribution analysis and distance measurement between labeled histidine and active site. Y.Z. and G.L. wrote the manuscript.

Competing interests: The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary20230524ZYS.pdf