

Cite this: *Chem. Sci.*, 2022, 13, 8289

All publication charges for this article have been paid for by the Royal Society of Chemistry

## Histidine-specific bioconjugation *via* visible-light-promoted thioacetal activation†

Chuan Wan,<sup>‡a</sup> Yuena Wang,<sup>‡a</sup> Chenshan Lian,<sup>‡b</sup> Qi Chang,<sup>b</sup> Yuhao An,<sup>b</sup> Jiean Chen,<sup>‡b</sup> Jinming Sun,<sup>a</sup> Zhanfeng Hou,<sup>b</sup> Dongyan Yang,<sup>c</sup> Xiaochun Guo,<sup>a</sup> Feng Yin,<sup>\*b</sup> Rui Wang<sup>‡\*b</sup> and Zigang Li<sup>‡\*ab</sup>

Histidine (His, H) undergoes various post-translational modifications (PTMs) and plays multiple roles in protein interactions and enzyme catalyzed reactions. However, compared with other amino acids such as Lys or Cys, His modification is much less explored. Herein we describe a novel visible-light-driven thioacetal activation reaction which enables facile modification on histidine residues. An efficient addition to histidine imidazole N3 under biocompatible conditions was achieved with an electrophilic thionium intermediate. This method allows chemo-selective modification on peptides and proteins with good conversions and efficient histidine-proteome profiling with cell lysates. 78 histidine containing proteins were for the first time found with significant enrichment, most functioning in metal accumulation in brain related diseases. This facile His modification method greatly expands the chemo-selective toolbox for histidine-targeted protein conjugation and helps to reveal histidine's role in protein functions.

Received 26th April 2022  
Accepted 25th June 2022

DOI: 10.1039/d2sc02353a

rsc.li/chemical-science

## Introduction

Post-translational modifications (PTMs) of proteins enable a marked increase in protein functional diversity<sup>1</sup> for manipulation of protein structure and function.<sup>2</sup> Various chemical tools have been developed to assess modified sites of proteins, exploring a large quantity of important information about protein interactions and/or enzymatic mechanisms in drug development, molecular biology and medicine.<sup>3–6</sup> However, most labeling strategies are limited to the functionalization of nucleophilic residues, namely cysteine, lysine and tyrosine.<sup>7–13</sup> Continuous efforts have been invested to target less-explored amino acids, including hydroxyl (threonine/serine),<sup>14</sup> carboxyl (aspartic acid/glutamic acid),<sup>15</sup> tryptophan,<sup>16–21</sup> and methionine.<sup>22–24</sup>

The His residue has an electron-deficient heteroaromatic imidazole side chain and a low abundance of ~2.2% (ref. 25) in proteins. It plays many important roles in protein functions including as a hydrogen bond donor/acceptor, and in proton

shuttling, metal binding coordination (Scheme S1a†),<sup>26–30</sup> metal-directed covalent modification (Scheme S1b†)<sup>31</sup> and nucleophilic catalysis.<sup>32–36</sup> Histidine phosphorylation has been extensively studied using neutral loss fragmentation<sup>37</sup> and has garnered increasing interest in recent years;<sup>38</sup> however, there is still no robust modification method specifically targeting the histidine residue.

Based on the nucleophilicity of the NH ring, Hamachi *et al.* reported an affinity-based labeling strategy to modify protein His residues *via* epoxide ring opening (Fig. 1a and Scheme S2a†).<sup>39</sup> A direct modification of His-tags by thiophosphorylation under weakly alkaline conditions (pH 8.5) (Fig. 1a and Scheme S2b†) was reported by Chang *et al.*<sup>34</sup> In addition to the modification on the N<sup>3</sup> position of His residues, Wang and Chen *et al.* developed a selective C–H alkylation on the C<sup>2</sup> position of His in peptides and/or proteins *via* a visible-light-promoted approach (Fig. 1a and Scheme S2c†) with the requirement of a strong acid (trifluoroacetic acid, TFA) and organic solvent (2,2,2-trifluoroethanol, TFE).<sup>40</sup> Besides, sequence-dependent strategies were demonstrated for PEGylation of protein His residues (Scheme S2d†).<sup>41,42</sup> Recently, Nakamura *et al.* utilized a nucleophilic small molecule (1-methyl-4-arylurazole) to selectively label histidine under singlet oxygen (<sup>1</sup>O<sub>2</sub>) conditions generated with a ruthenium catalyst under white LED light.<sup>36</sup> However, a biocompatible and selective modification of His on proteins still remains a challenge for chemo-proteomic study.

Organic sulfur(IV) molecules, such as sulfonium, are generally electrophilic and can undergo classical nucleophilic

<sup>a</sup>State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, 518055, P. R. China. E-mail: lizg.sz@pku.edu.cn

<sup>b</sup>Pingshan Translational Medicine Center, Shenzhen Bay Laboratory, Shenzhen, 518118, P. R. China. E-mail: yinfeng@szbl.ac.cn; wangrui@szbl.ac.cn; lizg@szbl.ac.cn

<sup>c</sup>College of Chemistry and Chemical Engineering, Zhongkai University of Agriculture and Engineering, Guangzhou, 510225, P. R. China

† Electronic supplementary information (ESI) available. See <https://doi.org/10.1039/d2sc02353a>

‡ These authors contributed equally to this work.



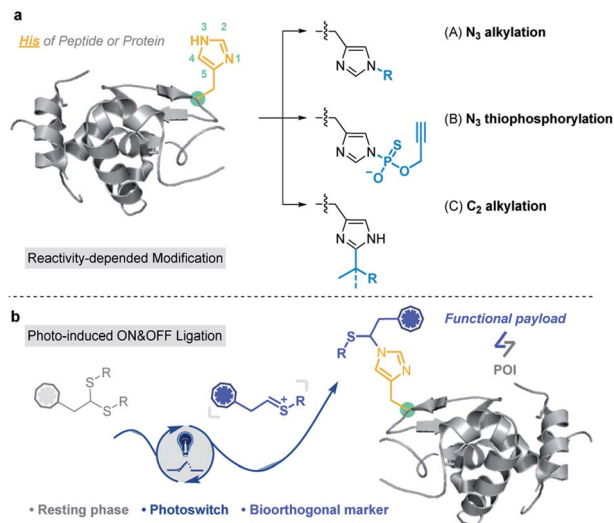


Fig. 1 Histidine-specific bioconjugation: (a) typical methods for His-specific modification of peptides and proteins; (b) visible-light-promoted His-specific bioconjugation.

substitution reactions as a suitable leaving group in  $S_NAr$  reactions.<sup>43–45</sup> For example, the methyltransferase-catalyzed nucleophilic methylation reaction between *S*-adenosylmethionine (SAM) and the histidine residue of proteins is an important PTM besides arginine and lysine methylations.<sup>46,47</sup> On the other hand, thionium is an important intermediate state of Pummerer-type reactions, and is a highly active substrate in nucleophilic  $S_EAr$  reactions.<sup>48–53</sup> Recently, our group developed various sulfonium-based chemistries for bioorthogonal applications with excellent biocompatibility.<sup>54–56</sup> Thus, we considered the application of electrophilic thionium intermediates for protein modification. Herein, we demonstrated a specifically biocompatible method for His-specific modification on proteins *via* visible-light-promoted nucleophilic substitution. After reaction condition screening and optimization, an optimal condition for His modification was achieved in a biocompatible manner. ~2000 reactive and exposed histidine-containing proteins were characterized from the MCF7 cell line using activity-based protein profiling (ABPP) (Fig. 1b).

## Results and discussion

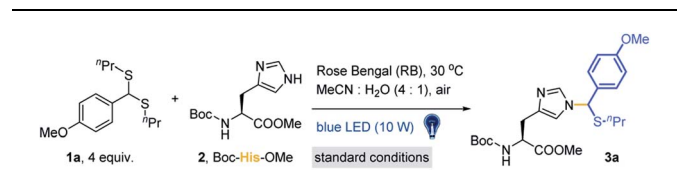
### Design and condition optimization

Lewis acid catalysed Pummerer-type  $S_EAr$  reactions are well documented for aromatic compound modifications (Scheme S3†).<sup>51,52,57</sup> However, the reaction conditions, such as organic solvents, Lewis acids and heating/cooling are generally bio-incompatible. The visible-light triggered cleavage of the C–S bond was also reported as an efficient pathway to provide electrophilic cationic intermediates.<sup>53,58</sup> Thus, we checked the possibility of thioacetal **1a** as a precursor of thionium for the Pummerer-type  $S_EAr$  reaction under visible light in the absence of Lewis acids. To develop a more bio-compatible condition, the reaction between Boc-His-OMe **2** and thioacetal **1a** in MeCN/20% H<sub>2</sub>O (MeCN used in condition screening to avoid solubility

issues) was performed under irradiation with blue LEDs (10 W, 450 nm). But this reaction gave very limited product. Then Lewis acids, transition metals and organic photocatalysts (see Fig. S1†) were examined in weakly polar MeCN/H<sub>2</sub>O solvent to study their effects. Rose Bengal (RB) was found to be the most efficient catalyst (10 mol% loading) (Table S1†, entries 1–13) with excellent yield (84%, Table 1, entry 1). The kinetic study was conducted at 4-fold excess of **1a** with RB as catalyst, and the reaction reached equilibrium at 1 hour (Fig. S2†). Other conditions were further tested, including dark, 50 °C heating, the absence of catalyst, lower catalyst loading (Table 1, entries 2–5) or non-blue LED light irradiation (Table S1†, entries 14–17). The results indicated that the reaction is initiated by light and favours blue light. The use of iridium and MesAcrClO<sub>4</sub> catalyst gave lower isolated yields (39% and 66%, Table 1, entries 6 and 7). Prolonging the reaction time (to 4 h) gave very limited improvement (85%, Table 1, entry 8). The addition of AcOH and (NH<sub>4</sub>)<sub>2</sub>CO<sub>3</sub> had a negligible effect on the reaction (83% and 80%, Table 1, entries 9 and 10), suggesting that the reaction is robust and tolerates bio-relevant pHs. Notably, due to the newly introduced chiral center on the imidazole ring, the products are mixtures of a pair of epimers. Also, the reacted position was further confirmed to be N3 in the imidazole ring by 2D NMR analysis (see detailed data and discussion in the ESI†).

The chemo-selectivity of this reaction was then checked and no reaction was detected for Phe, Tyr, Trp, Ser, Lys, Arg, Glu and Gln (Fig. S3 and S4†). Due to the potential oxidative conditions, the reaction of Met was checked, and about 62% oxidative conversion (sulfoxide) was observed. Taking the low abundance of Met (~2%)<sup>59</sup> in the human proteome and the highly reductive cellular circumstances into consideration, Met oxidation is not

Table 1 Optimization of the photocatalyzed reaction of Boc-His-OMe **2** and thioacetal **1a**



Entry	Change from standard conditions <sup>a</sup>	Yield <sup>b</sup> (%)
1	None	84
2	Dark	Trace <sup>c</sup>
3	50 °C in the dark	Trace <sup>c</sup>
4	Catalyst is absent	<10
5	5% catalyst loading	68
6	[Ir(ppy) <sub>2</sub> (dtbpy)]PF <sub>6</sub> instead of RB	39
7	MesAcrClO <sub>4</sub> instead of RB	66
8	4 h instead of 1 h	85
9	Addition of AcOH (2 equiv.)	83
10	Addition of (NH <sub>4</sub> ) <sub>2</sub> CO <sub>3</sub> (2 equiv.)	80
11	Addition of TEMPO (1 equiv.)	Trace <sup>c</sup>

<sup>a</sup> Standard conditions: thioacetal **1a** (20 mM), Boc-His-OMe **2** (5 mM) and 10 mol% Rose Bengal (RB) in MeCN/H<sub>2</sub>O (4/1) under irradiation with a blue LED (10 W) at 30 °C under air for 1 hour. <sup>b</sup> Isolated yield obtained by column chromatography. <sup>c</sup> Determined by LC-MS. TEMPO = 2,2,6,6-tetramethylpiperidine-1-oxyl.



likely to have a significant effect in protein profiling. It was discovered that few proteins changed as a result of oxidative damage in accordance with former reports.<sup>60</sup> Notably, the indole side chain of Trp reacted well with thioacetal **1a** to give the S<sub>E</sub>Ar product in TFE under irradiation with blue light (450 nm); however, no reaction was detected in aqueous media, which was also confirmed in our following profiling study (Scheme S4 and Fig. S4†). The thiol side chain of Cys rapidly reacts with **1a** under acidic conditions (both Lewis acid and Brønsted acid), but only gives a trace amount of substitution product under optimized conditions (Fig. S5†). In addition, the stability of product **3a** was checked under acidic conditions, including pH = 4 to 7 PBS buffer and HCOOH or TFA solvents. The results indicated partial hydrolysis of the product in weakly acidic PBS buffer after 24 hours (Fig. S6a†) and almost complete hydrolysis to **2** and 4-methoxybenzaldehyde in strong Brønsted acid-H<sub>2</sub>O solvent (Fig. S6b†).

Various substrates and peptides were prepared and reacted under the optimized conditions to further clarify the reaction between thioacetal and the His side chain. As shown in Fig. 2a, the substrates derived from hydroxy and propargyloxy benzaldehyde were observed to have similar yields (86% for **3b** and 81% for **3d**), and the thioacetals of ethanethiol and prop-2-ene-1-thiol provided slightly decreased yields (79% for **3c** and 62% for **3e**). The reaction of pentanal-propane-1-thiol acetal only gave a moderate yield (47% for **3f**), and no product was detected with formaldehyde thioacetal (**3g**). Three bioactive peptides, leuprorelin, angiotensin II and melanotan I, were applied for the further study of chemo-selectivity of this reaction. Four thioacetal substrates were reacted with peptide **4**, and good yields were observed for **4a** (70%). Likewise, the reaction of ethanethiol, prop-2-ene-1-thiol and hydroxyethyl derived thioacetals provided relatively lower yields (34% for **4b**, 47% for **4c** and 27% for **4d**). The standard reaction was conducted between **1a** and peptide **5** (51% yield) and peptide **6** (a Lys containing 13 AA peptide). Even though slightly lower yield was observed (36%) for peptide **6**, the MS/MS analysis gave further evidence for the chemo-selectivity of histidine modification (see the ESI† for additional data). Also, the Met containing peptides **7** and **8** were designed for the study of influence of the oxidative conditions. For the double-Met containing peptide **7**, 35% yield of thioacetal adducted products and 23% yield of products with both oxidized Met and modified His were observed under the standard conditions. Furthermore, peptide **8**, which contains all of the 14 reactive AA residues, was designed for the chemo-selectivity study. With the protection of acetamide on Cys, 34% target product and 33% Met-oxidation thioacetal-adducted product were observed (see the ESI† for additional data). As a result, the reaction exhibited good efficiency and chemo-selectivity, and the desired reaction is more likely to occur than oxidative side reactions. In addition, dithiol, mercaptoethanol and mercaptoethylamine derived acetals, thioketals and  $\alpha$ -carbamoylsulfides were examined and no desired products were detected under optimized conditions (Fig. S7†).

To clarify the reaction pathway, 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) was added under optimized conditions, and the reaction was largely inhibited (Table 1, entry 11), suggesting that

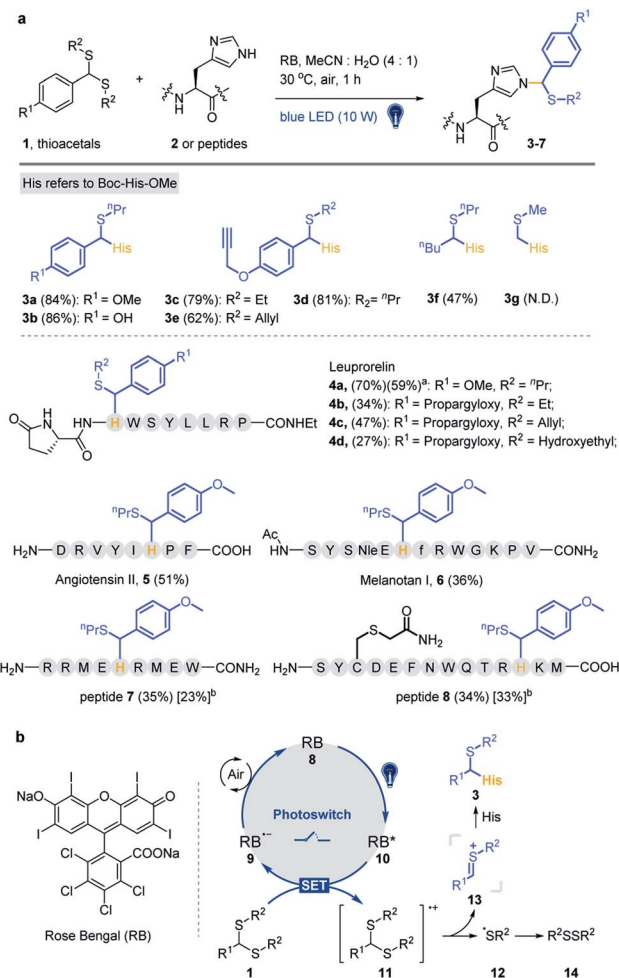


Fig. 2 Visible-light-promoted Pummerer-type reaction between thioacetal and the side chain of His. (a) Reaction between thioacetals **1** and Boc-His-OMe/peptides. Standard conditions: thioacetal (20 mM), Boc-His-OMe **2** (5 mM) or peptide (0.5–1 mM) and 10 mol% RB in MeCN/H<sub>2</sub>O (4/1, pH 7.4) under irradiation with a blue LED (10 W, 450 nm) at 30 °C under air for 1 h. The % yields of peptide products were determined from reverse-phase HPLC-MS with an internal standard (dibenzyl sulfoxide). <sup>a</sup>Isolated yield. <sup>b</sup>The yields in square parentheses refer to thioacetal-adducted products with Met oxidation. (b) Proposed mechanism.

a radical mechanism is involved.<sup>53,61,62</sup> The Stern-Volmer fluorescence quenching experiments suggest an energy transfer event between the excited-state RB and the thioacetal **1a** in a dose-dependent manner (Fig. S8a and S8b†). A reductive quenching mechanism is proposed in Fig. 2b. The reaction is initiated by irradiating RB **8** to its excited state with 10 W blue light and the strong single-electron oxidant **9** undergoes single-electron-transfer (SET) with thioacetal **1** to generate RB radical anion **10** and sulfur radical cation **11**. Subsequent loss of the sulfur radical from **11** furnishes thionium **13** and thiyl radical **12** which dimerizes to yield disulfide **14** (detected by GC-MS in Fig. S8c†). The transient intermediate **13** then undergoes nucleophilic addition with the 3-NH of His imidazole to afford the desired product **3**. Meanwhile, the photocatalyst RB is regenerated by the oxidation of **9** by air.





## Selective labeling in model protein and cell lysates

To expand the chemo-selective reaction of the thioacetal group on histidine residues, biotin-modified thioacetal probes (**TA1** and **TA2** in Fig. 3a) were synthesized and incubated with model protein carbonic anhydrase (CA, containing histidine residues but no cysteine residue) followed by the western blot. Two clearly visible bands confirmed the histidine labeling on CA (Fig. 3b). They also confirmed histidine labeling in the complex context of MCF7 cell lysates but **TA2** showed a reduced labeling efficiency (Fig. 3b). Then a series of alkyne-tagged probes (**TA3–7**, Fig. 3a) were prepared and incubated with model protein BSA in PBS (pH 7.4) for 2 h at 37 °C with 5% RB under blue light. Upon CuAAC “click” chemistry<sup>61</sup> with rhodamine-azide (**TAMRA-N<sub>3</sub>**), the reaction mixtures were subjected to SDS-PAGE and analyzed by in-gel fluorescence scanning. The strongest fluorescence was observed from **TA4** in PBS buffers with or without RB (Fig. 3c). Blue light was also found to be the most efficient among lights from different sources (Fig. 3d) and **TA4** showed stronger labeling under weakly alkaline conditions (Fig. 3e). Further kinetics and stoichiometry study indicated that 20 equiv. of thioacetal probe with 2 h reaction time was sufficient for labeling (Fig. 3f–g). The fluorescence intensity of the labeled BSA band was further inhibited with increased competitor **1a** (Fig. 3h), revealing that the labeling was thioacetal-dependent. Pretreatment with commercial cysteine-reactive reagent **IAA** (iodoacetamide) decreased the fluorescence intensity (Fig. 3i), but there was no obvious change observed from treatment with lysine-specific reagent **NHS-Ace** (Fig. 3j). ESI-TOF MS analysis indicated quantitative **1a** labeled myoglobin (MB, containing histidine residues but no cysteine residue, Fig. S9a†). Then the labeled peptides on BSA were analyzed by LC-MS/MS (Fig. 3k and Table S3†), verifying that **TA4** selectively labeled histidine in BSA when **IAA** was applied to block cysteine.

To increase the water-solubility of **TA4**, we synthesized **TA8** (Fig. 3l) to label proteins followed by CuAAC “click” chemistry with biotin-dadps-azide (DADPS biotin-N<sub>3</sub>). It's found that **TA8** did better labeling with CA protein (Fig. 3l). Therefore, we chose **TA8** to evaluate the proteome reactivity profiles in breast cancer cell line MCF7. We for the first time confirmed the histidine labeling of **TA8** with CA by LC-MS/MS (Fig. S9b and Table S2†). Concentration-dependent labeling profiles showed that **TA8** gave visible bands at 20 μM and reached saturation at 80 μM *in vitro* (Fig. 3m). Pretreatment with **1b** showed a weaker band throughout the whole lane (Fig. 3n), which was consistent with that of competitive labeling with BSA (Fig. 3h). Consistent with the above results, **IAA** pretreatment slightly decreased the fluorescence intensity (Fig. S9c†) and **NHS-Ace** pretreatment didn't affect the fluorescent signal (Fig. S9d†).

## Proteome-wide profiling in human cells

Furthermore, we expanded the chemo-selective probes in activity-based protein profiling (ABPP) to profile reactive and exposed histidine-containing proteomes. The thioacetal group was intended to serve as a reactive electrophile; the azide-tagged version was a latent affinity handle for conjugation by “click”

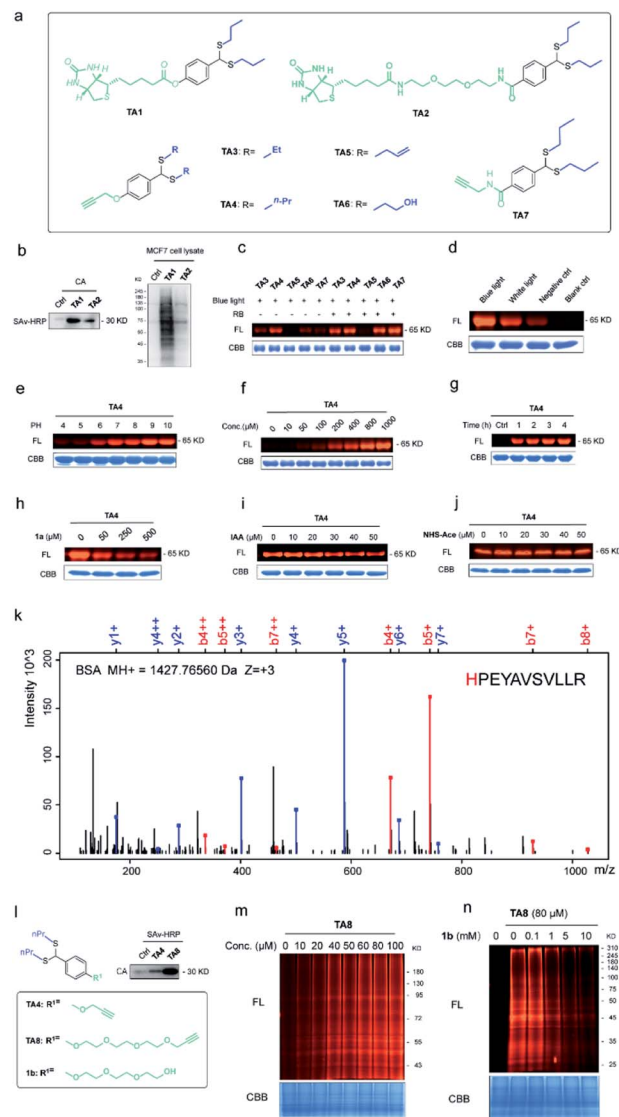
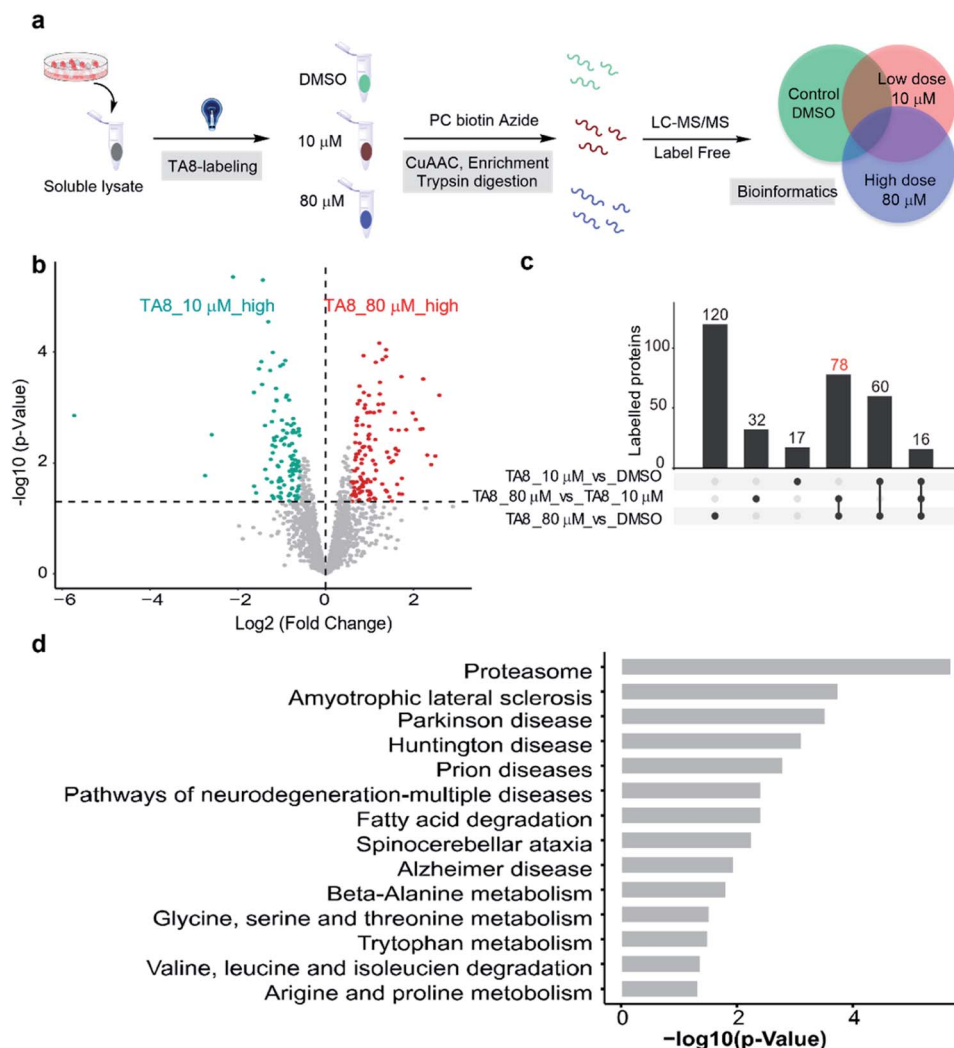


Fig. 3 Selective histidine labeling. (a) Structures of thioacetal probes. (b) Labeling of **TA1** and **TA2** with CA (CA/**TA1** and **TA2** 10/200 μM, 5% RB, pH 7.4, 37 °C for 2 h under blue light (10 W, 450 nm)) and MCF7 cell lysates (40 μg). (c) Labeling of BSA with the probes **TA3–7** (BSA/**TA3–7** 10/200 μM, pH 7.4, 37 °C for 2 h under blue light). FL = in-gel fluorescence scanning. CBB = Coomassie gel. (d) BSA (10 μM) incubated with **TA4** (200 μM) with or without light. (e) BSA (10 μM) incubated with **TA4** (200 μM) in phosphate buffer with different pHs. (f) Dose-dependent labeling of BSA (10 μM) with **TA4** for 2 h. (g) Time-dependent labeling of BSA (10 μM) with **TA4** (200 μM). (h) Competitive labeling of BSA (10 μM) with **TA4** (200 μM) in the presence or absence of competitor **1a**. (i) Labeling of BSA (10 μM) with **TA4** (200 μM) in the presence or absence of **IAA**. (j) Labeling of BSA (10 μM) with **TA4** (200 μM) in the presence or absence of **NHS-Ace**. (k) Analysis of the amino acid specificity of **TA4** with **IAA**-pretreated BSA. (l) **TA8** did better labeling with CA (BSA/**TA4** and **TA8** 10/200 μM, 1% RB, pH 7.4, 37 °C for 2 h under blue light). (m) Dose-dependent labeling of MCF7 cell lysates with **TA8**. (n) Competitive labeling of MCF7 cell lysates with **TA8** in the presence or absence of competitor **1b**.

chemistry to biotin tags for protein enrichment<sup>6</sup> (Fig. 4a). It's reasonable that the thioacetal labeled product was acid-labile; PC biotin-N<sub>3</sub> is more available for acid-labile thioacetal





**Fig. 4** Proteome-wide quantification of reactive and exposed histidine-containing proteomes. (a) General protocol for reactive and exposed histidine-containing proteome profiling by label-free TOP-ABPP. Cellular lysates are labeled with the thioacetal-group probe (TA8) at different concentrations and DMSO control. Labeled samples are conjugated to the photo-cleavable biotin tag (green, red, and blue for DMSO, and 10 and 80 μM TA8 probe treatment groups, respectively) by CuAAC click chemistry and incubated, and TA8 labeled proteins are enriched by neutravidin-conjugated beads and digested with trypsin to yield labeled proteins for LC-MS/MS analysis. (b) Volcano plot of differentially enriched proteins under dose dependent TA8 probe groups ( $n = 3$ ). (c) Venn diagram of highly enriched proteins in each dose-treatment group. (d) KEGG pathway analysis of highly reactive and exposed proteins in 80 μM TA8 proteome. Fold change  $\geq 1.5$ , false discovery rate (FDR)  $\leq 0.01$ ,  $q$ -value  $\leq 0.05$ . Proteins with significantly enriched proteome are shown in the ESI.†

modification, using 365 nm photo-irradiation to release the biotin tag. Subsequently, we selected PC biotin- $\text{N}_3$  modified TA8 to identify reactive and exposed histidine-containing proteomes by LC-MS/MS combined with label-free quantification (LFQ) methodology. Using parallel tandem orthogonal TOP-ABPP<sup>63</sup> ( $n = 3$  for three groups) in MCF7 cell lysates, and dose-dependent treatment with low-dose TA8 (10 μM), high-dose TA8 (80 μM) and DMSO-control without TA8, we identified ~1800 proteins (Fig. S10a†). Here, we used PCA (principal component analysis) to reduce the dimensionality of our datasets with a number of different variables in three experimental groups (Fig. S10b†). Comparison of TA8 (80 μM) versus TA8 (10 μM) (Fig. 4b) and TA8 (80 μM or 10 μM) versus DMSO (Fig. S10c†) indicated that 78 proteins were significantly enriched in both TA8 groups (80 μM/

10 μM and 80 μM/DMSO) (Fig. 4c and Table S4–S6†). Furthermore, we sought to characterize these highly enriched proteins using KEGG pathways. Notably, in KEGG analysis, most of the enriched proteins involved the nervous-system pathways, including Alzheimer's disease and Parkinson's disease (succinate dehydrogenase complex flavoprotein subunit A, SDHA and 26S proteasome ubiquitin receptor, ADRM1) associated with metal accumulation in the brain<sup>64</sup> (Fig. 4d and Table S7†). The significantly highly enriched proteasome here possibly comes from mammalian proteins of interest exhibiting histidine phosphorylation, including P-selectin, annexin I and the 20S proteasome, identified from emerging data.<sup>65,66</sup> Moreover, as active histidine residues are commonly found in enzyme active sites and metal-binding sites,<sup>67,68</sup> Gene Ontology (GO)<sup>69</sup> terms



analysis for 78 highly enriched proteins described their classes of biological process (BP), cellular component (CC) and molecular function (MF). It was indicated that reactive and exposed histidine-containing proteins were mainly involved in metabolic processes and it would be worth further studying their function, making this the first report of reactive and exposed histidine-containing proteomes to date.

## Conclusions

In summary, we have developed a biocompatible and selective approach for labeling histidine residues with thioacetal as a thionium precursor. Coupled with recent advances in photocatalysis, our work for the first time provides a selective and biocompatible chemical approach for histidine labeling in proteins *via* visible-light-driven nucleophilic substitution of the imidazole ring. We are optimistic that this method will provide a powerful tool to label histidine residues in native biological systems. Rapid labeling in living systems is currently in progress.

## Data availability

Primary data for the organic synthesis, in-gel fluorescence characterization, proteomic analysis, and bioinformatics analysis are provided in the ESI.†

## Author contributions

Research was conceived by all authors. Experiments were designed and performed by C. W., Y. W., C. L., R. W. and Z. L. Chemical synthesis was conducted by C. W., J. C., J. S. and Z. H. Biochemical assays and mass spectrometry experiments were performed by Y. W., C. L., Q. C., D. Y. and X. G. Bioinformatics analysis was performed by R. W. and Y. A. Cellular studies were conducted by Y. W., C. L. and F. Y. The manuscript was written and proofread by all authors.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We gratefully acknowledge financial support from the National Key Research and Development Program "Synthetic Biology" Key Special Project of China, 2018YFA0902504; the National Natural Science Foundation of China, grants 21778009 and 21977010; the Natural Science Foundation of Guangdong Province, 2019A151511184, 2020A1515010766 and 2020A1515011521; and the Shenzhen Science and Technology Innovation Committee, RCJC20200714114433053, JCYJ201805081522131455 and JCYJ20200109140406047. We acknowledge financial support from the Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions, grant 2019SHIBS0004. This work is supported by

the Proteomic Platform of Pingshan Translational Medicine Center, Shenzhen Bay Laboratory.

## Notes and references

- 1 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto Jr, *Angew. Chem., Int. Ed. Engl.*, 2005, **44**, 7342–7372.
- 2 Y. Xie, S. Du, Z. Liu, M. Liu, Z. Xu, X. Wang, J. X. Kee, F. Yi, H. Sun and S. Q. Yao, *Angew. Chem., Int. Ed. Engl.*, 2022, **61**, e202200303.
- 3 Y. Xie, L. Chen, R. Wang, J. Wang, J. Li, W. Xu, Y. Li, S. Q. Yao, L. Zhang, Q. Hao and H. Sun, *J. Am. Chem. Soc.*, 2019, **141**, 18428–18436.
- 4 O. Rinner, J. Seebacher, T. Walzthoeni, L. N. Mueller, M. Beck, A. Schmidt, M. Mueller and R. Aebersold, *Nat. Methods*, 2008, **5**, 315–318.
- 5 E. V. Vinogradova, X. Zhang, D. Remillard, D. C. Lazar, R. M. Suci, Y. Wang, G. Bianco, Y. Yamashita, V. M. Crowley, M. A. Schafroth, M. Yokoyama, D. B. Konrad, K. M. Lum, G. M. Simon, E. K. Kemper, M. R. Lazear, S. Yin, M. M. Blewett, M. M. Dix, N. Nguyen, M. N. Shokhirev, E. N. Chin, L. L. Lairson, B. Melillo, S. L. Schreiber, S. Forli, J. R. Teijaro and B. F. Cravatt, *Cell*, 2020, **182**, 1009–1026.
- 6 A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686–4687.
- 7 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.*, 2019, **3**, 147–171.
- 8 O. Koniev and A. Wagner, *Chem. Soc. Rev.*, 2015, **44**, 5495–5551.
- 9 O. Boutureira and G. J. Bernardes, *Chem. Rev.*, 2015, **115**, 2174–2195.
- 10 A. M. Marmelstein, M. J. Lobba, C. S. Mogilevsky, J. C. Maza, D. D. Brauer and M. B. Francis, *J. Am. Chem. Soc.*, 2020, **142**(11), 5078–5086.
- 11 J. Ohata, M. K. Miller, C. M. Mountain, F. Vohidov and Z. T. Ball, *Angew. Chem., Int. Ed. Engl.*, 2018, **57**, 2827–2830.
- 12 M. Lorenzi, C. Puppo, R. Lebrun, S. Lignon, V. Roubaud, M. Martinho, E. Mileo, P. Tordo, S. R. A. Marque, B. Gontero, B. Guigliarelli and V. Belle, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 9108–9111.
- 13 H. Ban, J. Gavriluk and C. F. Barbas 3rd, *J. Am. Chem. Soc.*, 2010, **132**, 1523–1525.
- 14 J. C. Vantourout, S. R. Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padiyal, A. Istrate, K. Maziarz, J. N. deGruyter, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes and P. S. Baran, *J. Am. Chem. Soc.*, 2020, **142**(41), 17236–17242.
- 15 N. Ma, J. Hu, Z. M. Zhang, W. Liu, M. Huang, Y. Fan, X. Yin, J. Wang, K. Ding, W. Ye and Z. Li, *J. Am. Chem. Soc.*, 2020, **142**(13), 6051–6059.
- 16 S. J. Tower, W. J. Hetcher, T. E. Myers, N. J. Kuehl and M. T. Taylor, *J. Am. Chem. Soc.*, 2020, **142**, 9112–9118.
- 17 M. Imiolek, G. Karunanithy, W. L. Ng, A. J. Baldwin, V. Gouverneur and B. G. Davis, *J. Am. Chem. Soc.*, 2018, **140**, 1568–1571.



- 18 Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki and M. Kanai, *J. Am. Chem. Soc.*, 2016, **138**, 10798–10801.
- 19 J. M. Antos and M. B. Francis, *J. Am. Chem. Soc.*, 2004, **126**, 10256–10257.
- 20 J. M. Antos, J. M. McFarland, A. T. Iavarone and M. B. Francis, *J. Am. Chem. Soc.*, 2009, **131**, 6301–6308.
- 21 B. V. Popp and Z. T. Ball, *J. Am. Chem. Soc.*, 2010, **132**, 6660–6662.
- 22 S. Lin, X. Yang, S. Jia, A. M. Weeks, M. Hornsby, P. S. Lee, R. V. Nichiporuk, A. T. Iavarone, J. A. Wells, F. D. Toste and C. J. Chang, *Science*, 2017, **355**, 597–602.
- 23 M. T. Taylor, J. E. Nelson, M. G. Suero and M. J. Gaunt, *Nature*, 2018, **562**, 563–568.
- 24 J. Kim, B. X. Li, R. Y. Huang, J. X. Qiao, W. R. Ewing and D. W. C. MacMillan, *J. Am. Chem. Soc.*, 2020, **142**, 21260–21266.
- 25 F. Tekaiia, E. Yeramian and B. Dujon, *Gene*, 2002, **297**, 51–60.
- 26 S. Learte-Aymami, C. Vidal, A. Gutierrez-Gonzalez and J. L. Mascarenas, *Angew. Chem., Int. Ed. Engl.*, 2020, **59**, 9149–9154.
- 27 J. Ohata, M. B. Minus, M. E. Abernathy and Z. T. Ball, *J. Am. Chem. Soc.*, 2016, **138**, 7472–7475.
- 28 R. Waibel, R. Alberto, J. Willuda, R. Finnern, R. Schibli, A. Stichelberger, A. Egli, U. Abram, J. P. Mach, A. Pluckthun and P. A. Schubiger, *Nat. Biotechnol.*, 1999, **17**, 897–901.
- 29 A. N. Kapanidis, Y. W. Ebright and R. H. Ebright, *J. Am. Chem. Soc.*, 2001, **123**, 12123–12125.
- 30 S. Lata, M. Gavutis, R. Tampe and J. Piehler, *J. Am. Chem. Soc.*, 2006, **128**, 2365–2372.
- 31 J. Ohata, Y. Zeng, L. Segatori and Z. T. Ball, *Angew. Chem., Int. Ed. Engl.*, 2018, **57**, 4015–4019.
- 32 F. Agostini, J. S. Voller, B. Kokschi, C. G. Acevedo-Rocha, V. Kubyshekin and N. Budisa, *Angew. Chem., Int. Ed. Engl.*, 2017, **56**, 9680–9703.
- 33 S. M. Liao, Q. S. Du, J. Z. Meng, Z. W. Pang and R. B. Huang, *Chem. Cent. J.*, 2013, **7**, 44.
- 34 S. Jia, D. He and C. J. Chang, *J. Am. Chem. Soc.*, 2019, **141**, 7294–7301.
- 35 X. Chen, F. Ye, X. Luo, X. Liu, J. Zhao, S. Wang, Q. Zhou, G. Chen and P. Wang, *J. Am. Chem. Soc.*, 2019, **141**, 18230–18237.
- 36 K. Nakane, S. Sato, T. Niwa, M. Tsushima, S. Tomoshige, H. Taguchi, M. Ishikawa and H. Nakamura, *J. Am. Chem. Soc.*, 2021, **143**(20), 7726–7731.
- 37 R. C. Oslund, J. M. Kee, A. D. Couvillon, V. N. Bhatia, D. H. Perlman and T. W. Muir, *J. Am. Chem. Soc.*, 2014, **136**, 12899–12911.
- 38 J. M. Kee and T. W. Muir, *ACS Chem. Biol.*, 2012, **7**, 44–51.
- 39 Y. Takaoka, H. Tsutsumi, N. Kasagi, E. Nakata and I. Hamachi, *J. Am. Chem. Soc.*, 2006, **128**, 3273–3280.
- 40 X. Chen, F. Ye, X. Luo, X. Liu, J. Zhao, S. Wang, Q. Zhou, G. Chen and P. Wang, *J. Am. Chem. Soc.*, 2019, **141**(45), 18230–18237.
- 41 K. Peciak, E. Laurine, R. Tommasi, J. W. Choi and S. Brocchini, *Chem. Sci.*, 2019, **10**, 427–439.
- 42 Y. Cong, E. Pawlisz, P. Bryant, S. Balan, E. Laurine, R. Tommasi, R. Singh, S. Dubey, K. Peciak, M. Bird, A. Sivasankar, J. Swierkosz, M. Muroi, S. Heidelberger, M. Farys, F. Khayrzad, J. Edwards, G. Badescu, I. Hodgson, C. Heise, S. Somavarapu, J. Liddell, K. Powell, M. Zloh, J. W. Choi, A. Godwin and S. Brocchini, *Bioconjugate Chem.*, 2012, **23**, 248–263.
- 43 F. Berger, M. B. Plutschack, J. Riegger, W. Yu, S. Speicher, M. Ho, N. Frank and T. Ritter, *Nature*, 2019, **567**, 223–228.
- 44 J. Li, J. Chen, R. Sang, W.-S. Ham, M. B. Plutschack, F. Berger, S. Chhabra, A. Schnegg, C. Genicot and T. Ritter, *Nat. Chem.*, 2019, **12**, 56–62.
- 45 M. H. Aukland, M. Šiaučiusis, A. West, G. J. P. Perry and D. J. Procter, *Nat. Catal.*, 2020, **3**, 163–169.
- 46 R. Cao, X. Zhang, X. Liu, Y. Li and H. Li, *Cell Res.*, 2018, **28**, 494–496.
- 47 A. W. Wilkinson, J. Diep, S. Dai, S. Liu, Y. S. Ooi, D. Song, T. M. Li, J. R. Horton, X. Zhang, C. Liu, D. V. Trivedi, K. M. Ruppel, J. G. Vilches-Moure, K. M. Casey, J. Mak, T. Cowan, J. E. Elias, C. M. Nagamine, J. A. Spudich, X. Cheng, J. E. Carette and O. Gozani, *Nature*, 2019, **565**, 372–376.
- 48 Z. He, A. P. Pulis and D. J. Procter, *Angew. Chem., Int. Ed. Engl.*, 2019, **58**, 7813–7817.
- 49 D. Chen, Q. Feng, Y. Yang, X. M. Cai, F. Wang and S. Huang, *Chem. Sci.*, 2017, **8**, 1601–1606.
- 50 T. Yanagi, S. Otsuka, Y. Kasuga, K. Fujimoto, K. Murakami, K. Nogi, H. Yorimitsu and A. Osuka, *J. Am. Chem. Soc.*, 2016, **138**, 14582–14585.
- 51 R. Parnes, H. Reiss and D. Pappo, *J. Org. Chem.*, 2018, **83**, 723–732.
- 52 R. Parnes and D. Pappo, *Org. Lett.*, 2015, **17**, 2924–2927.
- 53 T. Le, T. Courant, J. Merad, C. Allain, P. Audebert and G. Masson, *J. Org. Chem.*, 2019, **84**, 16139–16146.
- 54 Z. Hou, D. Wang, Y. Li, R. Zhao, C. Wan, Y. Ma, C. Lian, F. Yin and Z. Li, *J. Org. Chem.*, 2020, **85**, 1698–1705.
- 55 D. Wang, M. Yu, N. Liu, C. Lian, Z. Hou, R. Wang, R. Zhao, W. Li, Y. Jiang, X. Shi, S. Li, F. Yin and Z. Li, *Chem. Sci.*, 2019, **10**, 4966–4972.
- 56 X. Shi, R. Zhao, Y. Jiang, H. Zhao, Y. Tian, Y. Jiang, J. Li, W. Qin, F. Yin and Z. Li, *Chem. Sci.*, 2018, **9**, 3227–3232.
- 57 R. Parnes, S. Narute and D. Pappo, *Org. Lett.*, 2014, **16**, 5922–5925.
- 58 M. Lanzì, J. Merad, D. V. Boyarskaya, G. Maestri, C. Allain and G. Masson, *Org. Lett.*, 2018, **20**, 5247–5250.
- 59 S. Tyanova, T. Temu and J. Cox, *Nat. Protoc.*, 2016, **11**, 2301–2319.
- 60 C. Ramallo Guevara, O. Philipp, A. Hamann, A. Werner, H. D. Osiewacz, S. Rexroth, M. Rogner and A. Poetsch, *Mol. Cell. Proteomics*, 2016, **15**, 1692–1709.
- 61 C. Lebee, M. Languet, C. Allain and G. Masson, *Org. Lett.*, 2016, **18**, 1478–1481.
- 62 L. Jarrige, G. Levitre and G. Masson, *J. Org. Chem.*, 2016, **81**, 7230–7236.
- 63 A. E. Speers and B. F. Cravatt, *J. Am. Chem. Soc.*, 2005, **127**, 10018–10019.





- 64 M. S. U. Rasheed, S. Tripathi, S. Mishra and M. P. Singh, *Mol. Neurobiol.*, 2017, **54**, 4738–4755.
- 65 P. S. Steeg, D. Palmieri, T. Ouatas and M. Salerno, *Cancer Lett.*, 2003, **190**, 1–12.
- 66 B. Kim and K. J. Lee, *Exp. Mol. Med.*, 2021, **53**, 346–357.
- 67 A. Gutteridge and J. M. Thornton, *Trends Biochem. Sci.*, 2005, **30**, 622–629.
- 68 I. Dokmanic, M. Sikic and S. Tomic, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2008, **64**, 257–263.
- 69 M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin and G. Sherlock, *Nat. Genet.*, 2000, **25**, 25–29.

