

Spatiotemporal and direct capturing global substrates of lysine-modifying enzymes in living cells

Received: 13 December 2023

Accepted: 4 February 2024

Published online: 17 February 2024

 Check for updates

Hao Hu^{1,7}, Wei Hu^{1,7}, An-Di Guo^{1,7}, Linhui Zhai^{1,2,7}, Song Ma¹, Hui-Jun Nie¹, Bin-Shan Zhou¹, Tianxian Liu¹, Xinglong Jia¹, Xing Liu³, Xuebiao Yao³, Minjia Tan^{1,4,5} ✉ & Xiao-Hua Chen^{1,4,6} ✉

Protein-modifying enzymes regulate the dynamics of myriad post-translational modification (PTM) substrates. Precise characterization of enzyme-substrate associations is essential for the molecular basis of cellular function and phenotype. Methods for direct capturing global substrates of protein-modifying enzymes in living cells are with many challenges, and yet largely unexplored. Here, we report a strategy to directly capture substrates of lysine-modifying enzymes via PTM-acceptor residue crosslinking in living cells, enabling global profiling of substrates of PTM-enzymes and validation of PTM-sites in a straightforward manner. By integrating enzymatic PTM-mechanisms, and genetically encoding residue-selective photo-crosslinker into PTM-enzymes, our strategy expands the substrate profiles of both bacterial and mammalian lysine acylation enzymes, including bacterial lysine acylases PatZ, YiaC, LplA, TmcA, and YjaB, as well as mammalian acyltransferases GCN5 and Tip60, leading to discovery of distinct yet functionally important substrates and acylation sites. The concept of direct capturing substrates of PTM-enzymes via residue crosslinking may extend to the other types of amino acid residues beyond lysine, which has the potential to facilitate the investigation of diverse types of PTMs and substrate-enzyme interactive proteomics.

Post-translational modifications (PTMs) are covalent chemical alterations of proteins after ribosome biosynthesis, which are typically mediated by enzymes (Fig. 1a). Currently, there are more than 500 types of PTMs that extensively modulate protein functions¹. Many PTMs, such as acylation, phosphorylation, and ubiquitination, are reversibly controlled by protein-modifying enzymes in a dynamic manner at corresponding localization and time,

orchestrating complex cellular processes². Aberrant PTM states caused by dysfunction of regulatory enzymes are broadly involved in the development of various complex diseases³. Therefore, precise characterization of enzyme-substrate associations is essential for functional annotation of PTMs under physiological and pathological conditions⁴. Indeed, the PTM substrate-centric strategies with the aim to identify corresponding protein-modifying enzymes, such as

¹State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. ²Translational Research Institute of Brain and Brain-Like Intelligence, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai 200434, China. ³MOE Key Laboratory for Cellular Dynamics and Hefei National Center for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China. ⁴University of Chinese Academy of Sciences, Beijing 100049, China. ⁵Zhongshan Institute for Drug Discovery, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Zhongshan Guangdong 528400, China. ⁶School of Pharmaceutical Science and Technology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China. ⁷These authors contributed equally: Hao Hu, Wei Hu, An-Di Guo, Linhui Zhai. ✉e-mail: mjtan@simmm.ac.cn; xhchen@simmm.ac.cn

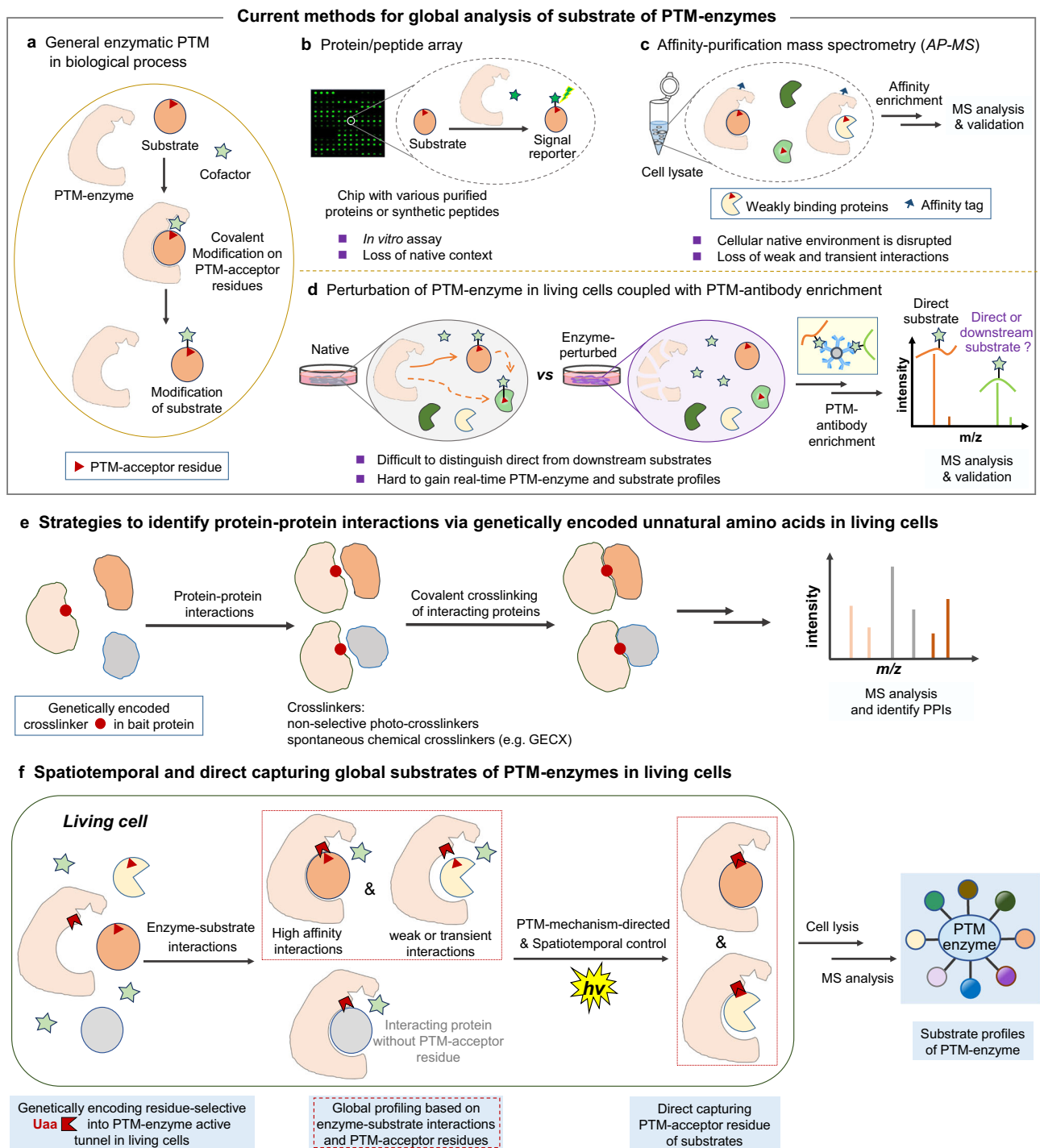


Fig. 1 | The development of post-translational modification (PTM)-mechanism-directed covalent capturing strategy via PTM-acceptor residue in living cells. **a** General enzymatic PTM in biological processes. **b** Protein or peptide array for validation of substrate modification via PTM-acceptor residue in the presence of enzymes *in vitro*. **c** Affinity purification coupled with MS analysis for validation of substrate-enzyme interactions. **d** Pharmacological or genetic perturbation of cellular enzyme activity coupled with PTM-antibody enrichment-based proteomics

e Strategies to identify protein-protein interactions (PPIs) via genetically encoded non-selective photo-crosslinkers or spontaneous chemical crosslinkers (e.g., GECX) in living cells. **f** Spatiotemporal and direct capturing global substrates of post-translational modification enzyme (PTM-enzyme) via PTM-acceptor residue of substrates, through genetically encoding residue-selective unnatural amino acid (Uaa) photo-crosslinker in living cells, offering substrate profiles of PTM-enzyme from MS analysis.

chemical proteomics based on PTM-decorated substrate surrogates^{5,6}, protein microarray (or peptide libraries)⁷, and tethered catalytic complex⁸, have successfully dissected various enzyme-substrate associations. On the other hand, enzyme-centric methods enable deciphering substrate profiles of PTM-enzymes to

determine enzymatic activities and biological functions⁴. Since protein-modifying enzymes usually control the dynamics for myriad PTM substrates, global profiling of substrates could dissect the complicated and elusive enzyme-substrate interactions, and further illuminate the landscape of complex regulatory networks^{2,4}. Thus,

the development of such enzyme-centric methods is particularly significant and highly desired.

Current enzyme-centric methods for substrate profiling could be categorized into those conducted *in vitro* and within cells. *In vitro* methods usually screen against protein or peptide microarrays (Fig. 1b) in an unbiased and high-throughput manner^{7,9}, but the cellular native context is eliminated. Although methods (e.g., affinity purification mass spectrometry, Fig. 1c)^{2,4,10} performed on cell lysate retain certain complex biological environments, the cellular native compartmentalization is disrupted, and they still suffer from the loss of weak and transient interactions, thus, resulting in reduced spatial resolution and accuracy^{2,10}. In contrast, approaches conducted in cells allow dissection of the enzyme-substrate relationships in a native milieu. Genetic perturbation of cellular enzyme activity coupled with PTM-antibody enrichment-based proteomics is the most widely used strategy (Fig. 1d)^{11,12}. However, such genetic perturbation not only impacts the direct substrates of the enzyme but also those downstream/bypassing events. Indeed, protein PTMs often modulate protein (e.g., enzymes) activity which could, in turn, reshape the PTM states of downstream proteins, such as phosphorylation cascade¹³. An obvious drawback of this approach is difficult to distinguish the direct substrates of the enzyme from those indirect ones¹². Moreover, what the genetic perturbation-based PTM-proteomic profiling experiment measures is the change of PTM level accumulated over a time course (e.g., from genetic perturbation of PTM-enzyme to cell harvest for PTM-profiling) instead of an instantaneous “snapshot” of substrate-enzyme interaction profiles at given time points. Thus, it is almost impossible to identify the real-time interacting substrates for the PTM-enzyme. Therefore, the genetic perturbation approach makes it difficult to distinguish the direct substrates from the indirectly altered PTM-proteins, and is not fit for the real-time PTM-enzyme and substrate interaction study.

In principle, higher decoding precision could be achieved by methods that directly capture enzyme-substrate interactions in living cells, conferring irreversibly covalent enzyme-substrate complexes for subsequent stringent purification, enrichment, and proteomic analysis^{14–16}. In fact, within highly dynamic and complicated native cellular contexts, a major challenge in developing a direct substrate-capturing strategy for post-translational modification enzymes is to capture bona fide substrates, including weak and transient interacting substrates, but not the non-substrate binding proteins. Indeed, the genetic code expansion technology to site-specifically incorporate unnatural amino acids (Uaas) renders minimal structural perturbation on bait proteins (Fig. 1e)^{17–22}, enabling the formation of covalent linkages to irreversibly crosslink protein complexes^{23–34}. Recently, an important advancement of the genetically encoded chemical crosslinking (GECX) method by Wang’s group, is an elegant strategy and exhibited excellent abilities to covalently capture the protein-protein interactions in living cells^{27,32–34}, while GECX method as yet is mainly focused on the investigation of the interacting proteins (e.g., interacting partners of non-enzyme or enzyme proteins). Thus, the potential and capability of PTM mechanism-guided direct capturing global substrates of post-translational modification enzymes via PTM-acceptor residues presents an exciting prospect yet largely unexplored.

Herein, we present a method to directly capture substrates of lysine-modifying enzymes via PTM-acceptor residue crosslinking chemistry in living cells, enabling global profiling of substrates of PTM-enzymes and their modification sites in a straightforward manner (Fig. 1f). This method covalently captures PTM-enzyme and interacting substrates in spatiotemporal resolution, with high sensitivity and accuracy. Harnessing these advantages, our developed strategy expands the substrate profiles of lysine acylation enzymes, such as the bacterial lysine acylases PatZ, YiaC, LplA, TmcA, YjaB, and the mammalian acyltransferases GCN5 and Tip60, leading to discovery of more

potential substrates and modification sites with biological significance. Notably, using this approach, we provide advanced insights into the acetylation on K131 of AcrA involved in the AcrA-TolC multidrug efflux pump assembly and the potential regulatory role in antibiotic resistance. Our work demonstrates that direct capture and validation of substrates of lysine-modifying enzymes via specific PTM-acceptor residue crosslinking chemistry in living cells (Fig. 1f) is a feasible and innovative paradigm for even more types of PTM research.

Results

Design of PTM mechanism-directed substrate capturing strategy

In general, the enzymatic PTMs are covalent alterations of the acceptor residue of substrates, which are carried out via enzyme-substrate complex formation and catalytic activities of the specific functional sites of the enzyme (Figs. 1a, 2a). Inspired by the widespread success of crosslinking chemistry for capturing protein and interacting partner (or ligand) interactions, as well as exemplified by the power of GECX method to capture protein-protein interactions via diverse residue-reactivities in living cells^{23–34}, we recently questioned whether it might be possible to combine the mechanism-based covalent process of a PTM with crosslinking chemistry, with an aim to directly capture enzyme-substrate interactions in living cells. More specifically, we reasoned that a genetically encoded residue-selective crosslinker in the active site tunnel of PTM enzyme should crosslink with the PTM-acceptor residues of the substrate via stable covalent linkage formation (Fig. 1f). Thus, the crosslinking would allow direct capturing global substrates of the subjected protein-modifying enzyme in living cells. However, because of the great complexity and dynamic nature of PTMs in living systems, several challenges are associated with the development of methods to directly capture substrates of PTM-enzymes. Most notably, it is elusive and difficult to discriminate the direct substrates from non-substrate binding proteins, due to the complexity of the multiprotein interacting networks^{2,27}. In addition, the modifications on substrate proteins are dynamic and reversible, thus, a temporal resolution of capturing enzyme-substrate interactions at a certain stage is indispensable. Furthermore, as enzyme-substrate interactions are weak and transient in many cases, capturing these unstable and low-stoichiometry enzyme-substrate complexes requires high crosslinking efficiency from PTM-enzymes with PTM-acceptor residues of substrates^{10,27}.

To fulfill these stringent demands, we reasoned that several characteristics are required for a generalizable direct capturing strategy: (i) a crosslinker with rigorous selectivity for the PTM-acceptor residue and located in the enzyme catalytic site tunnel to enable chemical (i.e., PTM-acceptor residue) and spatial dual-specificity in crosslinking, thus rendering high accuracy; (ii) temporal-control and fast-on activation of crosslinking that permits time-resolved dissection of the substrate interactome; (iii) relatively high reactivity and long reaction window of the activated crosslinker, allowing to capture sufficient amount of weak and transient enzyme-substrate complexes to satisfy the detection limit of mass spectrometry (MS) analysis; (iv) global substrate profiling criteria based on identification of captured proteins via crosslinking mass spectrometry (XL-MS) with stringent data filtration criteria built on properly designed control to further rule out non-specific background signals, which would help in identification accuracy.

As a proof-of-concept and general entry point, we herein showcase our strategy to directly capture the global substrates of several lysine-modifying enzymes. Lysine residues of proteins are among the most frequently modified amino acids with more than 40 discrete types of modification, such as acetylation, diverse short-chain acylations, lipidations, lipoylation, ubiquitination, and so on (Fig. 2a)¹. These lysine acylations regulate numerous protein functions (e.g., stability, enzymatic activity, cellular translocation, interactions, etc.) and almost

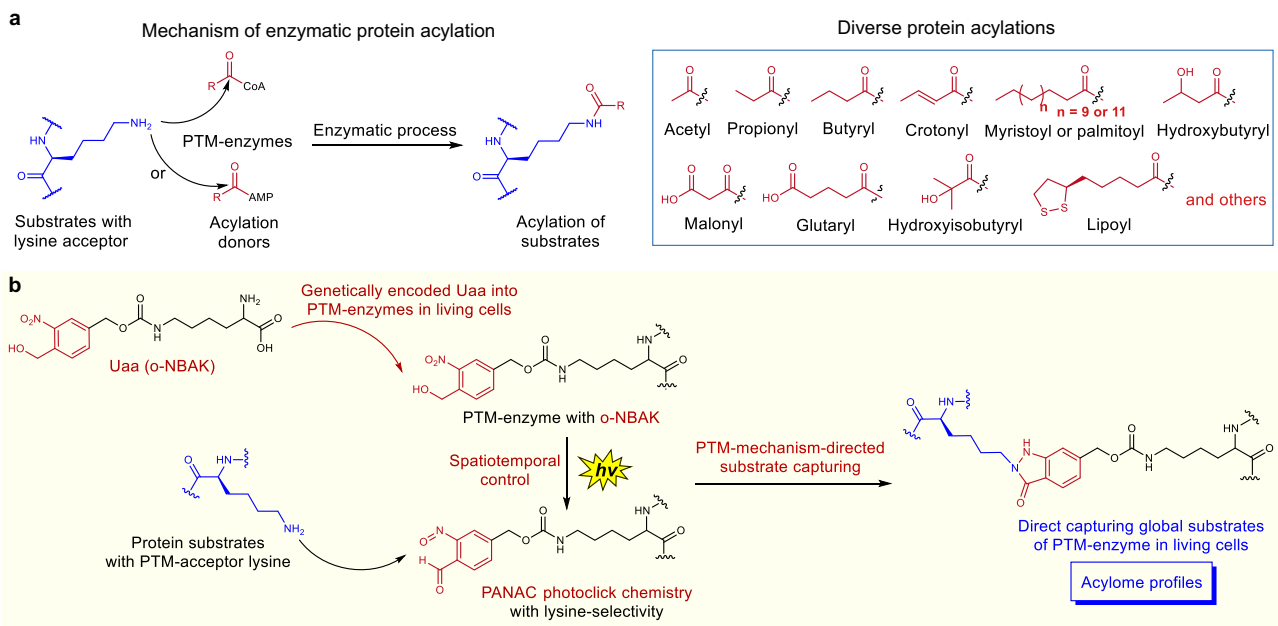


Fig. 2 | PTM mechanism-directed covalent capturing of substrates. **a** Mechanism of enzymatic protein acylation and diverse protein acylations. **b** Direct capturing substrates with PTM-acceptor lysine via spatiotemporal lysine-selective PANAC photoclick chemistry. CoA coenzyme A, AMP adenosine monophosphate.

are involved in many biological processes^{2,4,35}. In addition, several studies showed that lysine-modifying enzymes have promiscuous acylation activity other than certain specific modification^{36,37}. More importantly, short-chain acylations of bacterial virulence regulators can modulate host immunity and susceptibility to infection^{38,39}. However, the substrate profiles of acylation enzymes have been elusive^{36,40}. Thus, connecting lysine-modifying enzymes to their substrates (acylome) is crucial for understanding their regulatory mechanisms and functions^{41,42}.

In this context, the initial step of our study was to identify a genetically encoded crosslinker that was suitable for directly capturing acylation substrates of lysine-modifying enzymes. We reasoned that the advantages of the light-induced primary amines and *o*-nitrobenzyl alcohols cyclization (PANAC photoclick) are inherently linked to the direct capturing chemistry (Fig. 2b)⁴³. The photo-generated aryl-nitroso intermediates feature reliable chemo-selectivity with primary amines, high reactivity, good biocompatibility under physiological conditions, and a long half-life (with hours in buffer)⁴³. Based on these insights into the PANAC photoclick chemistry, we employed the *o*-nitrobenzyl alcohol-derived lysine (*o*-NBAK), an unnatural amino acid (Uaa) as a photo-crosslinker, which is able to selectively and efficiently react with lysine upon light-activation^{44,45}. Thus, genetically incorporating *o*-NBAK in the (or near) catalytic pocket of enzymes would capture the proximal PTM-acceptor lysine on the interacting substrates with high sensitivity and accuracy, in a spatiotemporally controlled manner (Fig. 2b).

Proteome-wide characterizing direct substrates of protein lysine acetyltransferase PatZ

We initially tested the feasibility of the designed strategy on global substrate profiling for PatZ enzyme in *E. coli*. PatZ and its homologs are the major lysine acetyltransferases in bacteria which regulate diverse biological processes, including growth, mobility, stress response, etc, but the corresponding acylation substrates are still elusive⁴⁶. In *E. coli*, most of the protein acetylation events were reported to be non-enzymatically mediated by acetyl phosphate, thus, genetic perturbation of PatZ resulted in negligible alteration on the level of protein acetylation⁴⁷. In addition, the previous studies from us and other groups showed PatZ and its homologs exhibit promiscuous acylation

activity other than acetylation^{36,48,49}. Therefore, it is challenging to assign acylation substrates of PatZ by genetic perturbation of enzymes coupled with PTM-antibody enrichment and MS analysis approach.

To explore the appropriate sites for genetic incorporation of the *o*-NBAK, we homology-modeled the structure of the catalytic domain (GNAT-domain) of PatZ since no crystal structure is currently available. The crystal structure of *S*PatA (PDB ID: 4NXY) was selected as the template because it has a known binding mode with substrate⁵⁰, which provided a hint of how PatZ would interact with its substrates (Fig. 3a). We initially selected three sites (I800, F810, L813) in PatZ for incorporation of the *o*-NBAK Uaa (Fig. 3b). I800 and F810 are located directly adjacent to a conserved residue E809 in the catalytic pocket⁵¹, but with distinct structural environments: F810 is deeply buried in the protein fold while I800 is exposed to solvent, and L813 is half-buried in the catalytic site (Fig. 3b). The PatZ mutants with site-specifically incorporated *o*-NBAK Uaa were successfully prepared according to our previous method⁴⁴, but the PatZ-F810-*o*-NBAK mutant showed limited expression of soluble protein (Supplementary Fig. 1a), and thus we chose those PatZ mutants incorporated with *o*-NBAK at sites I800 and L813.

After reaching the stationary phase, *E. coli* cells expressing PatZ-F810-*o*-NBAK protein or PatZ-L813-*o*-NBAK protein were subjected to 365 nm UV light activation for 20 min (+UV), and samples without UV light (-UV) activation were used as controls. After cell lysis and affinity purification, the eluent was further resolved by SDS-PAGE. UV-irradiated sample generated several higher molecular weight bands for PatZ-L813-*o*-NBAK mutant group, which indicated the efficient crosslinking of interacting proteins (Fig. 3c, d). In contrast, the other PatZ-I800-*o*-NBAK mutant distal to the catalytic site or unfavorably orientated, showed almost no difference between UV-irradiation samples and controls (Supplementary Fig. 1b). Moreover, we chose another distal site S816 for comparison (Fig. 3b). S816 is solvent-exposed and more far away from the active site tunnel. Thus, the incorporated *o*-NBAK in this distant site (S816-*o*-NBAK) could serve as an important control for capturing non-substrate binding proteins or distal PPIs, rather than interacting substrates of PatZ (Fig. 3b). In identical conditions, PatZ-S816-*o*-NBAK mutant showed almost non-crosslinking of interacting proteins, which was similar to the results of PatZ-I800-*o*-NBAK mutant (Supplementary Fig. 1b). Although these two

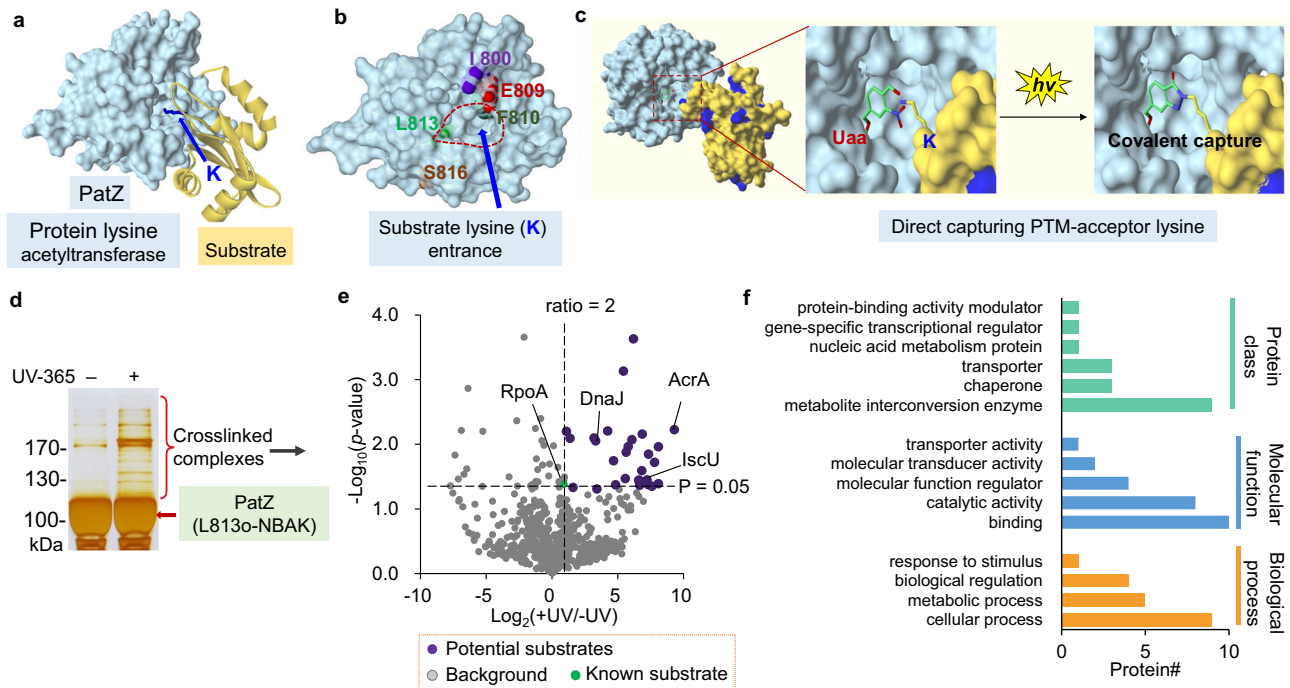


Fig. 3 | Global profiling substrates of PatZ enzyme via direct capturing strategy in living cells. **a** Schematic view of the homology-modeled catalytic domain of PatZ (shown in blue surface mode) interacting with substrate (yellow ribbon). Blue arrow indicates the substrate lysine shown in stick. **b** Structure of PatZ catalytic domain is shown in surface mode with I800, E809, F810, L813, and S816 indicated. **c** The mechanism of the light-activated primary amines and *o*-nitrobenzyl alcohols cyclization (PANAC photoclick reaction). **d** Sliver staining of samples from \pm UV-irradiation treatment and His6-tag affinity-enrichment, with red arrow indicating PatZ-L813o-NBAK. **e** Volcano plot displaying enriched proteins ($+UV/-UV$ ratio >2 ,

p value < 0.05 , purple dots) by PatZ-L813o-NBAK from biological duplicates. Statistical tests are unpaired, two-sided two-sided Student's t test. Protein identification and label-free quantification were performed with MaxQuant software (see Supplementary methods for details). The abundance of a known PatZ substrate RpoA in $+UV$ group is 1.9-fold of that in $-UV$ group (p value < 0.05). **f** Annotation of enriched proteins with regards to protein class, molecular function, and biological process according to PANTHER database (v16). Source data are provided as a Source Data file. UV ultraviolet light.

sites (I800, S816) of *o*-NBAKs are more solvent-exposed than L813o-NBAK, and expected to have more crosslinking possibility with surrounding proteins, the crosslinking efficiency mainly relied on suitable position of the genetically encoded photo-crosslinker *o*-NBAK which indicates the spatial resolution features (Fig. 3b, d, Supplementary Fig. 1b). These results demonstrate that the direct capturing efficiency largely depends on the substrate recognition and the engagement of the PTM-acceptor lysine.

Next, samples from PatZ-L813o-NBAK mutant were used for the subsequent substrate profiling. The bands with higher molecular weight than free PatZ mutant enzyme in SDS-PAGE gel with UV light ($+UV$) or without UV light exposure ($-UV$) groups were excised, respectively (Fig. 3d), which correspond to potential crosslinked enzyme-substrate assemblies. Based on the covalent crosslinking of mutant PatZ enzyme and substrates, the sequential affinity purification with washing process and SDS-PAGE separation would increase the detection sensitivity for substrates, due to the removal of high-abundance non-specific backgrounds and free PatZ mutant enzyme (Fig. 3d, PatZ-L813o-NBAK). The gel slices were then subjected to trypsin digestion, LC-MS/MS analysis and label-free quantification. A known PatZ substrate, DNA-directed RNA polymerase subunit alpha (RpoA), displayed significantly higher abundance in $+UV$ group over $-UV$ group. Proteins that were significantly enriched in $+UV$ group were considered as candidate substrates (Fig. 3e and Supplementary Data 1). In total, PatZ-L813o-NBAK mutant captured 28 candidate substrates with diverse functional classes and various biological processes (Fig. 3f). More importantly, the temporally-controlled crosslinking allows us to focus on the investigation of corresponding substrates in a given growth phase (at several time points, also see Supplementary Fig. 18). Taken together, our herein developed strategy

enabled direct capturing of PatZ-substrate interactions with tolerable perturbation (Supplementary Figs. 10–12) of the cellular milieu, and thus presents a straightforward and general method for the discovery of acylation substrates of lysine-modifying enzyme in a spatiotemporal manner.

Validation of PatZ substrates and acetylation sites

To validate the direct interaction between PatZ and candidate substrates, we performed *in vitro* crosslinking by selecting purified recombinant candidate proteins (Fig. 4a), including a reported PatZ substrate IscU⁵², a reported PatZ interacting protein DnaJ⁵³, and a distinct candidate substrate, multidrug efflux pump subunit AcrA (Supplementary Fig. 26a)⁵⁴. The purified proteins were crosslinked with PatZ-L813o-NBAK mutant under UV irradiation for 15 min on ice and resolved on SDS-PAGE. Compared with the corresponding control groups, extra bands consistent with the expected molecular weight of crosslinking were observed in UV-irradiated mixture groups (Supplementary Fig. 2). These bands were then excised and subjected to LC-MS/MS analysis for further validation of the covalent crosslinks. We identified high-confidence crosslink-spectrum matches (CSMs) for each substrate. Extensive *b*- and *y*-ions indicated that the covalent linkages exclusively formed between *o*-NBAK and the lysine residues on the candidate proteins (Fig. 4b–d and Supplementary Fig. 3), which indicated the direct and residue-specific crosslinking of the L813o-NBAK with lysine residues.

Notably, encouraged by the capability to capture K122 of substrate IscU (Fig. 4b), which had previously been implied as an enzymatic acetylation site⁵², we further investigated whether the crosslinked sites in the distinct candidate substrates (Fig. 4d) could actually be acetylated by PatZ. Therefore, we carried out validation for

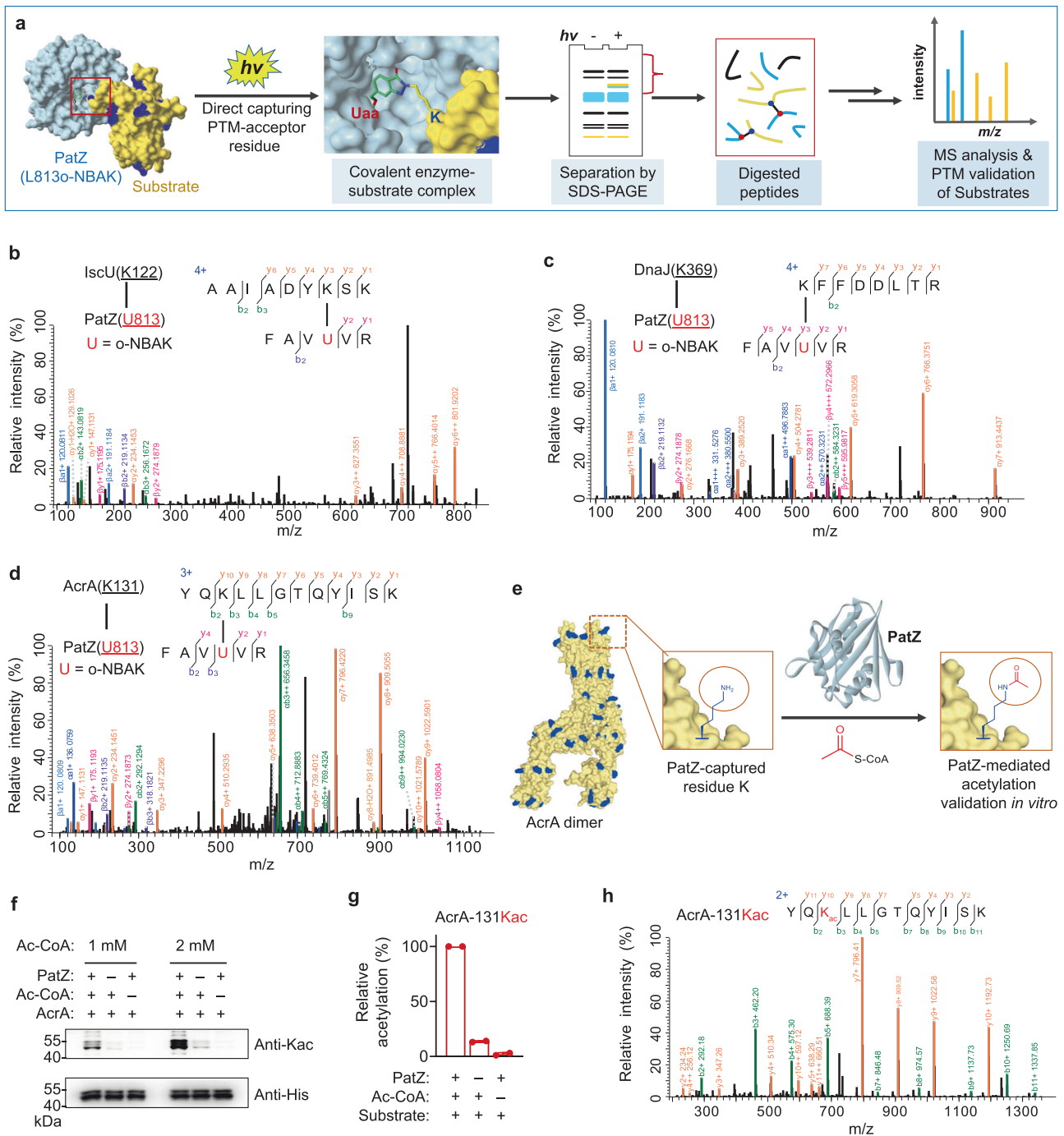


Fig. 4 | Validation of direct crosslinking of o-NBAK encoded enzyme (PatZ-L813o-NBAK) with candidate substrates, and acetylation of substrate Acra by PatZ enzyme. **a** Schematic of the workflow of light-induced direct capturing strategy for PatZ-L813o-NBAK mutant enzyme. **b–d** Annotated MS/MS spectra indicating o-NBAK of PatZ mutant crosslinks with IscU-K122 (**b**), DnaJ-K369 (**c**), AcrA-K131 (**d**), respectively, U = o-NBAK. The crosslinked peptides were identified with MeroX software (see Supplementary methods for details), and the quality of

the crosslink-spectrum matches was manually inspected. **e** Schematic of PatZ enzyme mediated acetylation confirming lysine acetylation (Kac) of Acra by PatZ. **f** Western blot analysis confirming lysine acetylation (Kac) of Acra by PatZ. **g** Relative acetylation level at Acra-131Kac under different conditions. Error bars represent standard errors of mean. $n = 2$ biologically independent experiments. **h** Annotated MS/MS spectrum of peptide bearing Acra-131Kac, and acetylated lysine residue was red-colored. Source data are provided as a Source Data file.

PatZ-mediated acetylation on Acra (Fig. 4e). Purified Acra protein was subjected to *in vitro* acetylation assay, compared with samples without PatZ or acetyl-coenzyme A (Ac-CoA) incubation as controls, and then analyzed by Western blot (Fig. 4f). Although intrinsic acetylation and Ac-CoA-mediated non-enzymatic acetylation were observed in controls, the acetylation level was significantly higher in the presence of both PatZ and Ac-CoA (Fig. 4f). In addition, the mutations of several

active site residues of PatZ (e.g. E809A, V814A and G824A)⁵¹ diminished the acetylation level to certain extents (Supplementary Fig. 4). These results confirmed the enzymatic acetylation of the substrate Acra by PatZ enzyme. After the MS analysis of the samples from the *in vitro* acetylation assay, we found K131 of Acra, which formed a covalent linkage with PatZ-L138o-NBAK in our direct substrate capturing analysis, was indeed acetylated by PatZ enzyme (Fig. 4g, h and

Supplementary Fig. 5), implicating AcrA is a substrate of PatZ in vitro. Moreover, we also identified some crosslinked peptides/sites from the cell lysate samples directly for several above-mentioned candidate substrates, such as *grpE*, *nusG*, *arcA*, *dxs*, etc. (Supplementary Figs. 13, 14). Thus, the crosslinking-MS analysis is complementary to enzymatic acetylation assays to validate the protein substrates and their modified sites. Altogether, these results clearly demonstrated that our strategy was able to directly capture enzymatic substrates via crosslinking the PTM-acceptor residues.

In vitro validation of the potential biological consequence of substrate AcrA acetylation

Protein AcrA is a previously unreported candidate substrate of PatZ enzyme, which is a component of AcrA-AcrB-AcrZ-ToIC multidrug efflux pump⁵⁴, leading to intrinsic resistance to a range of antibiotics⁵⁴. Although acetylation is a ubiquitous PTM in bacteria with different regulatory roles⁴⁶, the biological consequence of acetylation of AcrA has not been disclosed. The acetylation site K131 of AcrA belongs to one of the four co-evolved residue pairs in the interface between AcrA and ToIC (Supplementary Fig. 26a)⁵⁴. We reasoned that acetylation on K131 may interfere with its binding with ToIC. Thus, we prepared purified recombinant ToIC and wild-type AcrA (AcrA-wt), as well as genetically incorporated acetyl-lysine (Kac) into position 131 of AcrA (AcrA-K131Kac) (Supplementary Fig. 26b, c). The interactions of ToIC with AcrA-wt, and ToIC with AcrA-K131Kac were analyzed by bio-layer interferometry assay. After incubation of wild-type AcrA with ToIC, the interaction assay led to a substantial and concentration-dependent increase of signals. In contrast, minimal signal change was observed when the mutant AcrA-K131Kac was incubated with ToIC (Supplementary Fig. 26d, e), indicating the subassembly with ToIC was largely abolished by the acetylation on K131. To investigate whether AcrA-K131Kac affected *E. coli* drug resistance, we overexpressed wild-type AcrA and AcrA-K131Q mutant in the *acrA* gene-deleted *E. coli* strains for antibiotic susceptibility test (i.e., erythromycin, oxacillin, and linezolid), respectively. The K131Q mutation of AcrA neutralized its positive charge and therefore mimicked K131 acetylated AcrA. The growth inhibition assay results suggested that AcrA-K131Kac was potentially vital to the antibiotics susceptibility of *E. coli*, as AcrA-K131Q mutant exhibited severe growth inhibition in all three tested conditions (Supplementary Fig. 26f). Taken together, these results implicated a potential role of PatZ enzyme in regulating the formation of AcrA-ToIC complex via acetylation K131 of AcrA (Supplementary Fig. 26b–d), which possibly involved in the AcrA-AcrB-AcrZ-ToIC multidrug efflux pump assembly⁵⁴. These findings provide a perspective for understanding the regulatory mechanism of the multidrug efflux pump and, consequentially, the regulation of antibiotic resistance.

Application of direct capturing strategy in substrate profiling for the recently identified acetyltransferase YiaC

With proof of concept for the identification of both known and distinct substrates of PatZ, we next tested the applicability of our strategy for YiaC, an enzyme that is conserved across bacteria taxa. YiaC has been proposed as GNAT-domain lysine acetyltransferase, but no acetylation assay was performed to validate the potential substrates⁵². A recent study has shown that YiaC has specific impacts on the global level of lysine 2-hydroxyisobutyrylation in *E. coli*⁴⁰. In addition, YiaC in *S. enterica* has been characterized as N-terminal acetyltransferase⁵⁵. These studies implicate YiaC might have promiscuous acyltransferase activities, thus further investigation is required to understand its substrate repertoire and function. As there is no crystal structure of YiaC, we homology-modeled its structure (Fig. 5a). Since both YiaC and PatZ adopted standard GNAT fold⁵², they are expected to have a similar binding mode with substrate lysine. Similarly, we chose F75 as the site for genetic incorporation of o-NBAK, as F75 has spatial localization similar to L813 of PatZ, and F75 is near the E103 residue, which was

deduced to be engaged in catalytic reaction⁵². *E. coli* cells expressing YiaC-F75o-NBAK mutant (Supplementary Fig. 6a) with or without UV-irradiation, underwent SDS-PAGE separation and LC-MS/MS analysis. We obtained a total of 34 candidate substrates of YiaC (Fig. 5b, Supplementary Data 2). Of all these candidate substrates of YiaC, only 3 proteins were also identified as potential substrates of PatZ (Fig. 5c), which suggested minimal target redundancy between the two GNAT-domain acetyltransferases. Functional annotation analysis revealed these proteins were implicated in various biological processes (Fig. 5d).

We next validated YiaC-dependent acetylation activity for two selected candidate substrates, 50 S ribosomal protein L14 (RPL14) and FKBP-type 22 kDa peptidyl-prolyl cis-trans isomerase (FKBP22), by in vitro acetylation assay and Western blot (Supplementary Fig. 6b). The exact modification sites (K114 on RPL14, K86 and K130 on FKBP22) were confirmed by LC-MS/MS analyses (Fig. 5e–g and Supplementary Fig. 6c, d). Intriguingly, K114 of RPL14 is highly conserved and critical for binding with the ribosomal silencing factor RsfS⁵⁶. Thus, we inferred that acetylation of K114 would likely affect this interaction, implicating a role of YiaC in regulating the translation process. Altogether, with our developed direct global capturing strategy, we revealed a collection of potential acylation substrates of YiaC and the related biological processes. Moreover, our results verified the protein lysine acetylation activity of YiaC enzyme in biochemical experiments.

Profiling substrates of bacterial acylases LplA, TmcA, YjaB, and mammalian acyltransferases Tip60, and GCN5

To further evaluate our developed strategy for directly capturing substrates of another acylation type of PTM-enzymes, we selected the lysine-modifying enzyme lipoate-protein ligase A (LplA, Fig. 6a) in *E. coli*⁵⁷. Protein lipoylation (Fig. 6b), as well as its regulatory enzymes and substrates, are all evolutionarily conserved and play essential roles in energy metabolism and amino acid catabolism; however, there are only a limited number of lipoylated substrates previously reported⁵⁷. It remains to know the global lipoylated substrate proteins and there was no validation for lipoylation performed on the identified potential substrates. Thus, we questioned whether our direct capturing strategy could expand the inventory of potential substrates for LplA enzyme.

Based on the crystal structure of LplA (PDB ID: 3A7A), we genetically incorporated o-NBAK into site A138, which is located in the channel for substrate lysine accessing to the lipoic acid-AMP (LA-AMP) intermediate (Fig. 6a and Supplementary Fig. 7a). As a proof of the enzyme-substrate crosslinking feasibility, we performed in vitro crosslinking between o-NBAK encoded enzyme LplA-A138o-NBAK and apoH protein of the glycine cleavage system in which K65 is a reported lipoylation site by LplA⁵⁷. Compared to controls without apoH protein or without UV-irradiation (–UV), an extra band in the UV-irradiated mixture suggested the success of crosslinking (Supplementary Fig. 7b). The following LC-MS/MS analysis of proteins in this band unambiguously confirmed the crosslink between LplA-A138o-NBAK and the expected K65 of apoH protein (Fig. 6g).

Having confirmed the ability to capture the known substrate apoH, we next turned to a global profiling of substrates of LplA enzyme. *E. coli* cells expressing His6-tagged LplA-A138o-NBAK mutant were submitted to in vivo crosslinking, subsequent lysis, and affinity purification. The UV-irradiated sample showed multiple extra bands in SDS-PAGE gel, compared to the control group, indicating efficient crosslinking (Fig. 6c). The proteins with higher molecular weight than mutant enzyme LplA-A138o-NBAK (Fig. 6c), in both groups were submitted to LC-MS/MS analysis. We identified 121 proteins significantly enriched in +UV group over –UV group, including the known substrate apoH (Fig. 6d, Supplementary Data 3A, \log_2 ratio = 8.68, p value = 0.025), highlighting the suitability of our approach in the identification of substrates. Of note, the apoH and LplA binding is low affinity ($K_m = 1.2 \mu\text{M}$)⁵⁸, indicating our strategy could capture weak

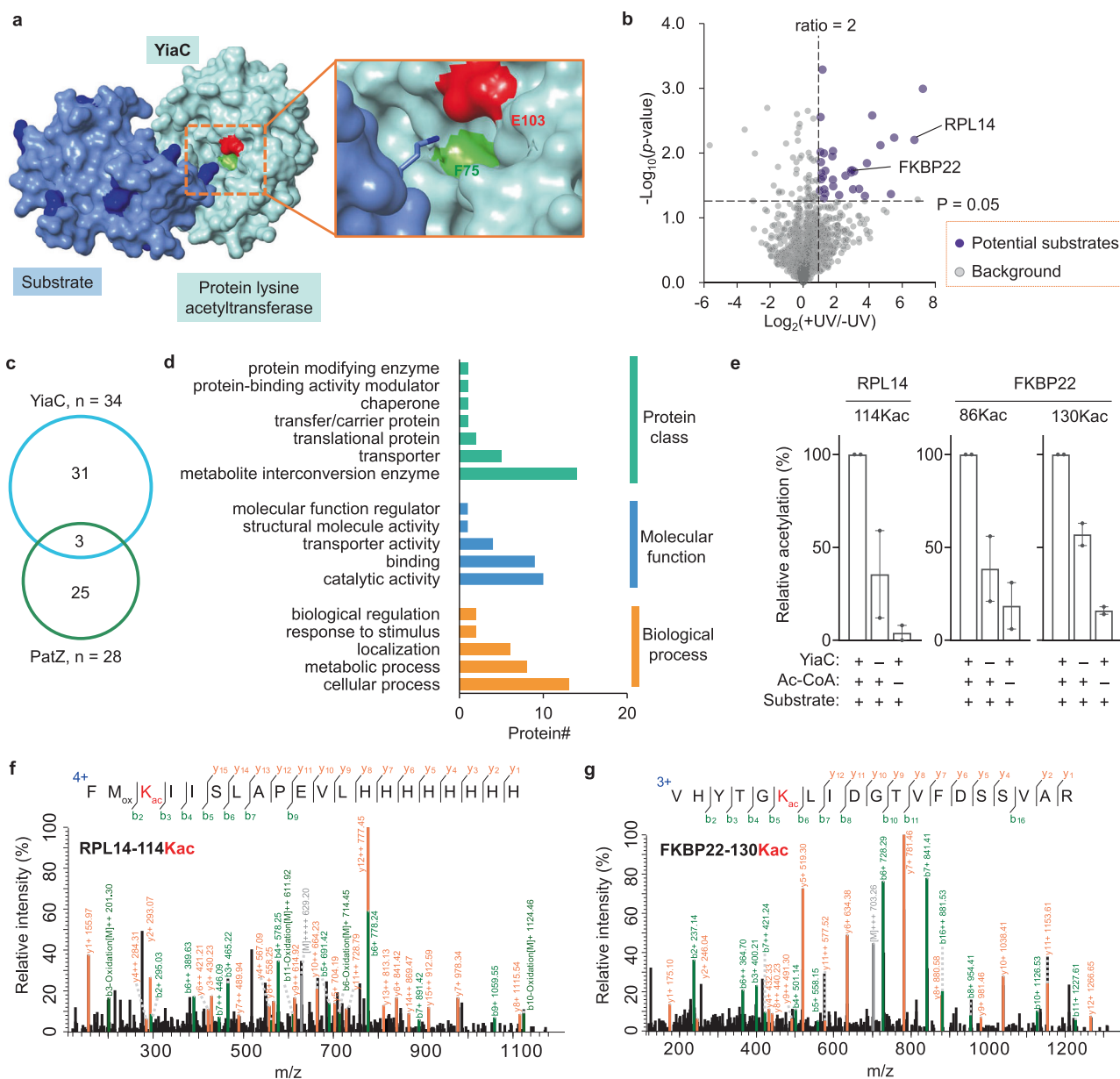


Fig. 5 | Global profiling of the substrates of YiaC enzyme. **a** Schematic view of homology-modeled YiaC (cyan) interacting with substrate (blue). The surfaces of residues F75 and E103 of YiaC enzyme are rendered in green and red, respectively. **b** Volcano plot displaying enriched proteins (+UV/-UV ratio > 2, p value < 0.05, purple dots) by YiaC-F75o-NBAK from biological triplicates. Statistical tests are unpaired, two-sided two-sided Student's t test. **c** Overlap between candidate substrates enriched by YiaC-F75o-NBAK and PatZ-L813o-NBAK. **d** Annotation of

enriched proteins with regards to protein class, molecular function, and biological process according to PANTHER database (v16). **e** Relative acetylation level at K114 of RPL14, K86 and K130 of FKBP22 under different conditions. Error bars represent standard errors of mean. $n = 2$ biologically independent experiments. **f**, **g** Annotated MS/MS spectra of peptides bearing RPL14-114Kac (**f**) and FKBP22-130Kac (**g**). Acetylated lysine (Kac) residues are red-colored. M_{ox} Oxidized methionine. Source data are provided as a Source Data file.

substrate and PTM-enzyme interactions. It is noteworthy that <10% of the proteins enriched by LplA-A138o-NBAK were identified as candidate substrates of acetyltransferases YiaC and PatZ (Fig. 6e), suggesting the proteins were enriched based on specific enzyme-substrate recognition rather than non-specific binding. Gene ontology enrichment analysis for these enriched proteins revealed that various biological processes were significantly enriched, such as RNA metabolism, ribosome biogenesis, translation, glucose, and pyruvate metabolism, etc. (Fig. 6f, Supplementary Data 3B).

We then performed in vitro crosslinking validation for a candidate substrate of LplA enzyme (Supplementary Fig. 27a), the transcriptional regulatory protein PhoP, which plays central roles in the control of

transcription of virulence genes in several gram-negative bacterial species and is considered an intriguing drug target⁵⁹. Compared to controls, additional bands with higher molecular weight showed up in the UV-irradiated LplA-A138o-NBAK and candidate mixture (Supplementary Fig. 7c). The following tandem MS analysis confirmed cross-linking between LplA and PhoP (Fig. 6h). Then, in vitro enzymatic assay was conducted to validate the lipoylation by wild type LplA. We identified 3 lysine lipoylation residues of PhoP in the presence of LplA enzyme and LA-ATP (Fig. 6i, and Supplementary Fig. 27b, c) based on the high-confidence peptide-MS spectrum matches and unambiguous site localization. In contrast, no lipoylation was identified in the control groups that lacked LplA or LA-ATP. In addition, we further confirmed

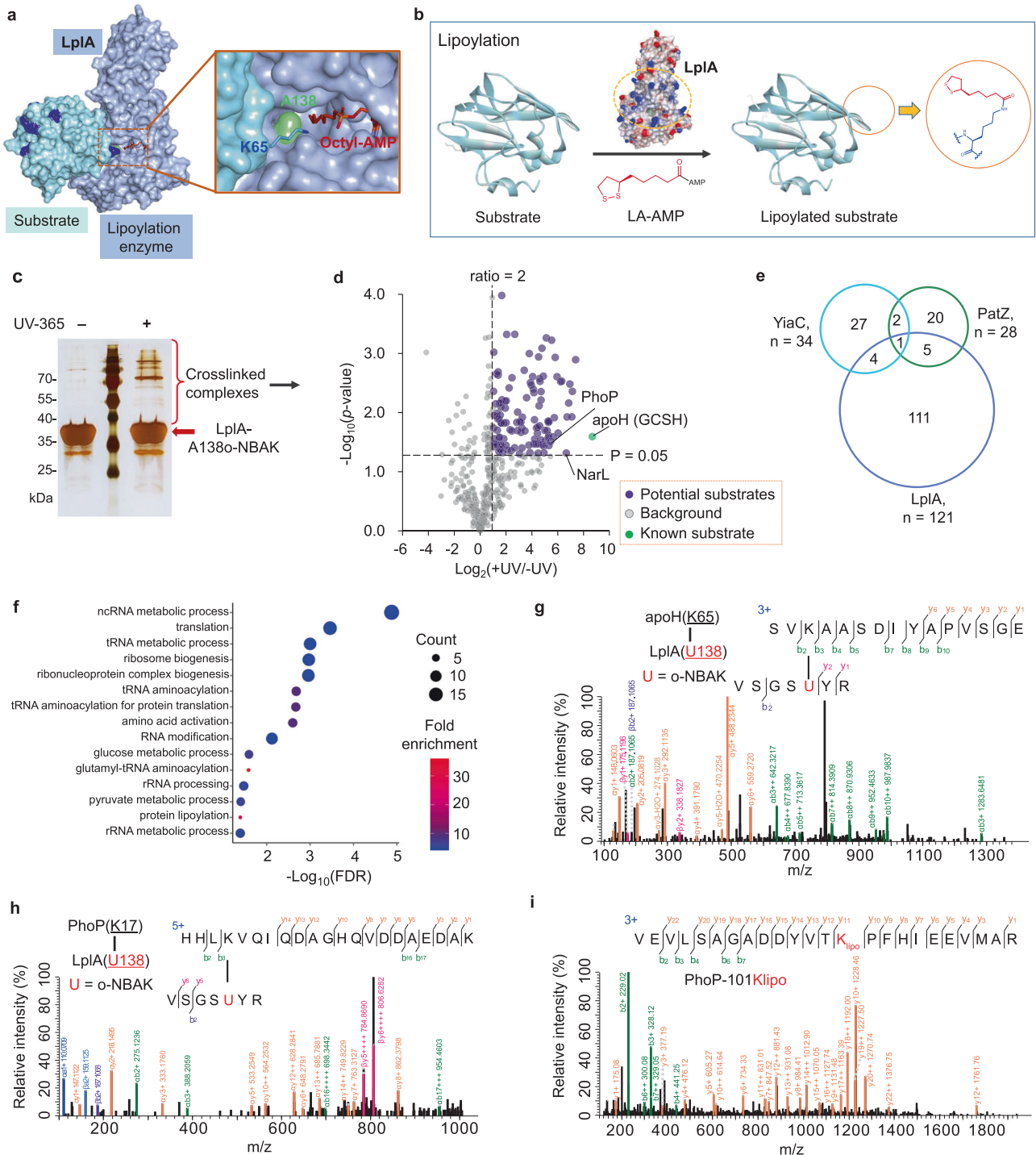


Fig. 6 | Global profiling of substrates of LplA enzyme and validation of substrate lipoylation by LplA enzyme. **a** Schematic of LplA (blue) interacting with its substrate protein (apoH protein, or called GCSH, cyan) (modified from structure, PDB ID: 3A7A). A138 (green) and the octyl-AMP (a stable surrogate for lipico acid-AMP conjugate, LA-AMP, red stick) are shown, respectively. **b** Schematic of LplA-mediated lipoylation. The surface of LplA is colored by atom charge, with red and blue indicating positively and negatively charged surfaces. The brown dashed oval indicates binding area with substrates. **c** Sliver staining of samples from \pm UV-irradiation treatment and His6-tag affinity-enrichment, red arrow indicating LplA-A138o-NBAK, and the higher molecular weight bands indicating crosslinked complexes. **d** Volcano plot displaying enriched proteins ($+UV/-UV$ ratio >2 ,

p value < 0.05 , purple and green dots) by LplA-A138o-NBAK from four replicates. Statistical tests are unpaired, two-sided two-sided Student's t test. **e** Venn diagram showing overlap of proteins enriched by LplA-A138o-NBAK, YiaC-F75o-NBAK, and PatZ-L813o-NBAK. **f** Gene ontology – biological process (GO-BP) enrichment analysis of enriched proteins. Significantly enriched (false discovery rate, FDR < 0.05) GO-BP terms of 15 highest fold enrichment value (color-coded) are shown, with size of the circle indicating number of proteins in the GO-BP term. **g, h** Annotated MS/MS spectra indicating o-NBAK of LplA mutant crosslinks with apoH-K65 (**g**) and PhoP-K17 (**h**) respectively, U = o-NBAK. **i** Annotated MS/MS spectrum of peptide bearing PhoP-101Klipo. lipo lipoylation. Source data are provided as a Source Data file.

another potential substrate, the nitrate/nitrite response regulator protein (NarL) of LplA enzyme. We identified 2 lysine lipoylation residues of NarL (Supplementary Fig. 27d, e), which suggested that NarL was the substrate of the LplA enzyme and was readily lipoylated by LplA. In addition, from the crosslinking-MS analysis, we could identify crosslinked peptides/sites from the cell lysate samples directly for some above-mentioned candidate substrates, such as galR, gapA, hisS, narH, rmlL, rpmB, etc. (Supplementary Figs. 13, 14), which suggested that our method was capable to capture PTM-acceptor lysine of substrates for other acylation, other than acetylation. Taken together, these results indicated that our strategy was able to capture the potential substrates of another type of acylation enzymes, such as lipoylation enzyme LplA.

Finally, to demonstrate more acyltransferases incorporating o-NBAK at the catalytic site tunnel to capture substrates, we expanded our strategy to other bacterial lysine acyltransferases TmcA, and YjaB, as well as mammalian acyltransferases Tip60 and GCN5. The TmcA enzyme, previously known as a RNA acetyltransferase, has recently been reported as a lysine 2-hydroxyisobutyryltransferase in the regulation of transcription and acid resistance of *E. coli*, but few substrates have been characterized⁴⁰. Our o-NBAK-incorporated TmcA captured dozens of candidate substrates (Supplementary Figs. 22, 23, Supplementary Data 4) in *E. coli* cells, and some of them were validated by in vitro acylation assay using wild-type enzyme (e.g., 2-hydroxyisobutyrylation of MTOX and FolC proteins, Supplementary Fig. 23). Indeed, we also could identify some crosslinked peptides/sites from the cell lysate samples directly, including lysines of substrates such as acnB-K571, aspS-K123, gatD-K341, hsdR-K669, leuS-K34, pepD-K59, etc. (Supplementary Figs. 13, 14). Another acetyltransferase YjaB was reported to regulate the acetylation levels for a number of proteins, but no direct substrate has been reported to date⁵². Again, our direct substrate capturing method could capture a variety of candidate substrates (Supplementary Fig. 22b, 24, Supplementary Data 5) in *E. coli* cells, and some of the candidate substrates were confirmed by in vitro enzymatic assay (e.g., EF-Ts, Rho and Rob proteins, Supplementary Fig. 24). We further demonstrated this method in mammalian acyltransferase substrate identification. The mammalian acyltransferase Tip60 enzyme plays important roles in many processes, such as chromosome segregation and DNA damage repair^{60,61}. In living human 293 F cells, using our direct substrate capturing strategy, we identified 89 candidate substrates (Supplementary Fig. 19, 20, Supplementary Data 6) for Tip60, including several previously reported substrates, such as H4 protein, H2A protein, and Ran protein. Another example is the mammalian acyltransferase GCN5, which is crucial for chromatin function regulation⁶¹. Our strategy has uncovered substrate profiles of GCN5, providing dozens of candidate substrates (Supplementary Fig. 19, 21, Supplementary Data 7), as well as previously reported substrates, such as H4 protein. Altogether, these results indicated our direct substrate-capturing strategy could successfully be applied in different acyltransferases in both bacteria and mammalian cells.

Discussion

By integrating enzymatic PTM mechanisms, residue-selective cross-linking chemistry, and tolerable perturbation to the cellular enzymes, we have established a direct capturing strategy for global substrate profiling of lysine-modifying enzymes in living cells, conferring irreversibly covalent enzyme-substrate complexes for purification and enrichment tolerating stringent processing, subsequent quantitative proteomic analysis. Based on our developed strategy, we have extended the inventory of acylation substrate candidates of PTM-enzymes, such as protein lysine acylases PatZ (Fig. 3a), YiaC (Fig. 5a), LplA (Fig. 6a), TmcA (Supplementary Fig. 22a) and YjaB (Supplementary Fig. 22b) in *E. coli*, as well as mammalian acyltransferases Tip60 (Supplementary Fig. 19a) and GCN5 (Supplementary Fig. 19b). In addition, we have confirmed the enzymatic modifications (i.e.,

acetylation by PatZ and YjaB, promiscuous acylation by YiaC, and lipoylation by LplA, 2-hydroxyisobutyrylation by TmcA) for a selection of the putative substrates. Our direct substrate capturing approach is capable of identifying the directly interacting substrates in a spatiotemporal-resolved manner (Fig. 3b, Supplementary Figs. 1b, 17, 18), as well as capturing and validating of acylation site of the substrates in a straightforward fashion.

PTMs of lysine residues are crucial for cellular processes from gene expression to signal transduction. In current studies, our strategy was applied to decode substrate profiles of five different protein lysine acylases PatZ, YiaC, TmcA, YjaB, and LplA, as well as two mammalian acyltransferases Tip60 and GCN5, which not only retrieved known substrates but also identified candidate substrates. Several candidates were validated by in vitro crosslinking and enzymatic assay, and the two assays agreed well with each other regarding the modification sites, which further corroborated the concept of our strategy. In addition, this approach also could identify crosslinked peptides/sites from the cell lysate samples directly (Supplementary Figs. 13, 14). Remarkably, biochemical studies on acetylation sites (AcrA-131Kac) regulated by PatZ showed a potential role of PatZ in intrinsic antibiotic drug resistance of bacteria (Fig. 4e–h and Supplementary Fig. 26), suggesting that the AcrA-131Kac was potentially vital to the antibiotic susceptibility of *E. coli* (Supplementary Fig. 26). Moreover, we revealed various potential functional substrates of YiaC, such as the YiaC-dependent acetylation on K114 of RPL14 in translation process. Furthermore, in our work, many putative substrates of LplA have been revealed, which are involved in diverse biological processes (Fig. 6f). After the analysis of the interfaces of enzyme-substrate (Supplementary Fig. 8a), we found that the lipoylation sites of potential substrates PhoP and NarL might be recognized by LplA enzyme in a mode similar to the known substrate apoH (Supplementary Fig. 8b–e). Compared with the acetylation of K102 and K201 of StPhoP in *S. Typhimurium*^{39,62}, we revealed the lipoylation of K101 (Supplementary Fig. 9) and K200 on PhoP in *E. coli*. possibly have an effect on the function of PhoP to suppress the transcriptional activity and block DNA binding, suggesting a potential role of LplA in the regulation of bacterial virulence. Of note, in these substrate profiling studies, the photo-crosslinker o-NBAK was incorporated into the catalytic site tunnel of the acyltransferases (Supplementary Figs. 10, 19, 22), rather than on the interaction interface of PPIs, thus the o-NBAK mainly captures the proximal PTM-acceptor lysines (within appropriate distance and suitable angle, Supplementary Fig. 17) on interacting substrates, conferring the spatial resolution in our substrate capturing process.

Indeed, despite several excellent crosslinking strategies for capturing protein and interacting partner (or ligand) interactions^{17–26,28–30}, as well as GECX strategy^{27,32–34}, methods for direct capturing substrates of protein-modifying enzymes remain largely unexplored with an aim to capture bona fide substrates but not the non-substrate binding proteins³¹. On the other hand, the potential steric hindrance would be caused by the side chain of Uaas, thus, the o-NBAK should be incorporated at the appropriate position in the active site tunnels of acylases (Supplementary method, guideline for the selection of appropriate sites). Based on the insights into the structures of acylases that the active site tunnels of acylases are relatively large³⁸ (Supplementary Fig. 10), and our investigation of the perturbation to the enzyme-substrate interactions with o-NBAK incorporation at appropriate position in active site tunnels of the target enzymes (Supplementary Fig. 11), as well as our success in capturing previously known substrates for several enzymes (e.g., Fig. 3e, RpoA; Fig. 6d, apoH; Supplementary Fig. 20, histone H4, histone H2A, Ran; Supplementary Fig. 21, histone H4), we demonstrated that the o-NBAK incorporated at appropriate sites on PTM-enzymes could have tolerable perturbation to the enzyme-substrate interactions. As demonstrated in our work, the direct capturing strategy for global substrates of PTM-enzymes crucially depends on tolerable perturbation to the cellular enzyme-

substrate interacting complex formation, the chemical mechanism of a certain PTM, and the accessibility of PTM-acceptor residues in proximity to the enzyme active site for residue-specific covalent cross-linking. Our developed strategy that spatiotemporally directly captures global substrates of PTM-enzymes in living cells has several features: (i) Genetic incorporation of a light-induced residue-selective crosslinker into enzyme catalytic site tunnel, which not only introduces tolerable perturbation to native enzymes (Supplementary Fig. 11), but also allows direct and spatiotemporal capture of PTM-acceptor residues on interacting substrates in situ (Fig. 3b, Supplementary Fig. 1b, 17, 18), deciphering substrates with multiple lysine modifications as well as the precise modification sites of the substrates. (ii) In comparison with genetically encoded non-selective photo-crosslinkers, our lysine-selectivity photo-crosslinkers o-NBAK yielded higher overall enrichment efficiency (Supplementary Figs. 15, 16, AbK Uaa, etc.), thus, conferring higher sensitivity for substrate capturing. (iii) As substrate profiling solely based on the identification of crosslinked peptides via MS/MS analysis would result in high-level of false negative result (i.e., failure to identify a large number of bona fide substrates), owing to the intrinsic technical challenge of XL-MS experiment^{63,64}, we used global substrate profiling criteria based on the identification of captured proteins via XL-MS with stringent data filtration criteria (i.e., proteins required to be identified across all replicates in either UV-treat or control group, and with ratio >2 and p value < 0.05 were considered as candidate substrates), which would help in decreasing false negative result and excluding elusive false positive substrates. (iv) Lysine acyltransferases are known to have promiscuous activity, i.e., one enzyme can catalyze different types of lysine acylation. Although our method could not identify a certain type of lysine modifications directly, we can determine the crosslinked peptides/sites by MS/MS analysis (Figs. 4b–d, Figs. 6g, h, Supplementary Figs. 3, 13, 14). In principle, our approach is capable of identifying all types of acylation substrates of a given acyltransferase in current studies rather than a single type of PTM substrate. Subsequent enzymatic acylation assays, or modification-specific antibody-based enrichment experiments, combined with crosslinked peptides/sites (Fig. 4b–d, Fig. 6g, h, Supplementary Figs. 3, 13, 14) will be helpful to further characterize a certain type of lysine acylation substrates and PTM-sites (Figs. 4f–h and 5e–g, Supplementary Figs. 12, 23b, c, 24b, c, 25). Thus, our method provides a complementary approach to the genetic perturbation-based method for acyltransferase research. Collectively, integrating these features, the developed strategy for direct capturing substrates of PTM-enzymes confers higher sensitivity and accuracy for the characterization of enzyme-substrate interactions, and, thus, the substrate profiles of PTM-enzymes.

The lysine-selectivity-based direct capturing strategy in living cells is expected to have broad applicability⁴¹. Over the past two decades, incorporation of Uaas into protein of interest has provided a widespread and powerful method for capturing protein-protein interactions in living cells^{19,20,22}, while the capturing strategy may not work in the case of structures unknown. Harnessing the power in protein structure prediction, for the PTM-enzymes with structure unknown, the structures could be predicted based on the recent advance of the accurate protein structure prediction with AlphaFold⁶⁵. Therefore, integrating the advancement of genetic code expansion technology and protein structure prediction, it is worth to note that our direct capturing strategy could be applied to capture substrates of other acylation enzymes, such as enzymes for malonylation, glutarylation, crotonylation, lipidation, and so on. In principle, the lysine-selectivity-based capturing chemistry also could be expanded to other lysine involved interactome, such as determining ligand-receptor interactions, mapping protein interfaces, and so on^{20–22}. Thus, this strategy expands the landscape of enzyme-centric methods to investigate substrate profiles of lysine-specific PTM-enzymes (e.g., lysine acylome) and protein interactive proteomics^{36,42}.

In conclusion, the direct capturing substrates of PTM-enzymes strategy offers a chemical tool for global substrate profiling of PTM-enzymes in living cells, largely expanding the substrate profiles of PTM-enzymes, such as the bacterial PatZ, YiaC, LplA, TmcA and YjaB, and the mammalian GCN5 and Tip60. With high residue-selectivity, spatiotemporal resolution, covalent capture, and tolerable perturbations introduced to the native PTM enzymes, this strategy enables decoding substrates in higher sensitivity and accuracy, as well as validation of the acylation site of the substrates in a straightforward manner. The concept of direct capturing substrates of PTM-enzyme could inspire and lead to further development of advanced capturing strategies based on other kinds of residue-selectivity crosslinking chemistry harnessing very recent advances²² for other types of PTM-enzymes, to expand the repertoire of enzyme-centric methods investigating more PTMs and the landscape of substrate-enzymes interactive proteomics.

Methods

Site-specific incorporation of o-NBAK into proteins, expression, and purification

The plasmid pBK-*Mm*PylRS-384F/306A encoding the evolved PylRS and its cognitive tRNA, which inserts o-NBAK into the in-frame amber codon site on target genes in *E. coli* cells was described in the previous literature⁴⁴. For more information on choosing the sites on target enzymes that could be incorporated with Uaas, please see Supplementary Information. Plasmids pTAK-PatZ-I800TAG, pTAK-PatZ-F810TAG, pTAK-PatZ-L813TAG, pTAK-PatZ-S816TAG, pTAK-YiaC-F75TAG, pTAK-LplA-A138TAG, pTAK-TmcA-R457TAG, or pTAK-YjaB-F77TAG were co-transformed with pBK-*Mm*PylRS-Y384F/Y306A into BL21(DE3) complement cells (50 μ L) using heat shock. The transformants were recovered in 500 μ L SOC media at 37 °C for 1 hour before plating to 2YT agar plate containing 50 μ g/mL of kanamycin and 34 μ g/mL of chloramphenicol. A single colony from the plate was picked and used to inoculate 5 mL 2YT containing 50 μ g/mL kanamycin and 34 μ g/mL chloramphenicol. A 2-mL aliquot of overnight culture was then used to inoculate 200 mL 2YT containing the same concentrations of antibiotics. The cells were grown at 37 °C until OD₆₀₀ reached ~0.5. The culture was divided into two 100 mL portions. One portion of the culture was supplemented with 1 mM o-NBAK, and the other portion served as a control by adding medium at the same volume. After culturing for another 20 min at 37 °C, the protein expression was induced by adding 0.5 mM IPTG at 30 °C for 8 hours. The cells were pelleted in 50-mL conical tubes by centrifugation at 4200 $\times g$ for 30 min at 4 °C and stored at –80 °C.

Next day, the cell pellets were re-suspended in 10 mL lysis buffer. After incubation on ice for 30 min, the lysates were further treated with ultrasonication. Following centrifuge, the supernatants were incubated with 100 μ L Ni-NTA agarose beads at 4 °C for 2 hours with gentle shaking. The resin was centrifuged briefly and washed with 10 column volume (CV) of lysis buffer (devoid of lysozyme) and 10 CV of washing buffer. Finally, the protein was eluted with elution buffer. A 10- μ L aliquot from the elution fraction was mixed with one-fourth amount of 5 \times SDS loading buffer and heated at 98 °C for 8 minutes, before loading onto SDS-PAGE gel.

Expression and photo-crosslinking of o-NBAK incorporated enzyme in 293-F cells

FreeStyle™ 293-F cells were cultured in SMM 293-TII Expression Medium and passaged every 2–3 days, starting with the cell density around 0.2×10^6 cells per ml. Transfection was performed at cell density $\sim 2.0 \times 10^6$ cells per ml. Polyethyleneimine molecular mass 40,000 (PEI, 1 mg/mL) of 60 μ L was diluted in 1 mL SMM 293-TII Expression Medium. DNA mixture (Tip60-C369TAG or GCN5-E575TAG: PylRS = 1: 3; or empty vector: PylRS = 1:3) of 20 μ g was diluted in 1 mL SMM 293-TII Expression Medium. Diluted DNA and PEI

solution were mixed and incubated at room temperature for 15 min before adding to the cell culture. o-NBAK (0.5 mM) was added immediately after transfection for o-NBAK incorporation.

Forty to forty-eight hours after transfection, the cells were collected, washed with cold PBS twice and illuminated for 15 min at 10 mm distance from a ZF-7A 16 W 365 nm UV light on ice. The cells were collected, washed with cold PBS, lysed ultrasonically in cooled RIPA lysis buffer supplemented with protease inhibitor cocktail. The lysates were centrifuged at $15,000 \times g$ for 30 min at 4 °C. The supernatant was obtained and normalized by BCA protein assay. The lysates were subjected to incubation with MagStrep type3 XT bead at 4 °C overnight. The beads were washed with twofold diluted RIPA lysis buffer and wash buffer (100 mM Tris•HCl, 150 mM NaCl, 1 mM EDTA, pH = 8) three times each. Proteins attached to the beads were eluted in 2× Laemmli buffer by heating at 95 °C for 8 min. The eluates were resolved by SDS-PAGE and visualized by silver staining. Gel lanes above the free Tip60 or GCN5 enzyme were cut for subsequent in-gel digestion.

In vitro photo-crosslinking reactions

Purified o-NBAK-containing proteins (in PBS buffer 150 mM NaCl, pH 7.5) were mixed with partner proteins (in PBS buffer 150 mM NaCl, pH 7.5), the mixture was divided and transferred to a 96-well plate (BD Biosciences), and irradiated at 10 mm distance from a handheld 365-nm UV lamp for 15 minutes on ice. Samples were collected, diluted with SDS loading buffer, and analyzed by SDS-PAGE.

Photo-crosslinking of o-NBAK incorporated proteins in living *E. coli* cells

For each sample, BL21(DE3) cells were co-transfected plasmid pBK-MmPylRS-384F/306 A, as well as pTAK-PatZ-I800TAG, pTAK-PatZ-F810TAG, pTAK-PatZ-L813TAG, pTAK-PatZ-S816TAG, pTAK-YiaC-F75TAG, pTAK-LplA-A138TAG, pTAK-TmcA-R457TAG, or pTAK-YjaB-F77TAG. The cells were grown in 2YT medium containing 50 µg/mL of kanamycin and 34 µg/mL of chloramphenicol overnight, and then were used to inoculate 200 mL 2YT containing the same concentrations of antibiotics. The cells were grown at 37 °C until OD600 reached ~0.5, and 1 mM o-NBAK was supplemented. After culturing for another 20 min at 37 °C, the protein expression was induced by adding 0.5 mM IPTG at 30 °C for 5 hours. Cells were harvested, washed twice with PBS (Gibco), and re-suspended to PBS. The mixture was divided into two portions. One portion was transferred to sixwell plate (3–4 mL/well) and irradiated at 10 mm distance with a handheld UV lamp (365-nm) for 20 min on ice. The other portion was kept in dark acting as the control group. The cells were harvested by centrifugation at $1000 \times g$ for 5 min and lysed in lysis buffer. The lysate was incubated with Ni-NTA agarose beads at 4 °C for 2 hours with gentle shaking. The beads were then washed with wash buffer three times. The proteins were finally eluted with elution buffer, concentrated using a 0.5 mL centrifugal filter, and separated by SDS-PAGE. Then the gel was visualized by coomassie blue stain or silver stain.

In vitro enzymatic acylation assay

The in vitro enzymatic acylation was performed in the reaction buffer containing 20 mM sodium phosphate (pH 8.0), 150 mM NaCl, 10% glycerol, 0.5 mM tris(2-chloroethyl) phosphate (TCEP). The in vitro acetylation by PatZ was carried out by incubation of PatZ (20 µg), acetyl-CoA (with concentration indicated in Fig. 4 for AcrA; or 0.1 mM for DnaJ, HflX), and candidate substrates (30 µg) in 100 µL reaction buffer at 37 °C for 2 hours. The in vitro acetylation by YiaC was performed in 50 µL reaction buffer containing YiaC (4 µg), acyl-CoA (1 mM for acetyl CoA; 0.5 mM for 2-hydroxyisobutyryl-CoA, propionyl-CoA, or malonyl-CoA), and candidate substrates (20 µg) at 37 °C for three hours. The in vitro acetylation by TmcA was performed in 50 µL reaction buffer containing TmcA (15 µg), 2-hydroxyisobutyryl-CoA (0.5 mM), ATP (1 mM), and candidate substrates (20 µg) at 37 °C for

three hours. The in vitro acetylation by YjaB was performed in 50 µL reaction buffer containing YjaB (4 µg), Ac-CoA (0.5 mM), and candidate substrates (20 µg) at 37 °C for three hours. The samples were resolved by SDS-PAGE for subsequent western blot analysis, or in-gel digestion and LC-MS/MS analysis.

Western blot analysis

Loading buffer with additional β-Me (2%) to denature the protein sample, protein samples in SDS-PAGE sample buffer were boiled for 8 min, resolved on an SDS-PAGE gel, and transferred to PVDF membrane (Thermo Fisher Scientific). The membrane was blocked in 5% skim milk in TBST (50 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH 7.5) at room temperature for 1 hour. Then the membrane was incubated with mouse anti-His tag antibody (1:10000, Proteintech, Cat. No. HRP-66005), mouse anti-vinculin monoclonal antibody (1:2000, Cat. No. 66305-1-Ig), mouse monoclonal anti-acetyl lysine antibody (1:1000, Cell Signaling Technology, Cat. No. 9681), mouse anti-strep-tag II monoclonal antibody (1:2000, Abbkine, Cat. No. ABT2230), rabbit polyclonal FtsZ antibody (1:1000, CUSABIO, Cat. No. CSB-PA359270HA01EGX) in TBST at 4 °C overnight. The membrane was washed with TBST (4 × 5 min) before the addition of the secondary goat anti-mouse horseradish peroxidase conjugate (1:4000, Proteintech, Cat. No. SA00001-1) or HRP-conjugated goat anti-rabbit antibody (1:5000, Cell Signaling Technology, Cat. No. 7074). After 1 hour, the membrane was washed with TBST (4 × 5 min). Final detection of HRP activity was performed using ECL Plus chemiluminescent substrate (Thermo Fisher Scientific). The image was acquired with Image Quant LAS 4000 or ChemoScope 6300. Uncropped images can be found in the Source Data File.

Bio-layer interferometry (BLI) assay

The binding characteristics between dN-AcrA-wt or dN-AcrA-K131Kac and TolC were determined using anti-HIS (HIS2) biosensors in the Octet RED96 system (ForteBio Inc., Menlo Park, CA, USA). The membrane protein TolC was loaded onto the HIS2 biosensors at 4 µM in PBS containing 150 mM NaCl, 0.03% DDM, and 5% glycerol. The biosensors were blocked with His8-tagged MBP protein (2 mg/ml) for 60 s. Diluted AcrA in PBS solution containing 150 mM NaCl, 0.03% DDM, and 5% glycerol was then added to the HIS2 biosensors loaded with TolC. The real-time binding response ($\Delta\lambda$ in nanometer, nm) between AcrA and TolC was calculated by subtracting the non-specific binding of AcrA to the HIS2 biosensors from the binding of AcrA with TolC.

In vitro lipoylation reaction

The in vitro lipoylation reaction was performed in the buffer containing 40 mM PBS (pH 7.0), 0.5 mM dithiothreitol (DTT), 10 µM LplA, 20 µM PhoP or NarL, 2 mM ATP, 2 mM MgCl₂, 1 mM lipoic acid for 3 h at 37 °C. Then, the reaction mixture was separated by SDS-PAGE and visualized by Coomassie blue stain.

In-gel digestion

The gel slices containing bands of interest were destained with 50% ethanol. After being balanced with distilled water, the bands were cut into cubes (~1 mm³) and dehydrated in acetonitrile (ACN). The cubes were then incubated with 25 mM NH₄HCO₃ containing 10 mM DTT at 56 °C for 1 h and then with 25 mM NH₄HCO₃ containing 55 mM iodoacetamide (IAA) at r.t. for 45 min in darkness. The gel cubes were then dehydrated in 50% ACN, vacuum-dried, and incubated with 50 mM NH₄HCO₃ containing 10–20 ng trypsin (sequencing grade, Hualishi Scientific, Beijing) at 37 °C overnight. For in vitro PatZ-substrate crosslinking, LplA-apoH crosslinking, crosslinking, and lipoylation reactions of LplA-NarL, samples were further incubated with 10–20 ng endoproteinase Glu-C (sequencing grade, Roche) at 37 °C for additional 4 h. The peptides were sequentially extracted with 50% ACN (containing 5% TFA), 75% ACN (containing 0.1% TFA) and

100% ACN, and vacuum-dried. The peptides were desalted with ZipTip C18 pipette tips (Merck-Millipore, Darmstadt, Germany) and ready for subsequent LC-MS/MS analysis.

LC-MS/MS analysis

The peptides were reconstituted in buffer A (0.1% formic acid, 2% ACN) and loaded onto 15 cm C18 reversed-phase capillary analytical column (3 μ m particle size, 90 Å pore size, Dikma Technologies, Lake Forest, CA) connected to an EASY-nLC 1000 HPLC system (Thermo Fisher Scientific, Waltham, MA).

For analyses of the samples from substrate profiling by PatZ-L813o-NBAK crosslinking in living cells, the peptides were eluted with a linear gradient of buffer B (0.1% formic acid, 90% ACN) from 8% to 32% in 58 min, followed by an increase to 48% in 6 min and further increase to 80% in 2 min at a constant flow rate of 300 nL/min. The samples from crosslinking substrate profiling in living cells (by YiaC-F75o-NBAK, LpIA-A138o-NBAK, TmcA-R457o-NBAK, YjaB-F77o-NBAK, GCN5-E575o-NBAK or Tip60-C369o-NBAK), substrate validation by in vitro enzymatic assays (catalyzed by wild type YiaC, LpIA, TmcA, and YjaB), the peptides were eluted with a linear gradient of buffer B from 8% to 13% in 20 min, followed by an increase to 26% in 31 min, to 45% in 5 min and to 80% in 1 min at a constant flow rate of 300 nL/min. The eluted peptides were ionized and sprayed into an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, Waltham, MA) in a positive mode. The mass spectrometric analysis was carried out in a “Top speed” data-dependent acquisition (DDA) mode with a cycle time of 3 s. The peptides with a range of m/z 350–1300 were analyzed by an Orbitrap mass analyzer with a resolution of 120,000, automatic gain control (AGC) target of 5×10^5 , and maximum ion injection time (IT) of 50 ms. The dynamic exclusion was set as 60 s, and the charge inclusion was set to 2–6+. The isolated precursor ions were subjected to fragmentation via higher-energy collisional dissociation (HCD) with a normalized collision energy (NCE) of 32%, and analyzed by an ion trap analyzer.

For analyses of samples from in vitro acetylation by wt YiaC or lipoylation by wt LpIA, the setting for LC and MS are the same as above except that the dynamic exclusion was set as 30 s, and the charge inclusion was set to 2–4+ for MS1 analysis.

For analyses of samples from in vitro acetylation by wt PatZ, the peptides were eluted with a linear gradient of buffer B from 7% to 35% in 24 min, followed by an increase to 80% in 3 min. The eluted peptides were ionized and sprayed into an Orbitrap Fusion mass spectrometer in a positive mode. The mass spectrometric analysis was carried out in a “Top 20” DDA mode. The peptides with a range of m/z 350–1500 were analyzed by an Orbitrap mass analyzer with a resolution of 240,000, AGC target of 7×10^5 , and IT of 50 ms. The dynamic exclusion was set as 10 s, and the charge inclusion was set to 2–4+. The isolated precursor ions were subjected to fragmentation via collision-induced dissociation with an NCE of 35%, and analyzed by ion trap analyzer.

To identify crosslinking sites, the peptides were eluted with a linear gradient of buffer B from 5% to 7% in 13 min, followed by an increase to 10% in 20 min, 25% in 55 min, 45% in 22 min, 80% in 3 min at a constant flow rate of 300 nL/min. The eluted peptides were ionized and sprayed into an Orbitrap Fusion mass spectrometer in a positive mode. The mass spectrometric analysis was carried out in a “Top speed” DDA mode with a cycle time of 3 s. The peptides with a range of m/z 300–1700 were analyzed by an Orbitrap mass analyzer with a resolution of 120,000, AGC target of 5×10^5 , and maximum IT of 100 ms. The dynamic exclusion was set as 60 s, and the charge inclusion was set to 3–7+. The isolated precursor ions were subjected to fragmentation via HCD with an NCE of 27–33% (stepped collision energy = 3%), and analyzed by Orbitrap analyzer at a resolution of 15,000, automatic AGC target of 2×10^5 and IT of 250 ms, the dynamic exclusion was set as 10 s.

All LC-MS/MS data were collected with XCalibur (3.0). For identification of proteins in experiments of substrate profiling by enzyme-o-NBAK mutants, raw MS data were processed with MaxQuant software (v. 1.6.6.0)⁶⁶ against UniProt *Escherichia coli* database (2019.2 release, 4350 sequences) or human database (2019.2 release, 95943 sequences). For data from the in vitro acetylation assay, MS raw data were analyzed by Mascot search engine (v2.3, Matrix Science, London, UK)⁶⁷ against the corresponding protein sequences. For the identification of crosslinks, raw data were searched against MeroX search engine (v2.0)⁶⁸ against the sequences for proteins of interest. For details, see Supplementary Information. The MS/MS spectra of identified acylated or crosslinked peptides were annotated by Mascot or MeroX software, respectively. In some cases, the annotated MS/MS spectra were displayed using pLabel⁶⁹ for better visualization.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the mass spectrometry proteomics raw data and the database search result files have been deposited to ProteomeXchange consortium via the iProX partner repository under the dataset identifier [PXD040318](https://doi.org/10.26434/chemrxiv-2024-pxd04). The GO-BP enrichment analysis was performed with PANTHER database (v16). Protein structures were retrieved from PDB: SPatA (PDB ID: [4NXY](https://doi.org/10.22541/au.16888888)), YjaB (PDB ID: [2KCW](https://doi.org/10.22541/au.16888888)), LpIA (PDB ID: [3A7A](https://doi.org/10.22541/au.16888888)), PhoP (PDB ID: [2PKX](https://doi.org/10.22541/au.16888888); [2PL1](https://doi.org/10.22541/au.16888888)), NarL (PDB ID: [1RNL](https://doi.org/10.22541/au.16888888)), TmcA (PDB ID: [2ZPA](https://doi.org/10.22541/au.16888888)), GCN5 (PDB ID: [1Z4R](https://doi.org/10.22541/au.16888888)), Tip60 (PDB ID: [2OU2](https://doi.org/10.22541/au.16888888)). The processed proteomics data and the GO-BP enrichment analysis data are available in the Supplementary Data files. Other data generated in this study are provided in the Supplementary Information/Source Data file. Source data are provided in this paper.

References

1. Keenan, E. K., Zachman, D. K. & Hirsche, M. D. Discovering the landscape of protein modifications. *Mol. Cell* **81**, 1868–1878 (2021).
2. Bludau, I. & Aebersold, R. Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nat. Rev. Mol. Cell. Biol.* **21**, 327–340 (2020).
3. Xu, H. et al. PTMD: A database of human disease-associated post-translational modifications. *Genomics Proteomics Bioinformatics* **16**, 244–251 (2018).
4. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).
5. Li, X. & Li, X. D. Chemical proteomics approaches to examine novel histone posttranslational modifications. *Curr. Opin. Chem. Biol.* **24**, 80–90 (2015).
6. Burton, A. J. et al. In situ chromatin interactomics using a chemical bait and trap approach. *Nat. Chem.* **12**, 520–527 (2020).
7. Syu, G.-D., Dunn, J. & Zhu, H. Developments and applications of functional protein microarrays. *Mol. Cell. Proteomics* **19**, 916–927 (2020).
8. Bidlingmaier, S. & Liu, B. Identification of posttranslational modification-dependent protein interactions using yeast surface displayed human proteome libraries. *In: Yeast surface display: methods, protocols, and applications* (ed. Liu, B.) 193–202 (Springer New York, New York, NY, 2015).
9. Ivry, S. L. et al. Global substrate specificity profiling of post-translational modifying enzymes. *Protein Sci.* **27**, 584–594 (2018).
10. Zheng, J., Chen, X., Yang, Y., Tan, C. S. H. & Tian, R. Mass spectrometry-based protein complex profiling in time and space. *Anal. Chem.* **93**, 598–619 (2021).
11. Jaeger, M. G. & Winter, G. E. Fast-acting chemical tools to delineate causality in transcriptional control. *Mol. Cell* **81**, 1617–1630 (2021).

12. Olsen, J. V. & Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteom.* **12**, 3444–3452 (2013).
13. Morrison, D. K. MAP kinase pathways. *Cold Spring Harb. Perspect. Biol.* **4**, a011254 (2012).
14. Sinz, A. Cross-linking/mass spectrometry for studying protein structures and protein–protein interactions: where are we now and where should we go from here? *Angew. Chem. Int. Ed.* **57**, 6390–6396 (2018).
15. Yang, B. et al. Identification of cross-linked peptides from complex samples. *Nat. Methods* **9**, 904–906 (2012).
16. Liu, F., Rijkers, D. T. S., Post, H. & Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).
17. Lang, K. & Chin, J. W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **114**, 4764–4806 (2014).
18. Xiao, H. & Schultz, P. G. At the interface of chemical and biological synthesis: an expanded genetic code. *Cold Spring Harb. Perspect. Biol.* **8**, a023945 (2016).
19. Zhang, S. et al. Conditional chaperone–client interactions revealed by genetically encoded photo-cross-linkers. *Acc. Chem. Res.* **50**, 1184–1192 (2017).
20. Coin, I. Application of non-canonical crosslinking amino acids to study protein–protein interactions in live cells. *Curr. Opin. Chem. Biol.* **46**, 156–163 (2018).
21. Nguyen, T.-A., Cigler, M. & Lang, K. Expanding the genetic code to study protein–protein interactions. *Angew. Chem. Int. Ed.* **57**, 14350–14361 (2018).
22. Wang, N. & Wang, L. Genetically encoding latent bioactive amino acids and the development of covalent protein drugs. *Curr. Opin. Chem. Biol.* **66**, 102106 (2022).
23. Chin, J. W. et al. Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
24. Hino, N. et al. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Methods* **2**, 201–206 (2005).
25. Coin, I. et al. Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CFR class B GPCR. *Cell* **155**, 1258–1269 (2013).
26. Xuan, W., Shao, S. & Schultz, P. G. Protein crosslinking by genetically encoded noncanonical amino acids with reactive aryl carbamate side chains. *Angew. Chem. Int. Ed.* **56**, 5096–5100 (2017).
27. Yang, B. et al. Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions. *Nat. Commun.* **8**, 2240 (2017).
28. He, D. et al. Quantitative and comparative profiling of protease substrates through a genetically encoded multifunctional photo-crosslinker. *Angew. Chem. Int. Ed.* **56**, 14521–14525 (2017).
29. Tang, H. et al. Proteomic identification of protein tyrosine phosphatase and substrate interactions in living mammalian cells by genetic encoding of irreversible enzyme inhibitors. *J. Am. Chem. Soc.* **140**, 13253–13259 (2018).
30. Tang, S. et al. Mechanism-based traps enable protease and hydrolase substrate discovery. *Nature* **602**, 701–707 (2022).
31. Lopez, J. E., Haynes, S. E., Majmudar, J. D., Martin, B. R. & Fierke, C. A. HDAC8 substrates identified by genetically encoded active site photocrosslinking. *J. Am. Chem. Soc.* **139**, 16222–16227 (2017).
32. Liu, C. et al. Identification of protein direct interactome with genetic code expansion and search engine OpenUaa. *Adv. Biol.* **5**, 2000308 (2021).
33. Liu, J. et al. A genetically encoded fluorosulfonyloxybenzoyl-L-lysine for expansive covalent bonding of proteins via SuFEX chemistry. *J. Am. Chem. Soc.* **143**, 10341–10351 (2021).
34. Liu, J. et al. Genetically encoding photocaged quinone methide to multitarget protein residues covalently in vivo. *J. Am. Chem. Soc.* **141**, 9458–9462 (2019).
35. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell. Biol.* **18**, 90–101 (2017).
36. Zhao, S., Zhang, X. & Li, H. Beyond histone acetylation-writing and erasing histone acylations. *Curr. Opin. Struct. Biol.* **53**, 169–177 (2018).
37. Su, D., Kosciuk, T., Yang, M., Price, I. R. & Lin, H. Binding affinity determines substrate specificity and enables discovery of substrates for N-myristoyltransferases. *ACS Catal.* **11**, 14877–14883 (2021).
38. Zhang, Z. J., Pedicord, V. A., Peng, T. & Hang, H. C. Site-specific acylation of a bacterial virulence regulator attenuates infection. *Nat. Chem. Biol.* **16**, 95–103 (2020).
39. Ren, J. et al. Acetylation of lysine 201 inhibits the DNA-binding ability of phop to regulate salmonella virulence. *PLoS Pathog.* **12**, e1005458 (2016).
40. Dong, H. et al. TmcA functions as a lysine 2-hydroxyisobutyryltransferase to regulate transcription. *Nat. Chem. Biol.* **18**, 142–151 (2022).
41. Neumann-Staubitz, P., Lammers, M. & Neumann, H. Genetic code expansion tools to study lysine acylation. *Adv. Biol.* **5**, 2100926 (2021).
42. Thinon, E. & Hang, H. C. Chemical reporters for exploring protein acylation. *Biochem. Soc. Trans.* **43**, 253–261 (2015).
43. Guo, A.-D. et al. Light-induced primary amines and o-nitrobenzyl alcohols cyclization as a versatile photoclick reaction for modular conjugation. *Nat. Commun.* **11**, 5472 (2020).
44. Hu, W. et al. Genetically encoded residue-selective photo-cross-linker to capture protein–protein interactions in living cells. *Chem* **5**, 2955–2968 (2019).
45. Guo, A.-D. et al. Spatiotemporal and global profiling of DNA–protein interactions enables discovery of low-affinity transcription factors. *Nat. Chem.* **15**, 803–814 (2023).
46. Christensen, D. G. et al. Post-translational protein acetylation: an elegant mechanism for bacteria to dynamically regulate metabolic functions. *Front. Microbiol.* **10**, 1604 (2019).
47. Castaño-Cerezo, S. et al. Protein acetylation affects acetate metabolism, motility and acid stress response in *Escherichia coli*. *Mol. Syst. Biol.* **10**, 762 (2014).
48. Sun, M. et al. Characterization of protein lysine propionylation in *Escherichia coli*: Global profiling, dynamic change, and enzymatic regulation. *J. Proteome Res.* **15**, 4696–4708 (2016).
49. Xu, J.-Y. et al. Characterization of the lysine acylomes and the substrates regulated by protein acyltransferase in *Mycobacterium smegmatis*. *ACS Chem. Biol.* **13**, 1588–1597 (2018).
50. Tucker, A. C., Taylor, K. C., Rank, K. C., Rayment, I. & Escalante-Semerena, J. C. Insights into the specificity of lysine acetyltransferases. *J. Biol. Chem.* **289**, 36249–36262 (2014).
51. de Diego Puente, T. et al. The protein acetyltransferase PatZ from *Escherichia coli* is regulated by autoacetylation-induced oligomerization. *J. Biol. Chem.* **290**, 23077–23093 (2015).
52. Christensen, D. G. et al. Identification of novel protein lysine acetyltransferases in *Escherichia coli*. *mBio* **9**, e01905–e01918 (2018).
53. Orchard, S. et al. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–D363 (2014).
54. Du, D. et al. Structure of the AcrAB–TolC multidrug efflux pump. *Nature* **509**, 512–515 (2014).
55. Parks, A. R. & Escalante-Semerena, J. C. Modulation of the bacterial CobB sirtuin deacetylase activity by N-terminal acetylation. *Proc. Natl Acad. Sci. USA* **117**, 15895 (2020).

56. Häuser, R. et al. RsfA (YbeB) proteins are conserved ribosomal silencing factors. *PLoS Genet.* **8**, e1002815 (2012).
57. Rowland, E. A., Snowden, C. K. & Cristea, I. M. Protein lipoylation: an evolutionarily conserved metabolic regulator of health and disease. *Curr. Opin. Chem. Biol.* **42**, 76–85 (2018).
58. Fujiwara, K. et al. Crystal structure of lipoate-protein ligase A from *Escherichia coli*: Determination of the lipoic acid-binding site. *J. Biol. Chem.* **280**, 33645–33651 (2005).
59. Groisman Eduardo, A., Duprey, A. & Choi, J. How the PhoP/PhoQ system controls virulence and Mg²⁺ homeostasis: Lessons in signal transduction, pathogenesis, physiology, and evolution. *Microbiol. Mol. Biol. Rev.* **85**, e00176–20 (2021).
60. Song, X. et al. Dynamic crotonylation of EB1 by Tip60 ensures accurate spindle positioning in mitosis. *Nat. Chem. Biol.* **17**, 1314–1323 (2021).
61. Shvedunova, M. & Akhtar, A. Modulation of cellular processes by histone and non-histone protein acetylation. *Nat. Rev. Mol. Cell. Biol.* **23**, 329–349 (2022).
62. Ren, J. et al. Metabolic intermediate acetyl phosphate modulates bacterial virulence via acetylation. *Emerg. Microbes Infect.* **8**, 55–69 (2019).
63. Yu, C. & Huang, L. Cross-linking mass spectrometry: An emerging technology for interactomics and structural biology. *Anal. Chem.* **90**, 144–165 (2018).
64. Piersimoni, L., Kastritis, P. L. & Sinz, A. Cross-linking mass spectrometry for investigating protein conformations and protein–protein interactions: a method for all seasons. *Chem. Rev.* **122**, 7500–7531 (2022).
65. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
66. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
67. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
68. Götze, M. et al. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **26**, 83–97 (2015).
69. Wang, L.-h et al. pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 2985–2991 (2007).
- MS analysis of selected substrates; in-cell crosslinking of PatZ, LplA, and TmcA for directly identifying crosslinked peptides; part of enzymes and substrates expression; bioinformatics analysis; BS2G-mediated chemical crosslinking. W.H. performed model design and experiments of PatZ, YiaC, and LplA enzymes, including plasmid construction, crosslinking experiments for substrate profiling in *E. coli* and in vitro crosslinking with selected substrates, substrate proteins expression of PatZ, YiaC, and LplA; in vitro PatZ acetylation and WB analysis; AcrA mutant and TolC expression; BLi assay for AcrA with TolC. A.-D.G. performed Tip60 and GCN5 plasmid construction, expression optimization, and crosslinking in mammalian cells. L.Z. performed part of LC-MS/MS experiments and analyzed the data in the identification of crosslinked peptides. S.M. assisted in plasmid construction and protein expression. H.-J.N. and B.-S.Z. synthesized the Uaas. L.T. performed antibiotics assays. X.J. performed part of the WB analysis. X.L. and X.Y. technically assisted in Tip60 experiments. X.-H.C. and M.T. directed and supervised the project and analyzed the data. X.-H.C., H.H. and M.T. wrote the manuscript with input from other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45765-3>.

Correspondence and requests for materials should be addressed to Minjia Tan or Xiao-Hua Chen.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Acknowledgements

This study was supported by the National Science Foundation of China No. (92053106 to X.-H.C.), (22225702 to M.T.), (92153302 to M.T.), (2022YFA1303100 to X.L.), (32090040 to X.Y.), (81821005 to M.T.), (21907100 to H.H.), China Postdoctoral Science Foundation (2019M661670 to W.H.), the National Key Science & Technology Program of China (2020YFE0202200 to M.T.); the Innovative Research Team of High-Level Local Universities in Shanghai (SHSMU-ZDCX20212700 to M.T.). We thank Dr. Chao Liu from Beihang University for the technical assistance of XL-MS data analysis. pBR322-TmcA was a gift from professor Kai Zhang; pcDNA3.1-HA-GCN5 and pCMV-Flag-Tip60 were gifts from professor Jiemin Wong).

Author contributions

H.H. performed the workflow in all MS analysis experiments; TmcA and YjaB substrate profiling and validation; enzymatic acylation assays and