Bioorthogonal Photocatalytic Profiling of Mitochondrial Proteomes from Primary Living Samples

Ziqi Liu[1,#], Fuhu Guo[1,#], Yufan Zhu[1], Shengnan Qin[1], Yuchen Hou[1], Feng Lin[2], Haotian Guo[1], Peng R Chen[1,2,*], Xinyuan Fan[1,2,*]

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

[2]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

[#]These authors contributed equally.

*E-mail: xinyuanfan@pku.edu.cn (X.F.), pengchen@pku.edu.cn (P.R.C.)

Abstract

In situ profiling of subcellular proteomic networks in primary and living systems, such as primary cells derived from native tissues or clinic samples, is crucial for the understanding of life processes and diseases, yet challenging for the current proximity labeling methods (e.g., BioID, APEX) due to their necessity of genetic engineering. Here we report CAT-S, a new-generation bioorthogonal photocatalytic chemistry-enabled proximity labeling method, that extends proximity labeling to a wide range of primary living samples for *in situ* profiling of subcellular proteomes. Powered by the newly introduced thioQM labeling warhead and targeted bioorthogonal photocatalytic decaging chemistry, CAT-S enables *in situ* labeling of mitochondrial proteins in living cells with high efficiency and specificity. We applied CAT-S to distinct cell cultures, mouse tissues (kidney and spleen) as well as primary T cells from human blood, portraying the mitochondrial proteomic characteristics for various primary living samples, and unveiled a set of hidden mitochondrial proteins in human proteome. Furthermore, CAT-S allows quantitative analysis of the *in situ* proteomic perturbations on dysfunctional tissue samples, exampled by the comparative proteomics of diabetic mouse kidneys, and revealed the alterations of lipid metabolism machinery that drive the disease progression. Given the unique advantages of non-genetic operation, generality, efficiency as well as high spatiotemporal resolution, we envision that CAT-S may open new avenues as a proximity labeling strategy for *in situ* investigation of the native-state subcellular proteomic landscape of primary living samples that are otherwise inaccessible, promoting our understanding of the molecular mechanisms underlying biological and pathological processes.

Introduction

Biological processes are rigorously regulated by proteomic networks organized within subcellular compartments. Perturbations of subcellular proteome networks are linked to various cell abnormalities and diseases, such as the tumorigenesis[1] and neurodegeneration[2] that are driven by mitochondrial dysfunctions. Given the high complexity, dynamicity and fragility of subcellular proteomes, technologies that directly capture subcellular proteomic information in living cells is thus in high demand, which would significantly facilitate biological discovery as well as therapeutic development[3]. In recent years, the proximity labeling approaches (e.g. BioID, APEX)[4,5] have emerged as a milestone for *in situ* biological study by enabling on-demand capture of proteomes from specific subcellular spaces in living cells. However, these enzymatic approaches require genetic engineering to introduce exogenous enzymes for intracellular labeling, hindering their utilities in cells from primary samples such as native tissues and clinic samples[6], which are crucial samples that are directly linked to mechanistic dissection of biological or pathological processes. Conventional subcellular fractionation or organelle isolation methods[7] could be applied on these samples for *ex situ* subcellular proteomics, but require cell lysis processing (e.g., shearing force, detergent) which would fundamentally alter the native state of cells and organelles[8], and also suffer from large material input, poor spatiotemporal resolution as well as inevitable contamination. Therefore, *in situ* proteomic dissection of these important primary living samples remains a long-standing but challenging goal.

We envision that a chemical approach that is able to *in situ* label the proteome with high spatiotemporal resolution without the need of genetic manipulation, might be able to address this unmet need. Recently, a series of novel proximity labeling approaches (e.g., μMap)[9-12] based on photocatalysts in the replacement of enzymes have been reported, but largely restricted to extracellular or *ex situ* applications (e.g., cell surface, isolated nuclei) as well as demonstrations on model cell lines. Alternatively, by coupling the emerging bioorthogonal photocatalytic decaging chemistry[13, 14] with an activable labeling probe Quinone Methide (QM), we conceived a chemical system, termed CAT-Prox[15]. This non-genetic system leveraged the subcellular-targeting photocatalyst for living cells, which activated chemical probe for intracellular proximity protein labeling in a photocatalytic manner. However, investigation of the more complexed and clinically relevant primary living samples was still unachieved due to the reactivity limitations of the labeling

warhead QM probe, which was the key component in the chemical labeling system.

Herein, we present CAT-S system (Figure 1), a new generation of transfection-free proximity labeling technique for efficient *in situ* capture of mitochondrial proteome of primary cells and tissues. Through the development, a new chemical labeling warhead, thio-quinone methide (thioQM), was created and introduced as a more efficient and accurate proximity labeling probe. Combining with the thoroughly evolved photocatalytic labeling system, subcellular proteome profiling has been largely improved in terms of efficiency and accuracy, enabling the dissection of the mitochondrial proteome of diverse samples including hard-to-transfect cells or primary cells from living tissues as well as clinical blood specimen, which otherwise remain inaccessible by current methods. Furthermore, the high labeling efficiency and negligible sample perturbations of our CAT-S system allowed the identification of new mitochondrial proteins as well as subcellular protein atlas of diseases within their native states, providing a landscape view as well as molecular insights of the primary and living samples from physiological or clinical settings.

Results and Discussions

Establishment of the CAT-S system

We started by the investigation on the chemical labeling warhead, which was the key factor for assuring the labeling efficiency and sensitivity. As a reactive Michael receptor, QM has been previously reported as a reactive intermediate for covalent labeling of protein residues[16]. However, the QM-based proteome capture[15] displayed significant lower efficiency than the enzymatic methods, resulting in inadequate proteomic coverage and sensitivity loss, likely due to their relatively lower reactivity than the reactive intermediate in enzymatic approaches such as phenolic radicals in APEX[5]. We turned to tune the reactivity of QM probes by adapting the chemical structures[17]. Notably, substitution of the oxygen atom of QM with sulfur atom would yield a new type of QM species, thio-quinone methide (thioQM)[18], which has been an underdeveloped reactive species but might possess much higher reactivity attributing to the following properties: i) faster formation of the thioQM intermediate upon the decaging reaction due to the less electronegativity of sulfur atom than the oxygen atom (2.4 vs 3.5), and ii) more stable state of the reactive zwitterionic resonance due to the larger atomic radius of sulfur (0.88 Å vs 0.48 Å), making the thioQM with 10^2-10^3 fold higher reactivity than the canonical QMs in nucleophilic addition reactions[16, 18]. Furthermore, the hydrolytic quenching rate of the thioQM intermediate is faster than the oxygen-

QM in aqueous environment[16, 18], which would lead to higher specificity in protein labeling due to their shorter diffusing radius in organelles. Given this significant enhancement in reactivity as well as labeling specificity, we reasoned that thioQM might hold great promise as an intriguing new warhead for labeling chemistry on proteins.

For the chemical evolution of labeling warheads, we designed and synthesized an array of activable probes based on QM and thioQM (Figure 2a). The phenolic oxygen or sulfur atom in these probes were protected by *para*-azidophenyl (PAB) group, which can be removed by the bioorthogonal photocatalytic decaging chemistry, releasing QM or thioQM upon light trigger to label proximal proteins *in situ*. The tethered biotin tag further enabled detection or isolation of the captured proteins. The *in vitro* labeling efficiency was first investigated by using bovine serum albumin (BSA) as the model protein (Figure 2b). Using the commercially available photocatalyst **Ir1** and additive NADH, we observed significant labeling within minutes upon mild blue light irradiation for all the probes (Figure 2c and Supplementary Figure 1). To our delight, the thioQM probes exhibited higher labeling efficiencies than the commonly used oxygen-QM probes as expected, with the difluoro-substituted thioQM probe (**SF2**) showing the highest labeling intensity, likely due to the further increased electrophilic reactivity by introducing electron-withdrawing fluorine atom. In order to exclude undesired potential labeling by the arylazide moiety other than the thioQM, we replaced the PAB group with a routinely used protecting group ONP[19] (removable by UV light) by synthesizing **SF2^{uv}** probe for comparison (Figure 2d). Decaging of **SF2^{uv}** upon UV irradiation led to similar intensive labeling as the **SF2** upon photocatalytic decaging, suggesting that the labeling was dependent on thioQM but not the arylazide moiety. By LC-MS/MS analysis after purifying the modified peptides from the thioQM-labeled BSA[20], we identified multiple potential sites with expected thioQM modifications (Figure 2e and Supplementary Figure 2). The **SF2**-derived thioQM warhead exhibited promiscuous reactivity to a wide range of nucleophilic protein residues (e.g., Lys, Arg, Cys, Tyr, Ser, Gln), similar to the previously reported QMs[21].

We next investigated the applicability of the new probes for capturing proteins in living cells (Figure 2f). After sequential incubation with photocatalyst **Ir1** and QM probe, the living HeLa cells were irradiated by blue LED to activate the QM probe for *in situ* protein labeling. Indeed, all four probes exhibited significant protein labeling in the presence of photocatalyst **Ir1**, among which **SF2** probe again presented the highest labeling efficiency (Figure 2f). Therefore, we selected **SF2** as the

optimal labeling probe for further investigations.

Since the photocatalyst is the key player for both mitochondria specificity and catalytic activity, systematic evaluation of the photocatalyst was conducted next, in terms of mitochondria targeting and photocatalytic activity synergistically (Figure 2g-i). By tracking the fluorescent signals of various iridium photocatalysts in living cells, we observed significant localization of cationic compounds **Ir1~Ir5** in mitochondria (Figure 2g and Supplementary Figure 3). In contrast, the neutrally charged compound **Ir6** displayed nearly no targeting ability, indicating that the positive charge played an essential role for the mitochondrial targeting, due to the electrochemical gradient known as mitochondrial membrane potential (MMP)[23, 24]. Next, we performed the photocatalytic decaging chemistry using **SF2** probe in living cells to evaluate the photocatalytic activity (Figure 2h). Remarkably, photocatalyst **Ir1**, which bears a *tert*-butyl-bipyridine ligand and two difluorophenyl-(trifluoromethyl)pyridine ligands, displayed much higher labeling activity than the other catalysts, suggesting that the electron-withdrawing group on phenylpyridine ligand might enhance the catalytic performance. Taken together, with its outstanding performance on both mitochondrial targeting specificity as well as photocatalytic activity in living cells, **Ir1** was chosen as the optimal photocatalyst for the CAT-S system (Figure 2i). Finally, by immunofluorescence detection, we visualized the **Ir1**-catalyzed labeling signals inside cells, with the biotinylation signal precisely colocalized with the mitochondria marker (Figure 2j and 2k). Statistical analysis further revealed high Pearson's R values (0.87 for HeLa and 0.90 for HEK293T) of colocalization, confirming the spatial specificity of our labeling system.

Validation and evaluation of the CAT-S system

With the chemical system established, we sought to capture mitochondrial proteomes in living cells with in-depth dissection by mass spectroscopy (Figure 3 and Supplementary Figure 4), especially for those samples difficult for enzymatic approaches. As an initial verification, we applied our CAT-S system in cancer cell line HeLa by using **SF2** as the labeling probe and **Ir1** as the photocatalyst (Figure 3a-c, top row). Living Hela cells were collected after 12-min irradiation by mild blue LED (450 nm, ~4 mW/cm^2) and the biotinylated proteins were enriched by streptavidin beads. After LC-MS/MS analysis with thousands of proteins detected, we applied ratiometric cutoff based on the distribution of true- and false-positives to filter off background signals (Figure 3b, top).

The final protein list for HeLa across biological triplicates revealed 391 enriched proteins, including 326 known mitochondrial proteins (83% specificity) according to MitoCarta3.0[25] and UniProt[26] database (Figure 3d). Surprisingly, the mitochondrial proteins accounted for 95% of the detected protein MS abundance, suggesting the highly efficient capture of mitochondria proteome by our method (Figure 3e). In contrast, previous chemical method CAT-Prox (187 mitochondrial proteins, 71% specificity)[15] and mitochondria-localizable reactive molecules (MRMs, 259 mitochondrial proteins, 69% specificity)[27] displayed much lower mitochondrial protein coverage as well as specificity, indicating the significant advancement of the current CAT-S system in terms of labeling efficiency and sensitivity (Figure 3f).

Next, in order to compare the efficiency of CAT-S with the enzymatic technology, we chose the genetic operation-accessible human embryonic kidney 293T (HEK293T) cells for evaluation (Figure 3a-c, middle row). Using the same CAT-S protocol as for HeLa cells, we obtained a protein list of 422 proteins including 334 known mitochondrial proteins (79% specificity, accounting for 96% of protein abundance) (Figure 3d), which showed both increased specificity and mitochondrial protein number than the previously reported enzymatic studies[28] using BioID (255, 70% specificity), TurboID (209, 67% specificity) or miniTurbo (269, 75% specificity) for the same cell type.

Given the independence of genetic operation, the CAT-S technique is advantageously accessible to hard-to-transfect cells. Therefore, we next evaluated CAT-S with K562 cells[29] (Figure 3a-c, bottom row), which are hematopoietic leukemia known to be difficult for genetic modification. Immunoblotting analysis on K562 cells applied with CAT-S displayed significant biotinylation signals, indicating the successful photocatalytic labeling. Notably, LC-MS/MS analysis yielded a list of 386 proteins containing 336 known mitochondrial proteins (87% specificity, accounting for 92% of protein abundance) (Figure 3d). Combined analysis of the aforementioned three cell lines indicated that the captured proteomes were largely overlapped, with mitochondrial proteins responsible for essential pathways in mitochondria (Figure 3g, 3h and Supplementary Figure 5). Sub-mitochondrial profiles of the captured mito-proteomes were all similar to the whole human mito-proteome, suggesting our system might unbiasedly capture proteins from the whole mitochondria (Figure 3i). Collectively, the successful applications to K562 and other cells validated the generality of CAT-S to portray mitochondrial proteomes for various cell types, especially the hard-to-transfect cells that are inaccessible for enzymatic approaches.

Additionally, to address the concern of potential photo-cytotoxic effects by photocatalyst, we meanwhile assessed the biocompatibility of our system (Supplementary Figure 6 and 7). After performing CAT-S to label mitochondrial proteome in living cells, we collected the cells for proliferation assay to evaluate cellular integrity and activity. Gratifyingly, all the three cell lines (HeLa, HEK293T and K562) after CAT-S labeling did not only stay alive, but also retained the same proliferation activity as the control group (Supplementary Figure 6), indicating CAT-S as an advantageous non-toxic proximity labeling technique. Furthermore, "Two-round proteome labeling" by applying CAT-S on the same living cell sample twice could be achieved, further confirming the high biocompatibility of our CAT-S system (Supplementary Figure 7).

New mitochondrial proteins discovered by CAT-S

Besides the known mitochondrial proteins, several non-mitochondrial proteins according to the MitoCarta3.0 and UniProt database were frequently observed by our CAT-S system, which prompted us to re-investigate their locations inside cells (Figure 4). Through analysis for cell lines, we identified a fraction of proteins as "mito orphans" without prior mitochondrial annotation but clustered with known mitochondrial proteins (Figure 4a).

Applying the well-established validation strategy[5], we investigated the subcellular location of potential "mito orphans" by expressing "orphan" proteins in HEK293T cells for imaging. To our delight, four proteins (PTPN1, SLC35A4 upstream ORF, TMEM160 and TRABD) were identified as new mitochondrial proteins, with predominant colocalization with mitochondria (Figure 4b and 4c). According to the zoom-in view, the phosphatase PTPN1[30] was mainly located on the mitochondrial outer membrane, possibly playing a role in the signal transduction on mitochondrial surface. Transmembrane proteins SL35A4 upstream ORF, TMEM160 and TRABD, which were highly uncharacterized proteins without clear knowledge of subcellular locations or functions in database, might be potentially intriguing targets for further investigation. These findings indicated the *in situ* profiling by CAT-S could also shed light on the hidden members of mitochondrial proteome.

CAT-S enabled mitochondrial proteome profiling for living primary tissues

Capturing subcellular proteomes in living tissues is highly desirable yet challenging. Applying

the enzymatic proximity labeling approaches to tissues normally requires the construction of transgenic animal model, which could be time-consuming and often inapplicable in many cases (e.g., clinical human tissues). Given the transfection-free advantage and the biocompatibility of CAT-S, we sought to capture mitochondrial proteomes directly from living tissues (Figure 5a). Since human metabolism and immunity critically rely on mitochondria, kidney- and splenocyte-related diseases have been often found in association with mitochondrial dysfunctions[31, 32], making dissection of mitochondrial protein networks crucial for mechanistic understanding. We chose to apply our CAT-S system in mitochondrial proteome profiling on kidney and spleen, the highly metabolic organ and the essential immune organ, respectively. As demonstrated, the tissues were excised from healthy B6 mice and disassociated into cell suspensions with red blood cells removed, and then subjected to CAT-S procedure for mitochondrial proteomic labeling and LC-MS/MS analysis. Quality control measurement by western blotting displayed significant biotinylation signals after photocatalytic labeling (Figure 5b). Across biological replicates and cutoff filter (Figure 5c), we finally revealed a list of 430 proteins (261 known mitochondrial proteins, 61% specificity) for mice kidney, and 357 proteins (228 known mitochondrial proteins, 64% specificity) for mice spleen(Figure 5d-f). Further analysis revealed that mitochondrial proteins contributed 75-80% of the protein abundance, indicating the effective mitochondrial proteomic capture in distinct primary living tissues (Figure 5e).

Seeking for biological insights on mitochondria at the tissue level, we performed Gene Ontology (GO) analysis for our captured proteins (Figure 5g, 5h and Supplementary Figure 8), especially for those proteins only detected in kidney (Figure 4f) or spleen (Figure 4g). For kidney-only dataset, aerobic respiration was ranked as the most enriched process, suggesting the preference of kidney mitochondria to energy metabolism (Figure 5g), which is consistent with kidney's role as the second highest oxygen consumer in body at rest[31]. Moreover, enzymes related with various small molecule metabolic pathways were enriched in kidney mitochondria, underpinning the kidney's role as a highly metabolic organ. In contrast, spleen samples were significantly enriched with proteins related with mito-gene expression, mitochondrial organization, and mitochondrial transport pathways (Figure 5h), suggesting the higher activity of mitochondrial biogenesis and morphological regulation in spleen. Given that spleen accommodates a huge quantity of immune cells (predominantly B cells and T cells), our observation indicated the elevated mitochondrial biogenesis

and dynamics within this organ that might be required by immune cell development and migration[33].

Mitochondrial proteome atlas for the diabetic disease enabled by CAT-S

The global diabetes pandemic, especially the obesity-induced type-2 diabetes (T2D), continues to grow as a severe threat to human health[34]. As a metabolic disorder, diabetes usually leads to various health complications including cardiovascular and kidney diseases, which are closely linked to mitochondrial dysfunctions [31, 35]. However, the molecular mechanisms underlying the diabetes-associated diseases remain largely unclear. To this end, we sought to apply CAT-S for profiling the native state of the mitochondrial proteomes from living kidneys under diabetic settings (Figure 6 and Supplementary Figure 9).

We utilized the well-documented *db/db* mice as the obesity-induced type-2 diabetes model[36], and *m/m* mice as the nondiabetic group for comparison (Figure 6a and 6b). The kidneys from 10-week-old obese-diabetic and healthy mice were excised, and subjected to CAT-S procedure as described above. Protein labeling was confirmed by immunoblotting, with labeling intensities at the similar level on both samples (Figure 6c). To quantitatively compare the mitochondrial proteomes from disease and healthy kidneys as well as filter out background signals, we used triplex dimethyl labeling for MS-based relative quantification. After the cutoff analysis for each experiment, we intersected the protein lists from biological triplicate to generate a high-confidence protein list (Figure 6b), consisting of 400 proteins including 281 known mitochondrial proteins (70% specificity) (Figure 6d and 6e). The data was further visualized as volcano plot (Figure 6f), showing over one hundred up- and down- regulated mitochondrial proteins from diabetic mouse kidney..

The identified differentially regulated proteins spanned a diverse biological functions and pathways. We conducted multiple bioinformatic investigations including mapping, clustering and enrichment analysis to elucidate biological processes and functional networks from our data (Figure 5g, 6h and Supplementary Figure 9). For upregulated mitochondrial proteins in obese-diabetic kidney (Figure 5g, right), the cellular respiration process was highly enriched, in harmony with the elevated oxygen consumption to sustain the high energy demand in early diabetic kidney disease[37]. According to the clustering analysis based on protein-protein interaction (PPI) (Figure 6g) and mitochondrial pathways (Supplementary Figure 9), we clearly observed the elevation of oxidative phosphorylation (OXPHOS) complexes in obese-diabetic group, along with the upregulations of

TCA-cycle, glucose oxidation, fatty acid oxidation and amino acid catabolism, concurrently underlying the hypermetabolic phenotypes in the early diabetic kidney[31].

Interestingly, we also detected apparent alterations for lipid metabolism, mitochondrial transmembrane transport, detoxification and other pathways, with complicated regulations under obese-diabetic conditions. Given the importance of lipid metabolism in diabetes[38] , we further chose three lipid-metabolism enzymes (Aldh3a2, Acsm2 and Cpt1b) with significant changes according to proteomic profile for further validation (Figure 6i, 6j and Supplementary Figure 10). Using both whole-cell lysate and mitochondrial lysate as samples, we discovered significant upregulation (Cpt1b) and downregulations (Aldh3a2 and Acsm2) for obese diabetic kidney, consistent with the CAT-S proteomic data. In conjunction with the previous evidences suggesting Cpt1b's contribution to insulin resistance[39] and Acsm2's relevance to kidney abnormality[40], we suspected that the reshaping of lipid metabolic network might play essential and multi-faceted roles during diabetic nephropathy. Particularly, the aberrant decrease of Aldh3a2 and Acsm2, which are enzymes responsible for converting fatty aldehydes and medium-chain/xenobiotic fatty acids respectively, would lead to the accumulation of such lipids and consequent cellular toxicities (e.g., metabolism disruption, biomolecule damage by aldehydes, signaling alterations)[41-43]. Thus, we concluded that the obese-diabetic condition could suppress the detoxification machinery in lipid metabolism, therefore driving the progression of kidney disorders and diseases. Noteworthy, these molecular features have not been revealed before in proteomic investigations of diabetic kidney[44], which may provide new insights for our understanding of the pathological processes, further demonstrating the advantages of our CAT-S technique for portraying *in situ* mitochondrial landscape in primary living tissue samples.

Mitochondrial proteomics for primary T cells from human PBMC

Applying the CAT-S system for clinically relevant samples, such as cells from human peripheral blood that is critical for practical diagnosis, would be another exciting direction for deciphering human biology and diseases. Among the cell types in peripheral blood mononuclear cells (PBMC), T cells are critical in adaptive immunity as well as diverse diseases including cancers, infections and autoimmunity, rendering the characterization of human blood T cells in high demand for diagnosis and mechanistic study. Moreover, mitochondria have been found to play fundamental

and multi-faceted roles underlying diverse T cell processes[33], particularly on activation[45, 46] and exhaustion[47], with the molecular characteristics yet resolved.

Given that primary T cells and other immune cells are difficult for transfection, we reasoned that CAT-S could be applied to T cells from clinical blood samples for *in situ* mitochondrial proteomics (Figure 7a)[6]. To obtain pure and living T cells, we firstly separated PBMCs from the whole blood by density-gradient centrifugation, and then performed magnetic negative cell sorting to isolate primary T cells. Immunoblotting assay showed that 20 nM **Ir1** generated intensive protein labeling (Figure 7b), suggesting the high sensitivity of CAT-S for T cell mitochondrial proteome labeling. With the limited volume of the clinical blood samples, we used ~3×10^7 primary T cells (< 0.5 mg total cellular protein due to the small cellular size) in each labeling and control group for MS analysis. We performed CAT-S experiments using samples from three healthy donors. The mitochondria-specific labeling was confirmed by distribution profile (Figure 7c). After cutoff analysis to reduce non-specific background, the protein lists of three donors were obtained, with 219 known mitochondrial proteins detected from more than two donors (62% specificity), which accounted for 75% of the detected protein abundance (Figure 7e and 7f).

Gene Ontology analysis further elucidated the major pathways (Figure 7g) and functions (Figure 7h) involved by our identified mitochondrial proteins in primary T cells from PBMC, exhibiting a profile similar to our dataset on mice splenocytes (Supplementary Figure 8). Cellular respiration and mitochondrion organization ranked as the most enriched biological processes, in accord with their indispensable roles for cell homeostasis. Interestingly, in contrast to the results on fast-proliferating cell lines mentioned above (Supplementary Figure 5), the process of mitochondrial gene expression was detected with relatively low enrichment for primary T cells, possibly due to the quiescent state of T cells under normal conditions. Although our method herein was demonstrated by using healthy human samples, further application of CAT-S for comparative mitochondrial profiling between disease and healthy conditions, or between different stages of T cell activation, would unveil more critical roles of mitochondria for immunobiology[48].

Conclusions

In summary, we have developed a new-generation transfection-free chemical platform CAT-S for proximity labeling in primary living cells and tissues, which leveraged the advantages of the bioorthogonal photocatalytic decaging methods and the newly developed thioQM-mediated labeling

chemistry. By coupling with in-depth MS-based proteomics, CAT-S was applied to a variety of living cell samples, including multiple cell lines, mouse tissues as well as primary T cells from human blood, providing mitochondria-centric views of proteome landscape as well as relative quantification between disease-related conditions.

Via the development and application of CAT-S, intriguing insights into both chemistry and biology studies have been unveiled. During the chemical exploration, we developed a new activatable warhead, thioQM, which showed excellent proximity labeling efficiency that enabled the otherwise difficult protein capturing in living primary tissue samples. The reactive thioQM could be generated on demand by various decaging strategy (e.g., photocatalytic or UV decaging), showing potentials for diverse applications involving covalent labeling. By profiling mitochondrial proteomes of distinct cell types, we meanwhile revealed a new set of mitochondrial located or associated proteins, including three newly identified proteins (phosphatase PTPN1, transmembrane proteins SLC35A4 upstream ORF, TMEM160 and TRABD) that are predominantly located in mitochondria. Moreover, by quantitatively comparing mitochondrial proteomes between obese-diabetic and nondiabetic kidneys, we observed major alterations of pathways involving lipid metabolism, with metabolic enzymes such as Cpt1b, Acsm2 and Aldh3a2 differentially regulated, suggesting that obese-diabetic condition suppresses the detoxifying lipid metabolic machinery to progress the diabetic nephropathy.

The CAT-S system exhibited a set of advantages as a new-generation *in situ* mitochondrial proteomics technique: i) CAT-S enabled *in situ* mitochondrial proteome capture in diverse living cell samples including hard-to-transfect cells, primary tissues and blood samples that are otherwise inaccessible by genetic engineering and enzymatic approaches; ii) by introducing the new thioQM warhead for efficient protein labeling, CAT-S significantly enhanced the capture efficiency and specificity over previous chemical methods; iii) CAT-S showed high biocompatibility with no obvious cellular toxicity under the experiment conditions, in contrast to the methods with harsh treatment; iv) CAT-S is flexibly controlled by external light to turn on/off the photocatalytic labeling system, with more precise control of labeling time-window than the substrate-controlled enzymatic methods, enabling high temporal resolution for potential dynamic investigations.

Future applications and extensions of CAT-S strategy can be foreseen in multiple aspects. For methodology development, CAT-S can be further coupled to the continuously evolving LC-MS/MS

techniques and bioinformatic analysis, thus enabling higher sensitivity and deeper dissection of mitochondria as well as other subcellular locations. Indeed, given that small molecule photocatalysts could be chemically engineered to target various subcellular spaces[49], further adaption of CAT-S system to profile other subcellular proteomes in primary tissue samples are currently under investigation in our laboratory. For further applications, CAT-S is advantageously accessible to living primary samples, and investigations on their *in situ* mitochondrial proteomes of different stages (e.g., T cell activation, exhaustion), disease conditions, or therapeutic effects (e.g., before/after immune checkpoint therapy) by CAT-S can also be foreseen. A complete chemical toolbox for the unprecedented mapping of subcellular-resolved proteomics of intact primary samples with excellent spatiotemporal resolution is expected, which will greatly facilitate our understanding of the native molecular mechanism for biological and pathological processes from real primary samples.

Methods

Materials, protocols and data characterizations for all biological and chemical experiments are described in detail in the Supplementary Information.

Acknowledgements

We acknowledge funding from the National Natural Science Foundation of China (22222701, 22077004, 92253301, 21937001, 22137001), the Ministry of Science and Technology (2019YFA0904201, 2022YFA1304700, 2022YFE0114900), Beijing Natural Science Foundation (Z200010) and Li Ge-Zhao Ning Life Science Junior Research Fellowship. Dr. Yicheng Weng is acknowledged for helpful discussion during the project.

Author contributions

X.F. and P.R.C. conceived and supervised the project. Z.L., F.G., Y.Z., S.Q., and Y.H. performed the experiments and analyzed the data. F.L. helped with the mice and human blood experiments. H.G. helped with material preparation and data analysis. Z.L., X.F. and P.R.C. wrote the paper with input from all authors.

Competing interests

X.F., P.R.C., and Z.L. have filed a patent application related to this work. All other authors declare no competing interests.

References

1. Vyas, S., Zaganjor, E. & Haigis, M.C. Mitochondria and Cancer. *Cell* **166**, 555-566 (2016).
2. Lin, M.T. & Beal, M.F. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* **443**, 787-795 (2006).
3. Lundberg, E. & Borner, G.H.H. Spatial proteomics: a powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology* **20**, 285-302 (2019).
4. Roux, K.J., Kim, D.I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology* **196**, 801-810 (2012).
5. Rhee, H.-W. et al. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* **339**, 1328-1331 (2013).
6. Al-Dosari, M.S. & Gao, X. Nonviral Gene Delivery: Principle, Limitations, and Recent Progress. *The AAPS Journal* **11** (2009).
7. Arslan, T. et al. SubCellBarCode: integrated workflow for robust spatial proteomics by mass spectrometry. *Nat. Protoc.* **17**, 1832-1867 (2022).
8. Nerem, R.M. Shear force and its effect on cell structure and function. *ASGSB Bull.* **4**, 87-94 (1991).
9. Geri, J.B. et al. Microenvironment mapping via Dexter energy transfer on immune cells. *Science* **367**, 1091-1097 (2020).
10. Oakley, J.V. et al. Radius measurement via super-resolution microscopy enables the development of a variable radii proximity labeling platform. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2203027119 (2022).
11. Seath, C.P. et al. Tracking chromatin state changes using nanoscale photo-proximity labelling. *Nature* **616**, 574-580 (2023).
12. Tay, N.E.S. et al. Targeted activation in localized protein environments via deep red photoredox catalysis. *Nat. Chem.* (2022).
13. Chen, Y., Kamlet, A.S., Steinman, J.B. & Liu, D.R. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat. Chem.* **3**, 146-153 (2011).
14. Liu, Z. et al. Spatially resolved cell tagging and surfaceome labeling via targeted photocatalytic decaging. *Chem* **8**, 2179-2191 (2022).
15. Huang, Z. et al. Bioorthogonal Photocatalytic Decaging-Enabled Mitochondrial Proteomics. *J Am Chem Soc* **143**, 18714-18720 (2021).
16. Toteva, M.M. & Richard, J.P. The generation and reactions of quinone methides. *Adv. Phys. Org. Chem.* **45**, 39-91 (2011).
17. Weinert, E.E. et al. Substituents on Quinone Methides Strongly Modulate Formation and Stability of Their Nucleophilic Adducts. *Journal of the American Chemical Society* **128**, 11940-11947 (2006).
18. Chiang, Y., Kresge, A.J., Sadovski, O. & Zhan, H.-Q. Kinetics and Mechanism of Hydration of o-Thioquinone Methide in Aqueous Solution. Rate-Determining Protonation of Sulfur. *The Journal of Organic Chemistry* **70**, 1643-1646 (2005).
19. Liu, Y. et al. Spatiotemporally resolved subcellular phosphoproteomics. *Proceedings of the National Academy of Sciences* **118** (2021).
20. Weerapana, E., Speers, A.E. & Cravatt, B.F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in

proteomes. *Nat. Protoc.* **2**, 1414-1425 (2007).
21. Liu, J. et al. Photocaged Quinone Methide Crosslinkers for Light-Controlled Chemical Crosslinking of Protein–Protein and Protein–DNA Complexes. *Angewandte Chemie International Edition* **58**, 18839-18843 (2019).
22. Wang, H. et al. Selective Mitochondrial Protein Labeling Enabled by Biocompatible Photocatalytic Reactions inside Live Cells. *JACS Au* **1**, 1066-1075 (2021).
23. Zorova, L.D. et al. Mitochondrial membrane potential. *Analytical Biochemistry* **552**, 50-59 (2018).
24. Ross, M.F. et al. Accumulation of lipophilic dications by mitochondria and cells. *Biochem J* **400**, 199-208 (2006).
25. Rath, S. et al. MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res.* **49**, D1541-D1547 (2021).
26. Bateman, A. et al. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* (2022).
27. Yasueda, Y. et al. A Set of Organelle-Localizable Reactive Molecules for Mitochondrial Chemical Proteomics in Living Cells and Brain Tissues. *Journal of the American Chemical Society* **138**, 7592-7602 (2016).
28. Branon, T.C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology* **36**, 880-887 (2018).
29. Richter, F. et al. Improved gene delivery to K-562 leukemia cells by lipoic acid modified block copolymer micelles. *Journal of Nanobiotechnology* **19** (2021).
30. Sangwan, V. et al. Regulation of the Met Receptor-tyrosine Kinase by the Protein-tyrosine Phosphatase 1B and T-cell Phosphatase. *Journal of Biological Chemistry* **283**, 34374-34383 (2008).
31. Forbes, J.M. & Thorburn, D.R. Mitochondrial dysfunction in diabetic kidney disease. *Nature Reviews Nephrology* **14**, 291-312 (2018).
32. Giri, J. et al. Leishmania donovani Exploits Myeloid Cell Leukemia 1 (MCL-1) Protein to Prevent Mitochondria-dependent Host Cell Apoptosis. *Journal of Biological Chemistry* **291**, 3496-3507 (2016).
33. Lisci, M. & Griffiths, G.M. Arming a killer: mitochondrial regulation of CD8+ T cell cytotoxicity. *Trends in Cell Biology* (2022).
34. Saeedi, P. et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* **157** (2019).
35. Chen, X., Wei, S. & Yang, F. Mitochondria in the pathogenesis of diabetes: a proteomic view. *Protein & Cell* **3**, 648-660 (2012).
36. Sharma, K., McCue, P. & Dunn, S.R. Diabetic kidney disease in thedb/dbmouse. *American Journal of Physiology-Renal Physiology* **284**, F1138-F1144 (2003).
37. Hansell, P., Welch, W.J., Blantz, R.C. & Palm, F. Determinants of kidney oxygen consumption and their relationship to tissue oxygen tension in diabetes and hypertension. *Clinical and Experimental Pharmacology and Physiology* **40**, 123-137 (2013).
38. Athyros, V.G. et al. Diabetes and lipid metabolism. *Hormones* **17**, 61-67 (2018).
39. Yang, Q. Carnitine Palmitoyltransferase 1b Deficiency Protects Mice from Diet-Induced Insulin Resistance. *J. Diabetes Metab.* **05** (2014).
40. Watanabe, H. et al. Expression of Acsm2, a kidney-specific gene, parallels the function and maturation of proximal tubular cells. *American Journal of Physiology-Renal Physiology* **319**,

F603-F611 (2020).
41. Yusuf, R.Z. et al. Aldehyde dehydrogenase 3a2 protects AML cells from oxidative death and the synthetic lethality of ferroptosis inducers. *Blood* **136**, 1303-1316 (2020).
42. van der Sluis, R. & Erasmus, E. Xenobiotic/medium chain fatty acid: CoA ligase - a critical review on its role in fatty acid metabolism and the detoxification of benzoic acid and aspirin. *Expert Opin. Drug Metab. Toxicol.* **12**, 1169-1179 (2016).
43. Thongnak, L., Pongchaidecha, A. & Lungkaphin, A. Renal Lipid Metabolism and Lipotoxicity in Diabetes. *Am. J. Med. Sci.* **359**, 84-99 (2020).
44. Zhang, D. et al. Proteomics analysis reveals diabetic kidney as a ketogenic organ in type 2 diabetes. *American Journal of Physiology-Endocrinology and Metabolism* **300**, E287-E295 (2011).
45. Ron-Harel, N. et al. Mitochondrial Biogenesis and Proteome Remodeling Promote One-Carbon Metabolism for T Cell Activation. *Cell Metabolism* **24**, 104-117 (2016).
46. Lisci, M. et al. Mitochondrial translation is required for sustained killing by cytotoxic T cells. *Science* **374** (2021).
47. Yu, Y.-R. et al. Disturbed mitochondrial dynamics in CD8+ TILs reinforce T cell exhaustion. *Nat. Immunol.* **21**, 1540-1551 (2020).
48. Mehta, M.M., Weinberg, S.E. & Chandel, N.S. Mitochondrial control of immunity: beyond ATP. *Nature Reviews Immunology* **17**, 608-620 (2017).
49. You, Y. Phosphorescence bioimaging using cyclometalated Ir(III) complexes. *Curr. Opin. Chem. Biol.* **17**, 699-707 (2013).
50. Szklarczyk, D. et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605-D612 (2021).

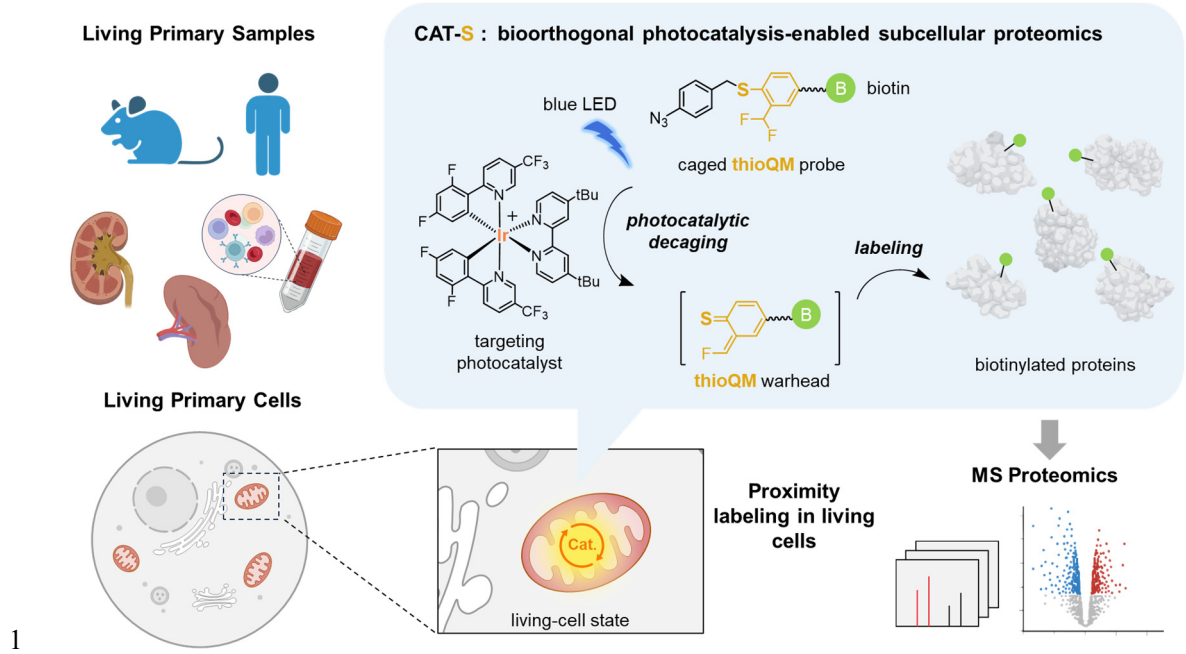

Figure 1. Schematic illustration of CAT-S

CAT-S is an *in situ* subcellular proteomics strategy that utilizes the cell-permeable Ir-based photocatalyst to target mitochondria in living cells from primary tissue or blood samples, which can catalyze the blue light-triggered bioorthogonal decaging reaction of the *p*-aziobenzyl (PAB) group to generate reactive thioQM probe for *in situ* proximity labeling of the mitochondrial proteome. No genetic engineering is required for CAT-S, rendering it widely applicable for various primary and living samples.

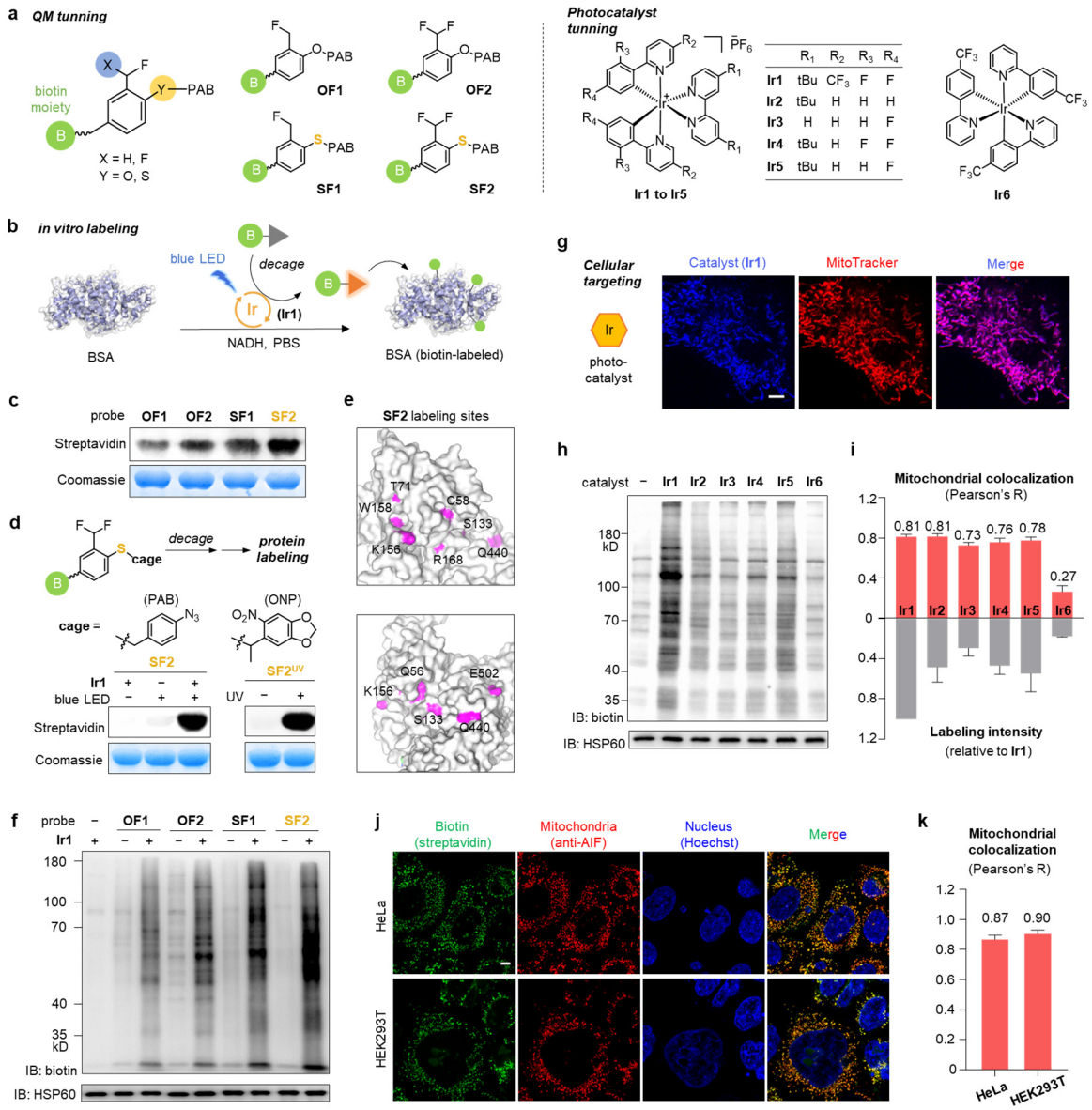

Figure 2. Establishment of the CAT-S system

a, Structural tunning of biotin-tethered QM/thioQM-based activatable probes and iridium photocatalysts. **b**, Schematic view of *in vitro* photocatalytic labeling. BSA was treated with **Ir1** (5 μM), QM probes (100 μM) and NADH (500 μM) for labeling (450-nm blue LED, ~4 mW/cm^2, 5 min). **c**, Comparison of *in vitro* labeling efficiencies between different QM probes. **d**, Decaging chemistries-controlled protein labeling by using the **SF2** warhead, as exampled by photocatalytic and UV-triggered decaging respectively. **e**, Potential photocatalytic labeling sites for **SF2** on BSA as identified by LC-MS/MS. **f**, Comparison of *in cellulo* labeling efficiencies between different probes. HeLa cells were treated with 100 nM **Ir1** and 100 μM probe for labeling. Mitochondrial protein HSP60 was blotted as the loading control. **g**, Confocal imaging of live HeLa cells treated with **Ir1** (0.5 μM, blue) and MitoTracker Deep Red (100 nM, red). Scale bar, 5 μm. **h**, Comparison of catalytic labeling efficiencies between different photocatalysts. HeLa cells were treated with photocatalyst (100 nM) and **SF2** probe (100 μM) for labeling. **i**, Quantification of mitochondrial colocalization (Pearson's R value, n = 15 cells) and catalytic labeling efficiency (anti-biotin blotting intensity, n = 2 biological replicates). **j**, Immunofluorescent images of HeLa (top) and HEK293T (bottom) cells after CAT-S labeling (100 nM **Ir1** and 100 μM **SF2**). Green, labeling signal detected

by streptavidin. Red, staining of mitochondrial marker AIF. Blue, staining of nucleus by Hoechst 33342. Scale bar, 5 μm. **k**, Quantification of mitochondrial colocalization (Pearson's R value, n = 15 cells) for CAT-S labeling.

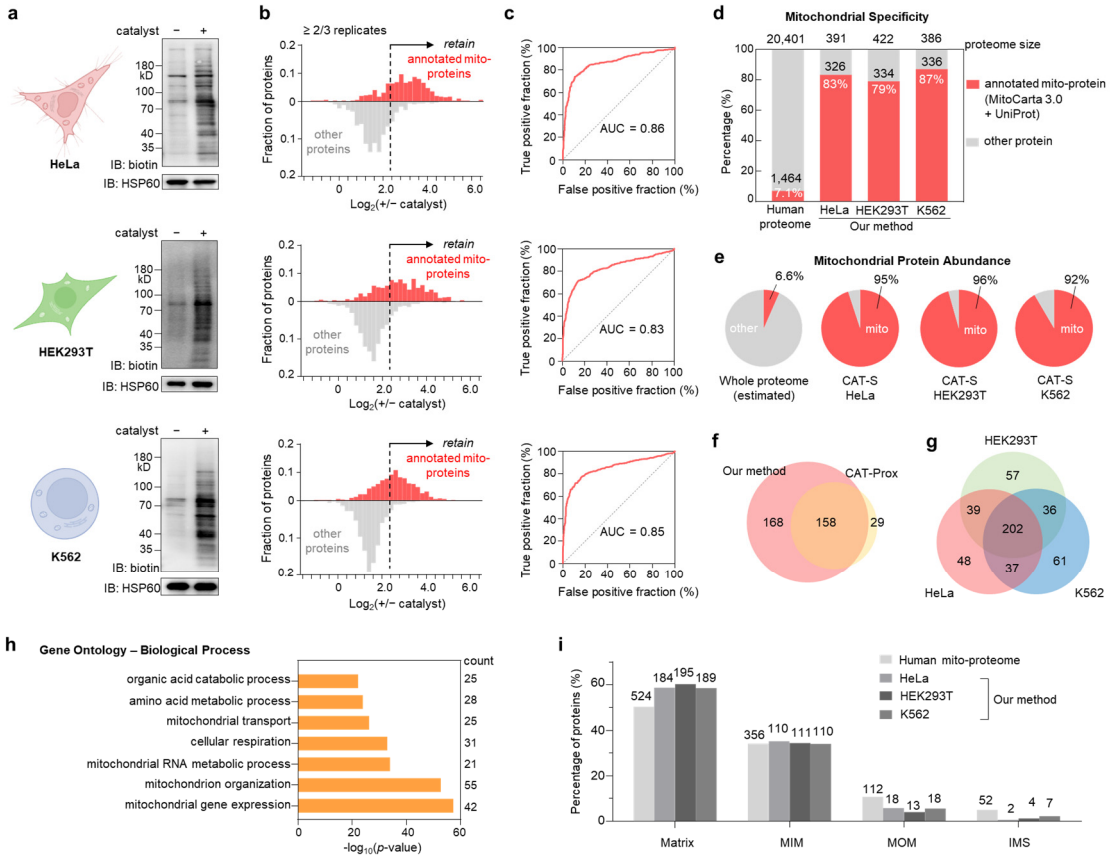

Figure 3. Validation and evaluation of the CAT-S system.

a-c, CAT-S-enabled *in situ* mitochondrial proteomics of living HeLa, HEK293T and K562cells (from top to bottom), treated with **Ir1** and **SF2** followed by blue LED irradiation. **a**, Labeling signals measured by immunoblotting. **b**, Frequency distribution profiles of MS-identified proteins (≥ 2/3 replicates) with (red) and without (gray) mitochondrial annotation according to the combined human MitoCarta3.0 and UniProt database. A filter based on +/− catalyst ratio (set to 5.0 for the three cell lines) was further applied to subtract background noise (non-specific proteins). **c**, ROC curves of CAT-S for detecting annotated mitochondrial proteins, with AUC (area under curve) values shown. **d**, Mitochondrial specificity analysis for HeLa, HEK293T and K562 cells with the columns showing the fraction of proteome with prior mitochondrial annotation in database. **e**, Contribution of annotated mitochondrial proteins (red) in the total protein MS abundance. The estimate abundance of mitochondrial proteins in the whole cellular proteome is shown for comparison. **f**, Comparison of mitochondrial proteomic coverage between this work and CAT-Prox. **g**, Venn diagram showing the overlap between captured mitochondrial proteomes. **h**, GO Biological Process analysis of the mutually captured mitochondrial proteome from HeLa, HEK293T and K562. Top terms ($P < 10^{-20}$) were shown. **i**, Sub-mitochondrial distribution profiles of captured mitochondrial proteomes from HeLa, HEK293T and K562 cells. Proteins were classified into “matrix”, “mitochondrial inner membrane (MIM)”, “intermembrane space (IMS)”, and “mitochondrial outer membrane (MOM)” according to MitoCarta3.0 database.

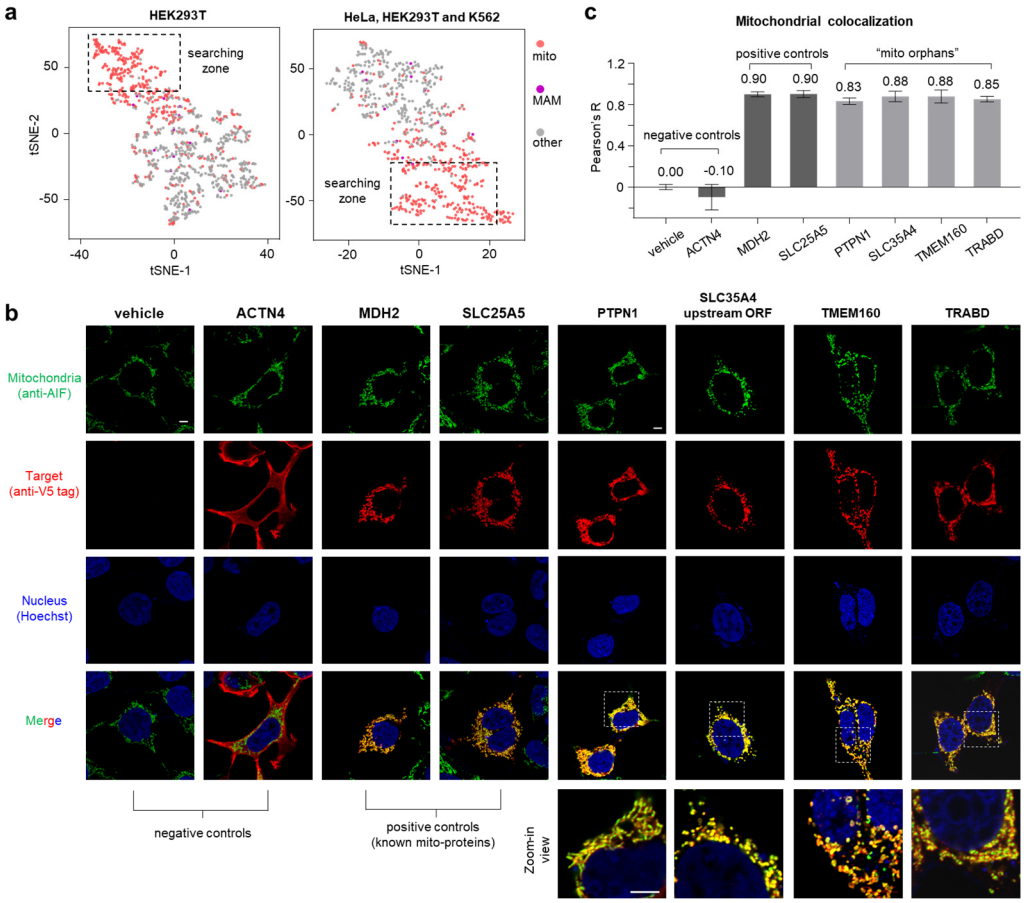

Figure 4. CAT-S-guided discovery of "mito orphans"

a, Visualization of CAT-S datasets including $\log_2$(+/− catalyst) values of individual experiments in two dimension by t-SNE analysis, for HEK293T and combined (3 cell lines) datasets respectively. Proteins are colored based on their prior annotations in database. Red, mitochondrial localization; purple, MAM (mitochondria-associated membrane) or ER-mitochondria contact site localization. Proteins without such annotation but clustered with mitochondrial proteins within "search zone" are potential "mito orphans". **b**, Imaging of recombinant orphan proteins in HEK293T cells. Green, staining of mitochondrial marker AIF. Red, staining of recombinant target protein by C-terminal V5 tag. Blue, staining of nucleus by Hoechst 33342. Scale bar, 5 μm. **c,** Quantification of mitochondrial colocalization (Pearson's R value, n = 12 cells) for "mito orphans" and controls.

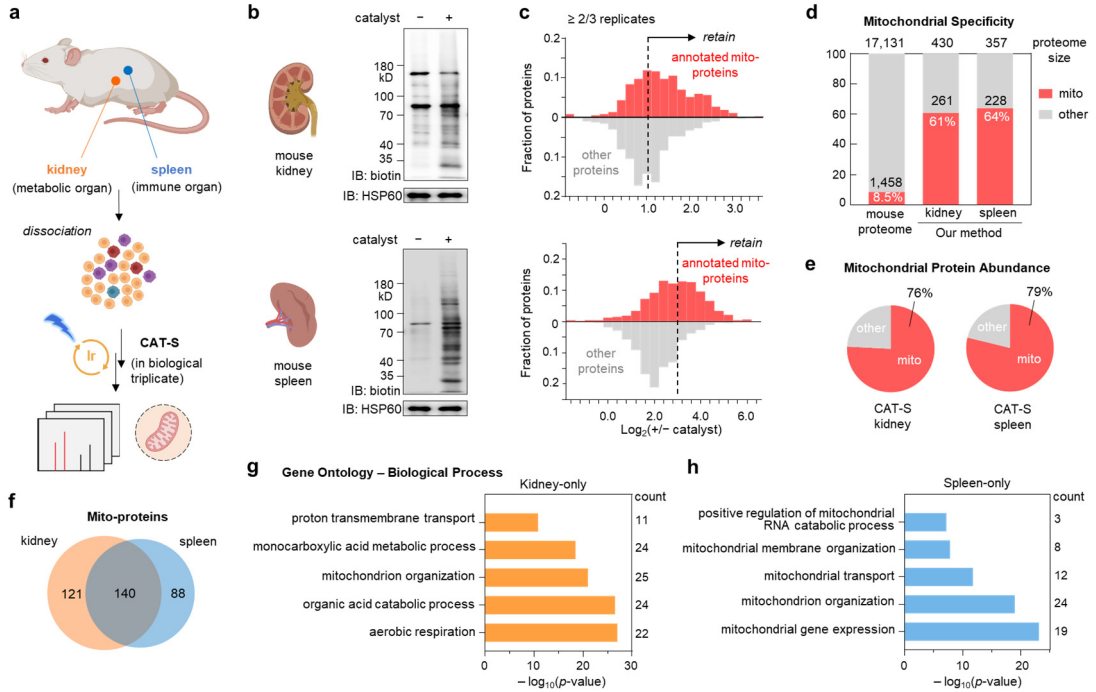

Figure 5. Mitochondrial proteome profiling for primary living tissues

a, Schematic view of CAT-S workflow for mouse tissues. **b-c**, CAT-S-enabled mitochondrial proteome profiling for mouse kidney and spleen, treated with **Ir1** and **SF2** followed by blue LED irradiation. **b**, labeling signals measured by immunoblotting analysis. **c**, frequency distribution profiles of MS-identified proteins (≥ 2/3 replicates) with (red) and without (gray) mitochondrial annotation according to the combined mouse MitoCarta3.0 and UniProt database. A filter based on +/− catalyst ratio (set to 2.0 for kidney and 8.0 for spleen, see also Materials and Methods) was further applied to subtract noise (non-specific proteins). **d**, Mitochondrial specificity analysis for mouse kidney and spleen with the column showing the fraction of proteome with prior mitochondrial annotation. **e**, Contribution of annotated mitochondrial proteins (red) in the total protein MS abundance. **f**, Venn diagram showing the overlap between the captured proteomes. **g-h**, Gene Ontology analysis (Biological Processes) for the mitochondrial proteomes captured only from kidney (**f**) or spleen (**g**). Top 5 enriched biological processes are shown.

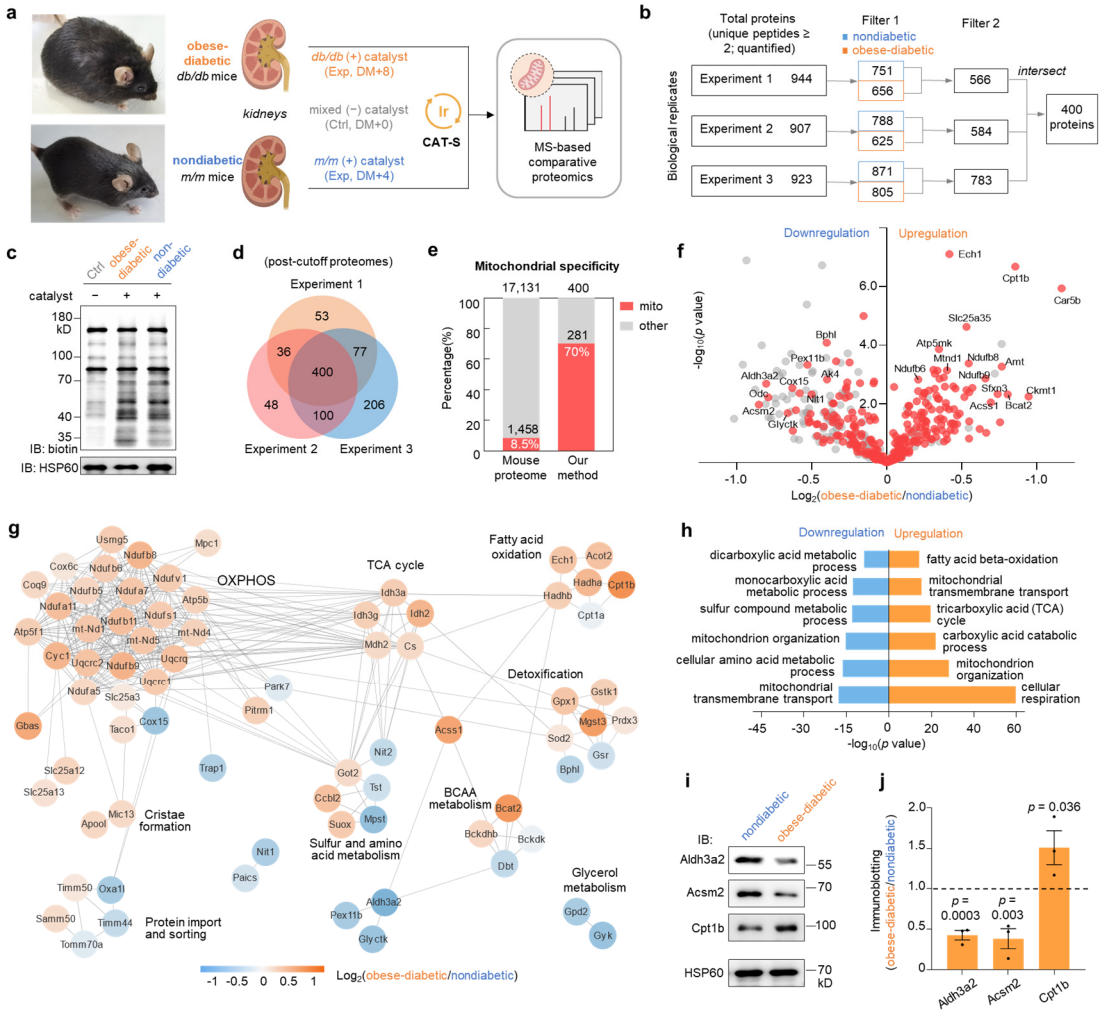

Figure 6. Comparative mitochondrial proteomics for diabetic and healthy mice

a, Schematic view of CAT-S-enabled comparative mitochondrial proteomics for diabetic and nondiabetic mouse kidneys. Samples were treated with **Ir1** and **SF2** followed by blue LED irradiation. **b**, Filtering of MS data to produce protein list for comparative proteomics. First column from left, proteins detected with ≥2 unique peptides and quantified ratios by MS. Second column, retained proteins after cutoff based on diabetic/Ctrl or nondiabetic/Ctrl ratios (set to 1.2) to remove non-biotinylated proteins while maximizing the retention of labeled proteome. The diabetic and nondiabetic protein lists were intersected to produce lists for relative quantification. Lists from three independent experiments were intersected to generate the final proteome. **c**, Labeling signals measured by immunoblotting analysis. **d**, Venn diagram showing the overlap between the comparative proteomes of three experiments. **e**, Mitochondrial specificity analysis with bars showing the fraction of proteome with prior mitochondrial annotation (red) in database. **f**, Volcano plot showing the regulation of proteins in diabetic and nondiabetic mouse kidney. Red dots, proteins with prior mitochondrial annotation. **g**, Cluster map showing regulations of functionally associated molecular complexes according to STRING analysis[50] and Markov clustering. Gray lines denote protein-protein interactions with high confidence (interaction score > 0.7). **h**, Top six enriched biological processes for up- and down-regulated mitochondrial proteins in obese-diabetic mouse kidney by Gene Ontology analysis. **i-j**, Validation for three lipid-metabolism mitochondrial proteins by immunoblotting. **i**, blot images. **j**, intensity quantification. Data are represented as mean ± SEM (n = 3 biological replicates) with p-values shown (unpaired one-tailed t test).

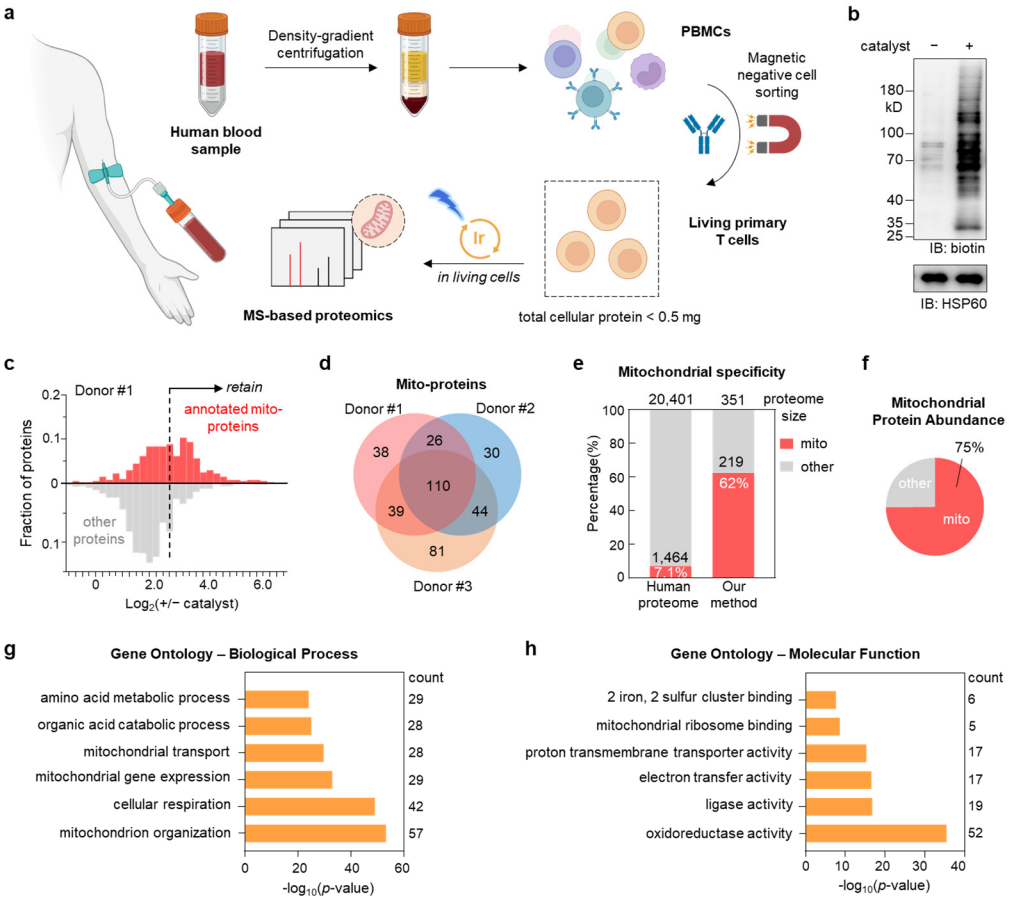

Figure 7. Mitochondrial proteomics for primary T cells from human peripheral blood

a, Schematic view of CAT-S mitochondrial proteomics for living primary T cells from human peripheral blood. Samples were treated with **Ir1** and **SF2** followed by blue LED irradiation. **b**, Labeling signals measured by immunoblotting analysis. **c**, CAT-S-enabled mitochondrial proteome profiling of primary T cells, as exampled by Donor#1. Left, frequency distribution profiles of MS-identified proteins with (red) and without (gray) mitochondrial annotation according to the combined human MitoCarta3.0 and UniProt databases. A filter based on +/− catalyst ratio (set to 7.0) for further noise subtraction (non-specific proteins) is illustrated. **d**, Venn diagram showing the overlap between captured mitochondrial proteomes of the three donors. **e**, Mitochondrial specificity analysis for the captured proteins (detected from two or more donors), with columns showing the fraction of proteome with prior mitochondrial annotation (red). **f**, Contribution of annotated mitochondrial proteins (red) in the total protein MS abundance. **g-h**, Gene Ontology analysis for the captured mitochondrial proteins (detected from two or more donors). Top six enriched biological processes (**f**) and molecular functions (**g**) are shown.