A Visible and Near-Infrared Light Activatable Diazo-Coumarin Probe for Fluorogenic Protein Labeling in Living Cells

Sheng-Yao Dai, Dan Yang*

Supporting Information Placeholder

ABSTRACT: Chemical modification of proteins in living cells permits valuable glimpses into the molecular interactions that underpin dynamic cellular events. While genetic engineering methods are often preferred, selective labeling of endogenous proteins in a complex intracellular milieu with chemical approaches represents a significant challenge. In this study, we report novel diazo-coumarin compounds that can be photo-activated by visible (430–490 nm) and near-infrared light (800 nm) irradiation to photo-uncage reactive carbene intermediates, which could subsequently undergo insertion reaction with concomitant fluorescence “turned-on”. With these new molecules in hand, we have developed a new approach for rapid, selective and fluorogenic labeling of endogenous protein in living cells. By using CA-II and eDHFR as model proteins, we demonstrated that subcellular localization of proteins can be precisely visualized by live-cell imaging and protein levels can be reliably quantified in multiple cell types using flow cytometry. Dynamic protein regulations such as hypoxia induced CA-IX accumulation can also be detected. In addition, by two-photon excitation with an 800 nm laser, cell-selective labeling can also be achieved with spatially controlled irradiation. Our method circumvents the cytotoxicity of UV light and obviates the need for introducing external reporters with “click chemistries”. We believe that this approach of fluorescence labeling of endogenous protein by bioorthogonal photo-irradiation opens up exciting opportunities for discoveries and mechanistic interrogation in chemical biology.

INTRODUCTION

Chemical modification of proteins is a powerful method for protein engineering and conjugation to construct biopharmaceuticals^{1–3}. The ability to selectively label proteins, particularly in living cells, is critical for characterizing protein function, localization and dynamics, but is challenging for chemical biology and drug development^{4–7}. Precise protein labeling can be achieved by genetic engineering approaches, such as fusion with fluorescent proteins, self-labeling tags⁸ (e.g. SNAP-tag⁹, CLIP-tag¹⁰ and HaloTag¹¹) or incorporation of unnatural amino acids (UAAs) by amber codon suppression¹². However, the considerably large fusion proteins (19–33 kDa) may perturb the functions of POIs. The challenges of UAA incorporation in mammalian cells also hindered its wide applications¹³. Most importantly, these methods can only be applied to cell types that are amenable to genetic manipulation. Direct chemical modification of endogenous proteins with small molecules is a prevailing avenue, as it minimally perturbs protein functions owing to the small sizes and bypasses the need for genetic manipulation. However, small molecule-based live-cell protein labeling was only achieved by limited examples including ligand-directed chemistry developed by Hamachi group and others^{14–19}. Their strategies rely on nucleophilic residues on proteins to react with the electrophilic moiety of affinity-based probes, which may intrinsically compromise the labeling efficiency due to restricted amino-acid coverage and competing hydrolysis reaction. Developing new chemistries for selective modification of native endogenous proteins has thus become one of the most important challenges.

Carbene-mediated insertion reactions have been demonstrated to be versatile methods in protein labeling because of their ability to modify a broad scope of residues. One of the most well-known carbene precursors are diazo compounds, which have recently been employed extensively in chemical biology²⁰. For instance, it have

been demonstrated that purified peptides or proteins can be labeled *in vitro* by metallocarbenoid with the use of dirhodium(II)-metallopeptides^{21–29} and ruthenium(II)-porphyrin³⁰ catalysts. Reactive carbene intermediates can also be photochemically generated from diazo and diazirine compounds to label proteins in cells for photoaffinity protein profiling³¹. However, their activation requires UV irradiation that is detrimental to biomolecules and cells via DNA damage^{32–34} and generation of reactive oxygen species³⁵, which hindered the subsequent live-cell applications. How to design new diazo compounds for specific protein modification in live cells in a bioorthogonal manner is still challenging.

Here we present a new photochemical approach by connecting a diazo group to an extended π conjugation system, such as a fluorophore, which allows the excitation wavelength of the diazo group to be drastically red-shifted from 254 nm^{36–37} to visible light region and enables specific intracellular protein labeling in a biocompatible manner without click-chemistry (Scheme 1a). Such protein modification strategy presents several advantages. Firstly, compared to methods utilizing diazirine^{38–39}, benzophenone⁴⁰, arylazide^{41–44}, tetrazole^{45–50} and α -ketoamide⁵¹, the use of long-wavelength (400–800 nm) irradiation can avoid the cytotoxicity of UV light. Secondly, copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) or other reactions that are often required to introduce fluorescence reporters can be obviated, which allows real-time report of labeled proteins. Furthermore, photochemical method provides rapid and non-invasive controls for labeling with spatiotemporal precision^{52–55}.

In our studies, we first analyzed the photolysis reaction of diazo-coumarins under visible light irradiation and further optimized the structure by introducing a trifluoromethyl substituent adjacent to the diazo group to improve the carbene insertion propensity. Covalent modification of purified proteins by diazo-coumarin can be accomplished with mild irradiation of blue LED lamp. It was con-

firmed that a range of amino acids can be covalently modified. Directed by small-molecule ligands, labeling of specific proteins was further advanced into living cells. Notably, the protein labeling in cells was found to be rapid, mild and specific. The coumarin fluorescence was recovered upon photo-irradiation to facilitate the instant report of labeled proteins in a variety of analyses, including SDS-PAGE, confocal microscopy and flow cytometry. Furthermore, *in situ* two-photon labeling was achieved in cells with a near-infrared laser, providing an exciting potential for achieving protein labeling with spatiotemporal control. Collectively, we envision that diazo-coumarin probes can be exploited as a new photochemical tool for live-cell protein modifications to facilitate the future biological studies, and also provide useful templates for the rational design of next generation protein modification reactions.

Scheme 1. (a) Schematic illustration of light-induced fluorogenic labeling of protein-of-interest (POI). (b) Photo-induced carbene insertion reaction of diazo-coumarins **1 and **2**.**

RESULTS AND DISCUSSION

Photochemical Characterization of Diazo-coumarins Under Blue Light. In the course of surveying different combinations of diazo group and fluorophores, we found that 4-diazomethylcoumarins⁵⁶⁻⁵⁹ were used as precursors for photo-releasing of compounds⁶⁰, but the photochemistry of those diazo compounds has been unexplored. We hypothesize that diazo-coumarin compounds can be activated by photo-irradiation with long-wavelength light to “uncage” reactive carbene species for covalent modification of proximal proteins (Scheme 1b). To test whether diazo-coumarins can be photo-activated by visible light, we synthesized compound **1** according to procedures reported in the literature (Scheme S1)⁵⁸. It was found that a large portion of the longest absorption peak fell into the visible light region (> 400 nm) (Figure S1a). Unlike ordinary coumarins, the fluorescence of **1** was found to be quenched by the diazo group, presumably via internal charge transfer (ICT)⁶¹. Consequently, photolysis of **1** was first carried out in methanol by irradiation with a household blue LED lamp (430–490 nm). Upon irradiation, changes in absorption and fluorescence spectra were observed (Figure S1a–b), indicating a photochemical transformation of **1** under blue light as we anticipated. It is highly conceivable that a carbene intermediate is formed and trapped by methanol. The reaction mixture was analyzed by LCMS, and multiple products were observed (Figure S1c). To identify these products, a large scale photolysis reaction was performed. The photolysis mixture was found to include several solvent insertion products (Scheme 2, Figure S1d), such as the methanol C–H insertion product **1a** (2.3%), the water insertion product **1b** (3.3%) and the methanol O–H insertion product **1d** (19%). While an aldehyde product **1c** resulting from the carbene oxidation by oxygen⁶² was isolated in 35%

yield, the major product was found to be a non-fluorescent ring expansion product **1e** (38%)⁶³, formed by Wolff rearrangement of the carbene intermediate to an allene⁶⁴ followed by methanol addition. This process is undesirable since the allene intermediate has a long lifetime and preferentially reacts with nucleophiles to give non-fluorescent products.

Scheme 2. Photolysis of **2 (30 μ M) in methanol upon blue LED irradiation and the isolated yields of photolysis products.**

Since trifluoromethyl substituent can prevent intramolecular rearrangement of carbene intermediate generated from trifluoromethylaryldiazirine⁶⁵, we designed and synthesized a trifluoromethyl analog **2** (Scheme S2). Compound **2** was found to have a longer absorption maximum positioned at 408 nm (extinction coefficient ϵ of 2259 M⁻¹cm⁻¹, Table S1) and was non-fluorescent (quantum yield Φ_f of 0.003). When **2** was subjected to blue LED irradiation, the peak at 408 nm only underwent a slight blue-shift and that at ~300 nm dropped sharply. This indicates the coumarin chromophore was preserved, whereas the diazo moiety that usually absorbs below 300 nm underwent decomposition. A significant fluorescence enhancement of 135-fold was also detected (Figure 1b,c). Subsequent analyses (Figure 1d,e) revealed that most products were fluorescent, including insertion products **2b** (17%) and **2d** (43%; Φ_f of 0.22), ketone product **2a** (15%), and cyclized⁶⁶ by-product **2c** (10%) (Figure S2, Table S1). Most importantly, no ring-expansion was found.

Figure 1. Photo-chemical characterizations of diazo-coumarins. (a)–(b) Change in absorption and emission spectra of **2** (30 μ M) in methanol upon blue LED irradiation. $\lambda_{ex} = 400$ nm. (c) Change in fluorescence intensity at $\lambda_{em} = 484$ nm upon blue LED irradiation. (d) HPLC analysis of **2** before and after photolysis. (e) Isolated yields of photolysis products.

Reactivity and Stability of Diazo-coumarins in Biologically Relevant Conditions. Encouraged by these results, we performed the blue light-prompted reaction of model protein bovine serum albumin (BSA) with **1** or **2**. The fluorogenic property of coumarin facilitated the visualization of proteins during in-gel fluorescence scanning without the need for “click” chemistry. While both compounds can covalently label BSA at high concentrations, **2** ap-

peared to be more efficient (Figure 2a) than **1**. Through tryptic digestion and tandem mass spectrometry (LC/MS/MS) analysis (Figures 2b, c, and S3), 35 and 8 modification sites by **1** and **2** were identified on BSA, respectively. It was found that although the two compounds modified similar regions, **2** labeled neutral and polar side-chains without strong preference, suggesting the formation of a reactive carbene intermediate. However, **1** additionally reacted with a large number of nucleophilic residues, such as His, Lys, Ser, Arg, Thr and Tyr, indicative of the generation of ketene intermediate along with the reactive carbene, which could potentially lead to non-specific labeling⁶⁷ in a more complex environment. These results suggested that **2** is superior to **1** in terms of sensitivity and selectivity in protein labeling. Since diazo groups are prone to decomposition, especially towards thiols⁶⁸ and acids⁶⁹, the stability of **2** was further examined. When **2** was incubated with an equal amount of reduced glutathione (GSH), an intracellularly abundant nucleophile and reductant, in phosphate buffer for 24 h in dark, no reaction was observed by NMR analysis (Figure 2d). In addition, as shown in Figure 2e, **2** was stable across a wide range of pH; the fluorescence increased only when the compound was exposed to light.

Figure 2. Reactivity and stability of diazo-coumarins. (a) SDS-PAGE of BSA (15 μM) labeled by **1** and **2** of various concentrations. (b)–(c) Bar chart showing the amino acid labeling frequencies of **1** and **2** obtained from tryptic peptide LC/MS/MS analysis of BSA. (d) Comparison of ¹H NMR spectra of **2** (5 mM) before and after 24 h incubation in sodium phosphate buffer (10 mM), pH 7.4 (D₂O)/CD₃OD 1:1, in the dark at room temperature, in the presence of reduced glutathione (5 mM). (e) Fluorescence intensity of **2** ($\lambda_{\text{ex}} = 400$ nm, $\lambda_{\text{em}} = 510$ nm) in sodium phosphate buffer of various pH values incubated in dark or under blue LED irradiation as indicated.

Labeling of Recombinant CA-II with Blue Light. The affinity-based approach was then employed to achieve specific protein labeling using small-molecule ligands. Carbonic anhydrase-II (CA-II, 29 kDa) was chosen as a target for selective labeling as it has been widely adopted in studies of affinity-based probes⁷⁰. CA-II can be reversibly inhibited by aromatic primary sulfonamides⁷¹, such as 4-carboxybenzene-sulfonamide (**CBS**) ($K_d \sim 3.2$ μM)⁷². Accordingly, we designed and synthesized **3** (Figure 3a, Scheme S2) by modifying one of the ethyl substituents on the 7-amino group to connect **CBS** through a simple amide linker. *In vitro* labeling of recombinant CA-II was then performed by mixing with **3** followed by blue LED irradiation (Figure 3b). As visualized by the resulting SDS-PAGE, labeling of CA-II by **3** was furnished in both

concentration (Figure 3c, EC₅₀: 2.48 μM) and time (Figure 3d) dependent manner, and was diminished by competition with an excess of **CBS** (Figure 3e, IC₅₀: 182 μM). An enzyme activity assay was performed, which showed that **3** has an IC₅₀ value comparable to that of **CBS** (Figure 3f, IC₅₀ of **3**: 0.90 μM, IC₅₀ of **CBS**: 2.14 μM). These results showed that the binding property of the sulfonamide was not disrupted despite its modest interaction with CA-II. Further, LC/MS/MS analysis revealed that **3** labeled His2 and Trp4 residues of the N-terminal peptide of CA-II (Figure 3g, Figure S4). Both residues were located at the entrance of the ligand-binding pocket (Figure 3h), suggesting that labeling is induced by the specific interaction between CA-II and **3**. Consistently, exclusive labeling of 29 kDa CA-II could also be observed in HeLa cell lysates spiked with the recombinant protein (Figure S5a).

Figure 3. Fluorescence labeling of recombinant CA-II with blue light. (a) Structures of 4-carboxybenzenesulfonamide (**CBS**) and **3**. (b) Schematic illustration of blue LED light-induced labeling of recombinant CA-II. (c)–(e) SDS-PAGE of recombinant CA-II (1.7 μM) labeled by **3** (10 μM) in (c) and (e) under blue LED irradiation. (f) Enzymatic activity assay of CA-II catalyzed the hydrolysis of *p*-nitrophenyl acetate. Error bar indicates SD, n = 3. (g) Mass spectrum of modified CA-II tryptic peptide by **3**. (h) Crystal structure of CA-II showing Zn²⁺ ion and the two labeled residues.

Labeling of Endogenous CA-II in Living Cells. Our ultimate goal is to label endogenously expressed proteins in living cells. Therefore, labeling was attempted with live MCF7 human breast cancer cells⁷³ and HCT116 human colorectal cancer cells⁷⁴, which endogenously express CA-II. After incubation with **3**, cells were irradiated with the blue LED lamp (Figure 4a). Subsequently, cell lysates were harvested and separated by SDS-PAGE to visualize the fluorescent signal. A strong single fluorescent band corresponds to CA-II was observed in SDS-PAGE of MCF7 cells, but

was blocked by **CBS** co-treatment (Figure 4b, lanes 3 and 4). A weaker band was observed without irradiation (Figure 4b, lane 2), suggesting that a minor reaction occurred in the dark. Western blot analysis confirms that CA-II was evenly expressed among the four groups and matched the molecular weight of fluorescent bands. Similar results were obtained in HCT116 cells (Figure S5b,c). Surprisingly, in the live-cell time-dependent experiments, the labeling was found to complete within 2 mins under irradiation (Figure S5d). The faster labeling in living cells is presumably attributed to the enriched concentrations of **3** and the intracellular proteins when confined within cells. When the CA-II expression of MCF7 cells was knocked down by siRNA, the in-gel fluorescent signal disappeared (Figure 4c). Taken altogether, these results convincingly demonstrated the visible-light-induced specific and rapid labeling of CA-II by diazo-coumarin in living cells.

Figure 4. Fluorescence labeling of endogenous CA-II in cells. (a) Schematic illustration of live-cell labeling. (b) SDS-PAGE of MCF7 live-cell labeling by **3** (10 μM) in the presence/absence of **CBS** (200 μM). (c) SDS-PAGE and western blot of MCF7 live-cell labeling by **3** (10 μM) after CA-II knockdown by siRNA. (d) Live-cell confocal imaging of HCT116 cells treated with **3** (5 μM), followed by blue LED irradiation. **3**: λ_{ex} = 405 nm, NucRed: λ_{ex} = 633 nm. Scale bar = 10 μm.

Due to the unique fluorogenic property of the diazo-coumarin probe **3** and biocompatibility of the labeling condition, live-cell imaging of the intracellular CA-II can be performed, which is otherwise difficult to achieve with other methods. After the treatment of **3** with blue LED irradiation, strong fluorescence was observed in the cytosolic region of HCT116 cells (Figure 4d) and MCF7 cells (Figure S6a). In addition, the fluorescent signal was diminished when **CBS** was added to compete with **3**. When a washout step was

performed prior to the irradiation, minimal labeling was observed (Figure S6b). The fluorescence signals were clearly excluded from the nucleus that was stained by the red nucleus probe (NucRed), which was consistent with the reported cytosolic distribution of CA-II⁷⁵. Immunofluorescence staining showed that the signal from **3** was highly co-localized with that of CA-II antibody (Figure S6c). Moreover, the fact that the fluorescence signal of **3** was retained after cell permeabilization is also an evidence that the covalent bonding was established in response to light.

By using fluorescence-activated cell sorting (FACS) analysis, the fluorogenic effect and competition can be quantitatively assessed (Figure 5a). As recent work has demonstrated the assessment of the expression level of the cannabinoid receptors by tandem live-cell photoaffinity labeling and click chemistry on fixed cells⁷⁶, we envisage that such an assessment can be carried out in living cells with our approach. As shown in Figure 5b, the strongest specific labeling for CA-II was obtained in MCF7 cells, followed by HCT116 cells, whereas HeLa cells showed minimal labeling. This result suggests that the expression level of CA-II should follow this order: MCF7 > HCT116 > HeLa, which was validated by Western blot analysis (Figure 5c).

Figure 5. FACS analysis of CA-II expression in different cell types. (a) FACS analysis of cells treated with: (i) **3** (5 μM), blue LED irradiation; (ii) **3** (5 μM) and **CBS** (100 μM), blue LED irradiation; (iii) **3** (5 μM), no irradiation; (iv) untreated. (b) Relative CA-II signals obtained from FACS analysis in MCF7, HCT116 and HeLa cells. (c) Western blot of CA-II in three cell lines.

Labeling of CA-IX under Hypoxia-mimetic Condition. CA-IX is a membrane-associated isoform of CAs, which is markedly overexpressed in hypoxic tumor cells⁷⁷. As it plays a pivot role in malignant progression, especially in regulating pH homeostasis under acidic extracellular matrix, CA-IX is thus regarded as a valuable biomarker as well as an attractive target in cancer therapy⁷⁸. The expression of CA-IX is modulated by hypoxia-inducible factor 1 (HIF-1), which under normoxia condition is hydroxylated at proline residues by prolyl hydroxylase domain-containing protein 2 (PHD2) and subsequently ubiquitinated by Von Hippel-Lindau (VHL) E3 ligase for proteasome degradation (Figure 6a)⁷⁹. Since oxygen-sensing of PHD2 relies on its catalytic iron(II) center, iron chelator deferoxamine mesylate (**DFO**) can inhibit PHD2 to stabilize HIF-1, thus acting as a hypoxia-mimetic agent⁸⁰. In our experiment, HeLa cells, which has no CA-II expression, were cultured under normal or hypoxia-mimetic conditions (treated with **DFO**). From the western blot result shown in Figure 6b, it can be observed that CA-IX was upregulated as a result of HIF-1 accumulation. After that, these cells were treated with **3** and followed by irradiation with blue LED light. The resulting live-cell confocal images

showed that substantial signals accumulated on the plasma membrane in cells treated with **DFO**, while untreated cells displayed minimal signal (Figure 6c). The majority of the signals could be displaced by co-treatment of **CBS**. This evidently indicated that **3** could fluorescently label the membrane-bound CA-IX under hypoxia-mimetic condition. Our results demonstrated that our visible light activated probe could be used for labeling proteins that are dynamically expressed under specific condition, which is otherwise hard to be accomplished with overexpression system of genetically modification approach.

Figure 6. Fluorescence labeling of CA-IX in cells under hypoxia-mimetic condition. (a) Schematic illustration of inducible CA-IX expression under hypoxia or treatment by **DFO**. (b) Western blot of HeLa cells cultured under normal or hypoxia-mimetic (**DFO**, 100 μ M for 24 h) conditions. (c) Live-cell confocal imaging of HeLa cells cultured under normal or hypoxia-mimetic (**DFO**, 100 μ M for 24 h) conditions. Cells were treated with **3** (5 μ M) in the presence/absence of **CBS** (100 μ M), followed by blue LED irradiation. **3**: λ_{ex} = 405 nm. Scale bar = 10 μ m.

In situ Two-photon Labeling and Imaging. Since 7-aminocoumarin has been applied in two-photon uncaging⁸¹, we wonder if two-photon protein labeling can be performed. In principle, the ability of two-photon activation would potentially lead to reduced photo-toxicity, better spatial control, and deeper tissue penetration⁸². We, therefore, tested the feasibility of two-photon activation of **2** to form carbene with the use of near-infrared (NIR) light to initiate the reaction⁸³⁻⁸⁴. When **2** was exposed to the femtosecond pulses of an 800 nm focused laser beam, a similar photo-chemical response to that of blue light irradiation was observed (Figure 7a). Dependence of the fluorescence increase on the focused laser power was recorded, however, no reaction occurred when the laser was unfocused (Figure 7b). These results suggest that **2** can be two-photon activated by NIR light. As a proof-of-concept experiment, *in situ* two-photon labeling and imaging of CA-II was performed in living cells with **3** (Figure 7c). In this experiment, MCF7 cells were treated with **3** and the cell location was first tracked with NucRed under confocal microscopy. Next, only selected cells were irradiated with a laser at 800 nm of the microscope. Followed by the washing of excess probes, the cells were imaged. As shown in Figure 7c, stronger fluorescence signals were detected in cells exposed to NIR light, indicating that two-photon activated labeling of CA-II can be achieved in live cells.

Figure 7. *In situ* two-photon labeling and imaging of CA-II. (a) Change in emission spectra of **2** (30 μ M) upon two-photon photolysis with a mode-locked Ti:sapphire 800 nm femtosecond pulses focused laser at 200 mW. Emission was measured at λ_{ex} = 400 nm. (b) Fluorescence intensity (λ_{ex} = 400 nm, λ_{ex} = 484 nm) measured at 30 min after irradiation with focused or unfocused laser beam of varying laser power. (c) Schematic illustration of two-photon induced labeling of CA-II with **3** (5 μ M) in living cells and confocal imaging of MCF7 cells before and after two-photon activation with 800 nm. Scale bar = 20 μ m.

Figure 8. Fluorescence labeling of mCherry-eDHFR in HeLa cells. (a) Chemical Structures of TMP and **4**. (b) Live-cell confocal imaging of HeLa cells with or without transient expression of mCherry-eDHFR, treated with **4** (5 μ M) and TMP (50 μ M) after blue LED irradiation. **4**: λ_{ex} = 405 nm, mCherry: λ_{ex} = 561 nm. Scale bar = 10 μ m.

Labeling of eDHFR in Living Cells. To demonstrate that our labeling strategy can be applied to other target proteins, **4** was synthesized by incorporating trimethoprim (**TMP**, K_d ~ 10 nM) as a

ligand moiety (Figure 8a)⁸⁵, which can selectively bind *E. coli* dihydrofolate reductase (eDHFR). HeLa cells were transfected plasmid DNA to overexpress eDHFR fused with a red fluorescent protein (mCherry-eDHFR). The transfection of this fusion protein helps to confirm the co-localization with diazo-coumarin in living cells under confocal microscopy. When the cells were treated with **4** (5 μ M) and irradiated with blue light, co-localization of the probe signal and mCherry was observed in live-cell imaging (Figure 8b). While the cells without transfection did not show any staining, the addition of **TMP** also predictably perturbed the labeling in transfected cells.

CONCLUSION

In the investigation described above, we have demonstrated that diazo-coumarin can be activated by visible light to “photo-uncage” reactive carbene intermediate, which can subsequently undergo insertion reaction with biomolecules. Through detailed chemical analysis, the structure of diazo-coumarin was optimized by the introduction of a trifluoromethyl group to improve the carbene insertion propensity. Moreover, we showed that efficient visible light photo-activation and fluorogenic property of diazo-coumarin can be exploited as a biorthogonal tool in endogenous protein labeling. As a result, proteins-of-interest can be covalently modified after mild blue LED irradiation, with instant fluorescence report bypassing the need to perform “click” chemistry. We demonstrated the applications of diazo-coumarin in affinity labeling of endogenous proteins in living cells using CA-II as a model protein. Remarkable labeling speed and specificity of endogenous CA-II has been achieved, which enables reliable visualization of localization by live-cell imaging and quantitative assessment of expression level among different cell types by flow cytometry. Live-cell imaging also allowed monitoring of protein dynamics as proved by hypoxia induced CA-IX accumulation. Furthermore, diazo-coumarin can be photo-activated by two-photon excitation, which allows spatiotemporally controlled protein labeling in living cells with the least phototoxic NIR light. As a further illustration, this strategy can also be successfully applied to the labeling of eDHFR. We believe that this photo-chemical approach should expand the chemical repertoire of the covalent labeling of biomolecules in living cells to facilitate biological discoveries.

ASSOCIATED CONTENT

Supporting Information

Supporting figures, tables, experimental procedures, synthesis, characterization of compounds, LC/MS/MS spectra and NMR spectra.

AUTHOR INFORMATION

Corresponding Author

Dan Yang – Morningside Laboratory for Chemical Biology, Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China; Email: yangdan@hku.hk

Authors

Sheng-Yao Dai – Morningside Laboratory for Chemical Biology, Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China

ACKNOWLEDGMENT

We thank Dr. Ming-De Li (Shantou University) and Prof. David Philip (The University of Hong Kong) for help in the two-photon photolysis experiment setup. We thank Dr. Fang-Fang Shen and Dr. Xiang David Li (The University of Hong Kong) as well as Dr.

Nai-Kei Wong (The Second Hospital Affiliated to Southern University of Science and Technology) for helpful advice on experiments and critical reading of the manuscript. We thank The University of Hong Kong Li Ka Shing Faculty of Medicine Faculty Core Facility for support in confocal microscopy and flow cytometry. We thank Dr. Ruijun Tian (Southern University of Science and Technology) for assistance in LC/MS/MS experiment. This work was supported by The University of Hong Kong, the University Development Fund, the Morningside Foundation, and the Hong Kong Research Grants Council under the Area of Excellence Scheme (AoE/P-705/16).

REFERENCES

- Hoyt, E. A.; Cal, P. M. S. D.; Oliveira, B. L.; Bernardes, G. J. L., Contemporary approaches to site-selective protein modification. *Nat. Rev. Chem.* **2019**, *3*, 147-171.
- Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L., Site-selective protein-modification chemistry for basic biology and drug development. *Nat. Chem.* **2015**, *8*, 103-113.
- Boutureira, O.; Bernardes, G. J., Advances in chemical protein modification. *Chem. Rev.* **2015**, *115*, 2174-95.
- Tamura, T.; Hamachi, I., Chemistry for Covalent Modification of Endogenous/Native Proteins: From Test Tubes to Complex Biological Systems. *J. Am. Chem. Soc.* **2019**, *141*, 2782-2799.
- Amaike, K.; Tamura, T.; Hamachi, I., Recognition-driven chemical labeling of endogenous proteins in multi-molecular crowding in live cells. *Chem. Commun.* **2017**, *53*, 11972-11983.
- Kubota, R.; Hamachi, I., Protein recognition using synthetic small-molecular binders toward optical protein sensing in vitro and in live cells. *Chem. Soc. Rev.* **2015**, *44*, 4454-71.
- Zhang, G.; Zheng, S.; Liu, H.; Chen, P. R., Illuminating biological processes through site-specific protein labeling. *Chem. Soc. Rev.* **2015**, *44*, 3405-17.
- Wang, Z.; Ding, X.; Li, S.; Shi, J.; Li, Y., Engineered fluorescence tags for in vivo protein labelling. *RSC Adv.* **2014**, *4*, 7235-7245.
- Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K., Directed Evolution of O⁶-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chem. Biol.* **2003**, *10*, 313-317.
- Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K., An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **2008**, *15*, 128-136.
- Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3*, 373-382.
- Lang, K.; Chin, J. W., Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **2014**, *114*, 4764-806.
- Lin, X.; Yu, A. C.; Chan, T. F., Efforts and Challenges in Engineering the Genetic Code. *Life (Basel)* **2017**, *7*.
- Tamura, T.; Ueda, T.; Goto, T.; Tsukidate, T.; Shapira, Y.; Nishikawa, Y.; Fujisawa, A.; Hamachi, I., Rapid labelling and covalent inhibition of intracellular native proteins using ligand-directed N-acyl-N-alkyl sulfonamide. *Nat. Commun.* **2018**, *9*, 1870.
- Matsuo, K.; Nishikawa, Y.; Masuda, M.; Hamachi, I., Live-Cell Protein Sulfonation Based on Proximity-driven N-Sulfonyl Pyridone Chemistry. *Angew. Chem. Int. Ed.* **2018**, *57*, 659-662.
- Yamaguchi, T.; Asanuma, M.; Nakanishi, S.; Saito, Y.; Okazaki, M.; Dodo, K.; Sodeoka, M., Turn-ON fluorescent affinity

- labeling using a small bifunctional O-nitrobenzoxadiazole unit. *Chem. Sci.* **2014**, *5*, 1021-1029.
17. Fujishima, S. H.; Yasui, R.; Miki, T.; Ojida, A.; Hamachi, I., Ligand-directed acyl imidazole chemistry for labeling of membrane-bound proteins on live cells. *J. Am. Chem. Soc.* **2012**, *134*, 3961-4.
 18. Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I., Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.* **2009**, *5*, 341-3.
 19. Hughes, C. C.; Yang, Y.-L.; Liu, W.-T.; Dorrestein, P. C.; Clair, J. J. L.; Fenical, W., Marinopyrrole A Target Elucidation by Acyl Dye Transfer. *J. Am. Chem. Soc.* **2009**, *131*, 12094-12096.
 20. Mix, K. A.; Aronoff, M. R.; Raines, R. T., Diazo Compounds: Versatile Tools for Chemical Biology. *ACS Chem. Biol.* **2016**, *11*, 3233-3244.
 21. Ohata, J.; Ball, Z. T., A Hexa-rhodium Metallopeptide Catalyst for Site-Specific Functionalization of Natural Antibodies. *J. Am. Chem. Soc.* **2017**, *139*, 12617-12622.
 22. Martin, S. C.; Vohidov, F.; Wang, H.; Knudsen, S. E.; Marzec, A. A.; Ball, Z. T., Designing Selectivity in Dirhodium Metallopeptide Catalysts for Protein Modification. *Bioconjug. Chem.* **2017**, *28*, 659-665.
 23. Vohidov, F.; Coughlin, J. M.; Ball, Z. T., Rhodium(II) metallopeptide catalyst design enables fine control in selective functionalization of natural SH3 domains. *Angew. Chem. Int. Ed.* **2015**, *54*, 4587-91.
 24. Minus, M. B.; Liu, W.; Vohidov, F.; Kasembeli, M. M.; Long, X.; Krueger, M. J.; Stevens, A.; Kolosov, M. I.; Tweardy, D. J.; Sison, E. A. R.; Redell, M. S.; Ball, Z. T., Rhodium(II) Proximity-Labeling Identifies a Novel Target Site on STAT3 for Inhibitors with Potent Anti-Leukemia Activity. *Angew. Chem. Int. Ed.* **2015**, *54*, 13085-13089.
 25. Kundu, R.; Ball, Z. T., Rhodium-catalyzed cysteine modification with diazo reagents. *Chem. Commun.* **2013**, *49*, 4166-8.
 26. Chen, Z.; Vohidov, F.; Coughlin, J. M.; Stagg, L. J.; Arold, S. T.; Ladbury, J. E.; Ball, Z. T., Catalytic protein modification with dirhodium metallopeptides: specificity in designed and natural systems. *J. Am. Chem. Soc.* **2012**, *134*, 10138-45.
 27. Popp, B. V.; Ball, Z. T., Proximity-driven metallopeptide catalysis: Remarkable side-chain scope enables modification of the Fos bZip domain. *Chem. Sci.* **2011**, *2*, 690-695.
 28. Chen, Z.; Popp, B. V.; Bovet, C. L.; Ball, Z. T., Site-specific protein modification with a dirhodium metallopeptide catalyst. *ACS Chem. Biol.* **2011**, *6*, 920-5.
 29. Popp, B. V.; Ball, Z. T., Structure-Selective Modification of Aromatic Side Chains with Dirhodium Metallopeptide Catalysts. *J. Am. Chem. Soc.* **2010**, *132*, 6660-6662.
 30. Ho, C.-M.; Zhang, J.-L.; Zhou, C.-Y.; Chan, O.-Y.; Yan, J. J.; Zhang, F.-Y.; Huang, J.-S.; Che, C.-M., A Water-Soluble Ruthenium Glycosylated Porphyrin Catalyst for Carbenoid Transfer Reactions in Aqueous Media with Applications in Bioconjugation Reactions. *J. Am. Chem. Soc.* **2010**, *132*, 1886-1894.
 31. Ge, S.-S.; Chen, B.; Wu, Y.-Y.; Long, Q.-S.; Zhao, Y.-L.; Wang, P.-Y.; Yang, S., Current advances of carbene-mediated photoaffinity labeling in medicinal chemistry. *RSC Advances* **2018**, *8*, 29428-29454.
 32. Ridley, A. J.; Whiteside, J. R.; McMillan, T. J.; Allinson, S. L., Cellular and sub-cellular responses to UVA in relation to carcinogenesis. *Int. J. Radiat Biol.* **2009**, *85*, 177-95.
 33. McMillan, T. J.; Leatherman, E.; Ridley, A.; Shorrocks, J.; Tobin, S. E.; Whiteside, J. R., Cellular effects of long wavelength UV light (UVA) in mammalian cells. *J. Pharm. Pharmacol.* **2008**, *60*, 969-76.
 34. Barlev, A.; Sen, D., DNA's Encounter with Ultraviolet Light: An Instinct for Self-Preservation? *Acc. Chem. Res.* **2018**, *51*, 526-533.
 35. Bai, X.; Huang, Y.; Lu, M.; Yang, D., HKOH-1: A Highly Sensitive and Selective Fluorescent Probe for Detecting Endogenous Hydroxyl Radicals in Living Cells. *Angew. Chem. Int. Ed.* **2017**, *56*, 12873-12877.
 36. Chowdhry, V.; Vaughan, R.; Westheimer, F. H., 2-diazo-3,3,3-trifluoropropionyl chloride: reagent for photoaffinity labeling. *Proc. Natl. Acad. Sci. U.S.A* **1976**, *73*, 1406-1408.
 37. Singh, A.; Thornton, E. R.; Westheimer, F. H., The Photolysis of Diazoacetylchymotrypsin. *J. Biol. Chem.* **1962**, *237*, 3006-3008.
 38. Tomohiro, T.; Morimoto, S.; Shima, T.; Chiba, J.; Hatanaka, Y., An Isotope-Coded Fluorogenic Cross-Linker for High-Performance Target Identification Based on Photoaffinity Labeling. *Angew. Chem. Int. Ed.* **2014**, *53*, 13502-5.
 39. Tomohiro, T.; Yamamoto, A.; Tatsumi, Y.; Hatanaka, Y., [3-(Trifluoromethyl)-3H-diazirin-3-yl]coumarin as a carbene-generating photocross-linker with masked fluorogenic beacon. *Chem. Commun.* **2013**, *49*, 11551-3.
 40. Murale, D. P.; Hong, S. C.; Yun, J.; Yoon, C. N.; Lee, J. S., Rational design of a photo-crosslinking BODIPY for in situ protein labeling. *Chem. Commun.* **2015**, *51*, 6643-6.
 41. Wang, Y.; Hu, L.; Xu, F.; Quan, Q.; Lai, Y.-T.; Xia, W.; Yang, Y.; Chang, Y.-Y.; Yang, X.; Chai, Z.; Wang, J.; Chu, I. K.; Li, H.; Sun, H., Integrative approach for the analysis of the proteome-wide response to bismuth drugs in *Helicobacter pylori*. *Chem. Sci.* **2017**, *8*, 4626-4633.
 42. Chao, A.; Jiang, N.; Yang, Y.; Li, H.; Sun, H., A Ni-NTA-based red fluorescence probe for protein labelling in live cells. *J. Mater. Chem. B* **2017**, *5*, 1166-1173.
 43. Lai, Y. T.; Chang, Y. Y.; Hu, L.; Yang, Y.; Chao, A.; Du, Z. Y.; Tanner, J. A.; Chye, M. L.; Qian, C.; Ng, K. M.; Li, H.; Sun, H., Rapid labeling of intracellular His-tagged proteins in living cells. *Proc Natl Acad Sci U S A* **2015**, *112*, 2948-53.
 44. Chiba, K.; Asanuma, M.; Ishikawa, M.; Hashimoto, Y.; Dodo, K.; Sodeoka, M.; Yamaguchi, T., Specific fluorescence labeling of target proteins by using a ligand-4-azidophthalimide conjugate. *Chem. Commun.* **2017**, *53*, 8751-8754.
 45. Tian, Y.; Lin, Q., Genetic encoding of 2-aryl-5-carboxytetrazole-based protein photo-cross-linkers. *Chem. Commun.* **2018**, *54*, 4449-4452.
 46. Tian, Y.; Jacinto, M. P.; Zeng, Y.; Yu, Z.; Qu, J.; Liu, W. R.; Lin, Q., Genetically Encoded 2-Aryl-5-carboxytetrazoles for Site-Selective Protein Photo-Cross-Linking. *J. Am. Chem. Soc.* **2017**, *139*, 6078-6081.
 47. Cheng, K.; Lee, J. S.; Hao, P.; Yao, S. Q.; Ding, K.; Li, Z., Tetrazole-Based Probes for Integrated Phenotypic Screening, Affinity-Based Proteome Profiling, and Sensitive Detection of a Cancer Biomarker. *Angew. Chem. Int. Ed.* **2017**, *56*, 15044-15048.
 48. Zhao, S.; Dai, J.; Hu, M.; Liu, C.; Meng, R.; Liu, X.; Wang, C.; Luo, T., Photo-induced coupling reactions of tetrazoles with carboxylic acids in aqueous solution: application in protein labelling. *Chem. Commun.* **2016**, *52*, 4702-5.
 49. Li, Z.; Qian, L.; Li, L.; Bernhammer, J. C.; Huynh, H. V.; Lee, J. S.; Yao, S. Q., Tetrazole Photoclick Chemistry: Reinvestigating Its Suitability as a Bioorthogonal Reaction and Potential Applications. *Angew. Chem. Int. Ed.* **2016**, *55*, 2002-6.
 50. Herner, A.; Marjanovic, J.; Lewandowski, T.; Marin, V. L.; Patterson, M. J.; Miesbauer, L.; Ready, D.; Williams, J.; Vasudevan, A.; Lin, Q., 2-Aryl-5-carboxytetrazole as a New Photoaffinity Label for Drug Target Identification. *J. Am. Chem. Soc.* **2016**, *138*, 14609-14615.
 51. Ota, E.; Usui, K.; Oonuma, K.; Koshino, H.; Nishiyama, S.; Hirai, G.; Sodeoka, M., Thienyl-Substituted alpha-Ketoamide: A Less Hydrophobic Reactive Group for Photo-Affinity Labeling. *ACS Chem. Biol.* **2018**, *13*, 876-880.
 52. Ankenbruck, N.; Courtney, T.; Naro, Y.; Deiters, A., Optochemical Control of Biological Processes in Cells and Animals. *Angew. Chem. Int. Ed.* **2018**, *57*, 2768-2798.

53. Sato, S.; Morita, K.; Nakamura, H., Regulation of target protein knockdown and labeling using ligand-directed Ru(bpy)₃ photocatalyst. *Bioconjug Chem* **2015**, *26*, 250-6.
54. Sato, S.; Nakamura, H., Ligand-directed selective protein modification based on local single-electron-transfer catalysis. *Angew. Chem. Int. Ed.* **2013**, *52*, 8681-4.
55. Geri, J. B.; Oakley, J. V.; Reyes-Robles, T.; Wang, T.; McCarver, S. J.; White, C. H.; Rodriguez-Rivera, F. P.; Parker, D. L.; Hett, E. C.; Fadeyi, O. O.; Oslund, R. C.; MacMillan, D. W. C., Microenvironment mapping via Dexter energy transfer on immune cells. *Science* **2020**, *367*, 1091.
56. Ito, K.; Sawanobori, J., 4-Diazomethyl-7-Methoxycoumarin as a New Type of Stable Aryldiazomethane Reagent. *Synth. Commun.* **1982**, *12*, 665-671.
57. Takadate, A.; Tamura, T.; Fujino, H.; Goya, S., Synthesis and Properties of 4-Diazomethyl-7-methoxycoumarin as a New Fluorescent Labeling Reagent for Alcohols and Carboxylic acids. *Chem. Pharm. Bull.* **1982**, *30*, 4120-4125.
58. Ito, K.; Maruyama, J., Studies on Stable Diazoalkanes as Potential Fluorogenic Reagents. I. 7-Substituted 4-Diazomethylcoumarins. *Chem. Pharm. Bull.* **1983**, *31*, 3012-3014.
59. Ito, K.; Maruyama, J., Studies on Stable Diazoalkanes as Potential Fluorogenic Reagents. II. : Ring-Fused 4-Diazomethylcoumarins. *Chem. Pharm. Bull.* **1986**, *34*, 390-395.
60. Ando, H.; Furuta, T.; Tsien, R. Y.; Okamoto, H., Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos. *Nat. Genet.* **2001**, *28*, 317-25.
61. Sivakumar, K.; Xie, F.; Cash, B. M.; Long, S.; Barnhill, H. N.; Wang, Q., A Fluorogenic 1,3-Dipolar Cycloaddition Reaction of 3-Azidocoumarins and Acetylenes. *Org. Lett.* **2004**, *6*, 4603-4606.
62. Bethell, D.; Stevens, G.; Tickle, P., The reaction of diphenylmethylenes with isopropyl alcohol and oxygen: the question of reversibility of singlet-triplet interconversion of carbenes. *J. Chem. Soc. D* **1970**, 792b-794.
63. Dean, F. M.; Park, B. K., Activating groups for the ring expansion of coumarin by diazoethane: benzoyl, pivaloyl, arylsulfonyl, arylsulfinyl, and nitro. *J. Chem. Soc., Perkin Trans. 1* **1976**, 1260-1268.
64. McMahon, R. J.; Abelt, C. J.; Chapman, O. L.; Johnson, J. W.; Kreil, C. L.; LeRoux, J. P.; Mooring, A. M.; West, P. R., 1,2,4,6-Cycloheptatetraene: the key intermediate in arylcarbene interconversions and related C7H6 rearrangements. *J. Am. Chem. Soc.* **1987**, *109*, 2456-2469.
65. Dubinsky, L.; Krom, B. P.; Meijler, M. M., Diazirine based photoaffinity labeling. *Bioorg. Med. Chem.* **2012**, *20*, 554-70.
66. Ito, K.; Maruyama, J., A Facile Intramolecular Cyclization of 4-Diazomethylcoumarins. A Convenient Route to Benzopyrano[3,4-c]pyrazol-4(3H)-ones. *Heterocycles* **1984**, *22*, 1057-1059.
67. Preston, G. W.; Wilson, A. J., Photo-induced covalent cross-linking for the analysis of biomolecular interactions. *Chem. Soc. Rev.* **2013**, *42*, 3289-3301.
68. Takagaki, Y.; Gupta, C. M.; Khorana, H. G., Thiols and the diazo group in photoaffinity labels. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 589-595.
69. McGrath, N. A.; Andersen, K. A.; Davis, A. K.; Lomax, J. E.; Raines, R. T., Diazo compounds for the bioreversible esterification of proteins. *Chem. Sci.* **2015**, *6*, 752-755.
70. Teruya, K.; Tonissen, K. F.; Poulsen, S.-A., Recent developments of small molecule chemical probes for fluorescence-based detection of human carbonic anhydrase II and IX. *Med. Chem. Commun.* **2016**, *7*, 2045-2062.
71. Supuran, C. T., How many carbonic anhydrase inhibition mechanisms exist? *J. Enzyme Inhib. Med. Chem.* **2016**, *31*, 345-60.
72. Li, G.; Liu, Y.; Liu, Y.; Chen, L.; Wu, S.; Liu, Y.; Li, X., Photoaffinity labeling of small-molecule-binding proteins by DNA-templated chemistry. *Angew. Chem. Int. Ed.* **2013**, *52*, 9544-9.
73. Mallory, J. C.; Crudden, G.; Oliva, A.; Saunders, C.; Stromberg, A.; Craven, R. J., A novel group of genes regulates susceptibility to antineoplastic drugs in highly tumorigenic breast cancer cells. *Mol. Pharmacol.* **2005**, *68*, 1747-56.
74. Hulikova, A.; Aveyard, N.; Harris, A. L.; Vaughan-Jones, R. D.; Swietach, P., Intracellular carbonic anhydrase activity sensitizes cancer cell pH signaling to dynamic changes in CO₂ partial pressure. *J. Biol. Chem.* **2014**, *289*, 25418-30.
75. Supuran, C. T., Carbonic Anhydrases - An Overview. *Curr. Pharm. Des.* **2008**, *14*, 603-614.
76. Soethoudt, M.; Stolze, S. C.; Westphal, M. V.; van Stralen, L.; Martella, A.; van Rooden, E. J.; Guba, W.; Varga, Z. V.; Deng, H.; van Kasteren, S. I.; Grether, U.; AP, I. J.; Pacher, P.; Carreira, E. M.; Overkleeft, H. S.; Ioan-Facsinay, A.; Heitman, L. H.; van der Stelt, M., Selective Photoaffinity Probe That Enables Assessment of Cannabinoid CB₂ Receptor Expression and Ligand Engagement in Human Cells. *J. Am. Chem. Soc.* **2018**, *140*, 6067-6075.
77. Yang, J.-S.; Lin, C.-W.; Chuang, C.-Y.; Su, S.-C.; Lin, S.-H.; Yang, S.-F., Carbonic anhydrase IX overexpression regulates the migration and progression in oral squamous cell carcinoma. *Tumor Biol.* **2015**, *36*, 9517-9524.
78. Tafreshi, N. K.; Lloyd, M. C.; Bui, M. M.; Gillies, R. J.; Morse, D. L., Carbonic Anhydrase IX as an Imaging and Therapeutic Target for Tumors and Metastases. In *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*, Frost, S. C.; McKenna, R., Eds. Springer Netherlands: Dordrecht, 2014; pp 221-254.
79. Benej, M.; Pastorekova, S.; Pastorek, J., Carbonic Anhydrase IX: Regulation and Role in Cancer. In *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*, Frost, S. C.; McKenna, R., Eds. Springer Netherlands: Dordrecht, 2014; pp 199-219.
80. An, W. G.; Kanekal, M.; Simon, M. C.; Maltepe, E.; Blagosklonny, M. V.; Neckers, L. M., Stabilization of wild-type p53 by hypoxia-inducible factor 1 α . *Nature* **1998**, *392*, 405-408.
81. Geißler, D.; Antonenko, Y. N.; Schmidt, R.; Keller, S.; Krylova, O. O.; Wiesner, B.; Bendig, J.; Pohl, P.; Hagen, V., (Coumarin - 4 - yl)methyl Esters as Highly Efficient, Ultrafast Phototriggers for Protons and Their Application to Acidifying Membrane Surfaces. *Angew. Chem. Int. Ed.* **2005**, *44*, 1195-1198.
82. Bort, G.; Gallavardin, T.; Ogden, D.; Dalko, P. I., From one-photon to two-photon probes: "caged" compounds, actuators, and photoswitches. *Angew. Chem. Int. Ed.* **2013**, *52*, 4526-37.
83. Dyer, J.; Jockusch, S.; Balsanek, V.; Sames, D.; Turro, N. J., Two-Photon Induced Uncaging of a Reactive Intermediate. Multiphoton In Situ Detection of a Potentially Valuable Label for Biological Applications. *J. Org. Chem.* **2005**, *70*, 2143-2147.
84. Urdabayev, N. K.; Popik, V. V., Wolff Rearrangement of 2-Diazo-1(2H)-Naphthalenone Induced by Nonresonant Two-Photon Absorption of NIR Radiation. *J. Am. Chem. Soc.* **2004**, *126*, 4058-4059.
85. Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W., In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. *Nat. Methods* **2005**, *2*, 255-257.

Supporting information for

**A Visible and Near-Infrared Light Photoactivatable Diazo-Coumarin
Probe for Fluorogenic Protein Labeling in Living Cells**

Sheng-Yao Dai, Dan Yang*

Morningside Laboratory for Chemical Biology, Department of Chemistry, The University of Hong Kong, Pokfulam Road,
Hong Kong, P. R. China.

Table of content

Supplemental Figures and Tables	2
General Information	7
Experimental Procedures	9
Chemical synthesis	16
Appendix	23
References	67

Supplemental Figures and Tables

Figure S1. (a)–(b) Change in absorption and emission spectra of **1** (30 μM) in methanol upon irradiation with a blue LED lamp. $\lambda_{\text{ex}} = 400 \text{ nm}$. (c) HPLC analysis of **1** before and after photolysis. (d) Proposed mechanism of formation of each product and isolated yields from photolysis of **1**.

Figure S2. (a) Proposed mechanism of formation of each product from photolysis of **2**. (b)–(c) Absorption and emission spectra of **2a–d** (30 μM) in methanol. $\lambda_{\text{ex}} = 400 \text{ nm}$.

Table S1. Summary of photophysical properties of **2** and **2a–d**.

Compound	Solvent	λ_{max} (nm) ^a	ϵ (10 ⁴ M ⁻¹ cm ⁻¹) ^b	λ_{em} (nm) ^c	Φ_{fl} ^d
2	CH ₃ OH	408	2.26	490	0.003
2a	CH ₃ OH	392	1.17	489	0.23
2b	CH ₃ OH	392	1.34	481	0.20
2c	CH ₃ OH	351	0.92	489	0.02
2d	CH ₃ OH	393	3.12	485	0.22
2	H ₂ O	414	1.93	500	0.0002
2d	H ₂ O	402	2.89	500	0.015

All the measurements were obtained at 25 °C with 10 μ M of compounds. ^a Wavelength of the longest absorption maxima. ^b Molar extinction coefficient, calculated from the equation $A = \epsilon lc$, where A is the absorbance, c is the concentration of the compound, and l is the path length (1 cm of the cuvette). ^c Wavelength of the emission maxima. ^d Fluorescence quantum yield, measured by using 7-diethylamino-4-methylcoumarin as reference ($\Phi_{\text{fl}} = 0.73$ in EtOH).

Figure S6. (a) Live-cell confocal imaging of MCF7 cells labeled with **3** (5 μM) with or without CBS (200 μM). Scale bar = 5 μm. (b) Live-cell confocal imaging of HCT116 cells treated with **3** (5 μM) and vigorously washed with HBSS for 3 times prior to blue LED irradiation. The imaging setting was identical to Figure 3c. Scale bar = 10 μm. (c) Confocal imaging of fixed MCF7 cells labeled with **3** (5 μM) and immunofluorescence (IF) stained with CA-II antibody. Scale bar = 10 μm.

General Information

All chemicals were purchased from commercial vendors and used without further purification unless indicated otherwise. All reactions requiring anhydrous conditions were carried out under argon atmosphere using oven-dried glassware. AR-grade solvents were used for all reactions. Dichloromethane, DMF, and trimethylamine were distilled from calcium hydride. THF was distilled from sodium. Air and moisture sensitive compounds were introduced via syringes through rubber septa under argon atmosphere. Reaction progress was monitored by TLC on pre-coated silica plates (Merck 60 F254 nm, 0.25 μ m) and spots were visualized by UV and/or staining in phosphomolybdic acid (PMA) or KMnO₄ solution followed by heating. Flash column chromatography was carried out using silica gel (Merck 60 F254 nm, 0.040-0.063 μ m).

NMR spectra were acquired on Bruker Avance DPX 300 Fourier Transform Spectrometer operating at 300 MHz for ¹H and at 75 MHz for ¹³C, Bruker Avance DPX 400 Fourier Transform Spectrometer operating at 400 MHz for ¹H, 100 MHz for ¹³C, 376 MHz for ¹⁹F and Bruker Avance III HD 500 Spectrometer operating at 500 MHz for ¹H, 125 MHz for ¹³C, 470 MHz for ¹⁹F. Chemical shifts were reported in parts per million (ppm) referenced with respect to appropriate internal standards or residual solvent peaks (CDCl₃ = 7.26 ppm, DMSO-*d*₆ = 2.50 ppm, CD₃OD = 3.31 ppm for ¹H; CDCl₃ = 77.16 ppm, DMSO-*d*₆ = 39.52 ppm, CD₃OD = 49.00 ppm for ¹³C). The following abbreviations were used in reporting spectra, br s (broad singlet), s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets) and m (multiplet). High-resolution mass spectra were obtained on Bruker maXis II High-Resolution QTOF using ESI mode. All LCMS analysis was carried out on Agilent 1100 HPLC system equipped with an auto-sampler using reverse-phase Phenomenex Luna 5 μ m C18(2) 100 Å 50 × 3.0 mm columns. Water and acetonitrile containing 0.1% TFA were used as eluents and the flow rate was 0.4 mL/min. Ultraviolet absorption spectra were recorded on a Cary 50 UV-visible spectrophotometer. Fluorescence measurements were carried out on a Hitachi F-7000 fluorescence spectrophotometer. The slit width was set at 2.5 nm. The emission spectrum was scanned from 300 to 700 nm at 1200 nm/min and the photomultiplier voltage was set at 700 V.

MCF7 and HeLa cells were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco) containing 10% fetal bovine serum (FBS; Gibco), 100 U/mL penicillin and 100 μ g/mL streptomycin. HCT116 cells were cultured in McCoy's 5A media (Gibco) supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 μ g/mL streptomycin. Cells were maintained in a humidified 37 °C incubator with 5% CO₂. To harvest the cell lysate proteins, cells were washed twice with phosphate-buffered saline (PBS), treated with 1× trypsin, and collected by centrifugation. Cell pellets were then washed and lysed with lysis buffer (0.1%

SDS, 1% Triton X-100, 150 mM NaCl, 0.1 U Benzonase, Roche EDTA-free protease inhibitor cocktail, 20 mM Tris-HCl, pH 8.0). Protein concentration was determined by BCA protein assay. Cell lysate samples were resolved by hand-cast SDS-PAGE. In-gel fluorescence was visualized by Typhoon FLA 9000 Gel Imaging Scanner (GE Healthcare) using 473 nm blue LD laser and detected by emission filter of ≥ 510 nm. The SDS-PAGE gels were stained with Coomassie brilliant blue (CBB). For Western blot, the resolved proteins were transferred to poly(vinylidene difluoride) (PVDF) membranes. Membranes were then blocked with 3% BSA in 0.1% Tween in Tris-buffered saline (TBST) for 1 h at room temperature. Membranes were incubated with the corresponding primary antibody overnight at 4°C, washed with TBST (3 $\times$ 5 min) and then incubated with an appropriate secondary antibody. Finally, blots were washed again with TBST (3 $\times$ 5 min) and developed with SuperSignal™ West Pico Chemiluminescent Substrate (Thermo Scientific). The membranes were visualized by chemiluminescent mode of ThermoFischer myECL™ Imager.

For confocal imaging, cells were seeded in glass-bottom dishes at 5×10^4 cells/mL density 24 h prior to the experiment. All imaging data of treated cells were collected on a Carl Zeiss LSM 780 confocal microscope system equipped with 40 $\times$ /1.3 Oil and a photomultiplier tube (PMT) detector. For flow cytometry analysis, cells were seeded into a 150-mm culture dish at 5×10^5 cells/mL density 24 h prior to the experiment. Treated cells were then harvested by trypsinization and collected by centrifugation. After discarding the supernatant, cell pellets were re-suspended gently into single cells with 1 mL HBSS. The fluorescence intensity was measured by using BD LSR Fortessa Analyzer. The data were analyzed by FlowJo software.

Experimental Procedures

Small scale photolysis

Compound to be tested was dissolved in 3 mL of methanol in a quartz cuvette and was exposed to various light sources with constant stirring at room temperature. For one-photon excitation, a 7W blue LED lamp covering 430–490 nm (maxima at around 460 nm) was used. For two-photon excitation, a mode-locked Ti:sapphire laser was used to project femtosecond pulses of an 800 nm focused light beam. At the various time intervals, the absorption and emission spectra were measured. All fluorescence measurements were carried out on a Hitachi F-7000 fluorescence spectrophotometer. The slit width was set at 2.5 nm. The emission spectrum was scanned from 400 to 700 nm at 1200 nm/min and the photomultiplier voltage was set at 700 V. The photolysis mixture was concentrated *in vacuo*, re-dissolved in 0.6 mL MeCN and subjected to LC/MS analysis.

Large scale photolysis of **1**

A solution of **1** (43 mg, 0.85 mM) in 250 mL MeOH was irradiated with blue LED lamp for 48 h. The reaction was monitored by HPLC to confirm complete reaction. The reaction mixture was then concentrated *in vacuo* and purified by preparative TLC to afford **1a** (1.0 mg, 2.3%), **1b** (1.4 mg, 3.3%), **1c** (15 mg, 35%), **1d** (8 mg, 19%) and **1e** (17 mg, 38%). **1a**: ¹H NMR (400 MHz, CDCl₃) δ 7.43 (d, *J* = 9.0 Hz, 1H), 6.58 (dd, *J* = 8.8, 2.5 Hz, 1H), 6.51 (d, *J* = 2.5 Hz, 1H), 5.99 (s, 1H), 3.97 (t, *J* = 6.5 Hz, 2H), 3.41 (q, *J* = 7.0 Hz, 4H), 2.95 (t, *J* = 6.6 Hz, 2H), 1.21 (t, *J* = 7.0 Hz, 6H). **1b**: ¹H NMR (300 MHz, CDCl₃) δ 7.32 (d, *J* = 9.0 Hz, 1H), 6.60–6.54 (m, 1H), 6.54–6.50 (m, 1H), 6.25 (s, 1H), 4.83 (s, 1H), 3.41 (q, *J* = 7.1 Hz, 4H), 1.20 (t, *J* = 7.1 Hz, 6H). **1c** is identical to **S2**. **1d**: ¹H NMR (300 MHz, CDCl₃) δ 7.36 (d, *J* = 9.0 Hz, 1H), 6.56 (dd, *J* = 9.0, 2.6 Hz, 1H), 6.51 (d, *J* = 2.6 Hz, 1H), 6.16 (s, 1H), 4.54 (s, 2H), 3.46 (s, 3H), 3.40 (q, *J* = 7.1 Hz, 4H), 1.20 (t, *J* = 7.0 Hz, 6H). **1e**: ¹H NMR (300 MHz, CDCl₃) δ 7.12 (d, *J* = 8.6 Hz, 1H), 6.91 (dd, *J* = 11.2, 5.1 Hz, 1H), 6.54 (d, *J* = 2.3 Hz, 1H), 6.47 (dd, *J* = 8.6, 2.3 Hz, 1H), 5.86 (d, *J* = 11.2 Hz, 1H), 4.85 (d, *J* = 5.1 Hz, 1H), 3.41 (s, 3H), 3.33 (q, *J* = 7.0 Hz, 4H), 1.15 (t, *J* = 7.0 Hz, 6H).

Large scale photolysis of **2**

A solution of **2** (16 mg, 0.05 mM) in 250 mL MeOH was irradiated with blue LED lamp for 48 h. The reaction was monitored by HPLC to confirm complete reaction. The reaction mixture was then concentrated *in vacuo* and purified by preparative TLC to afford **2a** (2.4 mg, 15%), **2b** (2.7 mg, 17%), **2c** (1.7 mg, 10%) and **2d** (7 mg, 43%). **2a** is identical to **S5** and **2b** is identical to **S4**. **2c**: ¹H NMR (300 MHz, CDCl₃) δ 7.51 (d, *J* = 9.0 Hz, 1H), 6.61–6.55 (m, 1H), 6.53 (d, *J* = 2.6 Hz, 1H), 4.77 (q, *J*_{H,F} = 3.0 Hz, 1H), 3.50 (s, 3H), 3.42 (q, *J* = 7.0 Hz, 4H), 1.22

(t, $J = 7.0$ Hz, 6H). **2d** ^1H NMR (300 MHz, CDCl_3) δ 12.01 (s, 1H), 7.75 (d, $J = 9.0$ Hz, 1H), 6.74–6.68 (m, 1H), 6.65 (d, $J = 2.6$ Hz, 1H), 3.42 (q, $J = 7.0$ Hz, 4H), 1.22 (br t, $J = 7.0$ Hz, 6H).

Assay for the stability in the aqueous medium and towards glutathione

A solution of **2** (5 mM) in 0.1 M sodium phosphate buffer (pH values 7.4) was incubated in dark for 24 h with or without the addition of an equivalent of reduced glutathione (GSH). The ^1H NMR before and after the incubation was recorded and compared.

Assay for the stability of **2** in different pH values

A solution of **2** (10 μM) in 0.1 M sodium phosphate buffer that was adjusted to various pH values (3–10) was incubated in dark for 1 h. The fluorescence intensity was recorded after incubation. As a control, the fluorescence intensity of **2** (10 μM) after irradiation in water with a blue LED lamp was measured. Emission was recorded at 510 nm with an excitation at 400 nm.

In-gel fluorescence analysis of BSA labeling

BSA (15 μM) and **1** or **2** (0.5 μL DMSO stock solution) were mixed in a total volume of 50 μL solution in PBS buffer, which was then incubated under darkness for 1 h at room temperature. The samples were then irradiated with a blue LED lamp (430–490 nm) at a distance of 10 cm for 1 h at room temperature. The reaction was treated with 4 $\times$ SDS sample buffer, then separated by 10% SDS-PAGE.

Enzymatic activity assay of CA-II

CA-II activity can be calorimetrically determined by the catalytic hydrolysis of *p*-nitrophenyl acetate into *p*-nitrophenol, which shows strong absorbance at 405 nm. CA-II (1 μM) and *p*-nitrophenyl acetate (2.5 mM) were incubated in the presence of **3** or CBS at various concentrations in a 96-well plate in PBS buffer of pH 7.4 (final volume of 100 μL). After incubation at 24 $^\circ\text{C}$ for 30 min, the absorbance at 405 nm was measured using a multimode microplate detector (DTX 800, Beckman Coulter). The percentage inhibition was defined relative to the maximal activity which was measured without the inhibitor. The IC_{50} value was obtained from a graph of percent inhibition versus inhibitor concentration and calculated using Prism Graphpad software.

In-gel fluorescence analysis of recombinant CA-II labeling

CA-II (1.7 μ M) and **3** (0.1 μ L DMSO stock solution) were mixed in a total volume of 10 μ L solution in PBS buffer, which was incubated under darkness for 1 h at room temperature. The samples were then irradiated with a blue LED lamp (430–490 nm) at a distance of 10 cm for 1 h at room temperature. The reaction was treated with 4 $\times$ SDS sample buffer, then separated by 12% SDS-PAGE. In-gel fluorescence was visualized by Typhoon FLA 9000 Gel Imaging Scanner (GE Healthcare) using 473 nm blue LD laser and detected by emission filter of $\geq$ 510 nm.

LC/MS/MS analysis of recombinant CA-II labeling by tryptic digestion

CA-II (1.7 μ M) and **3** (50 μ M) were mixed in a total volume of 100 μ L solution in PBS buffer, which was incubated under darkness for 1 h at room temperature. The samples were then irradiated with a blue LED lamp (430–490 nm) at a distance of 10 cm for 1 h at room temperature. The buffer was exchanged by size exclusion mini spin column (10 kDa) into sample buffer that contains 50 mM Tris-HCl (pH 8), 5 mM DTT and 6 M urea, incubated at 56°C for 30 min. The sample was cooled to room temperature and alkylated with 12 mM iodoacetamide in the dark for 30 min at room temperature, then quenched with 1 mM DTT. The sample was diluted with 7 times volume of Tris-HCl buffer to bring the urea concentration to less than 1 M. MS grade trypsin (1 μ g, trypsin:protein 1:50 w/w) was added and the digestion reaction was incubated for 16 h at 37°C. A final concentration of 5% acetic acid was added to terminate the reaction. The obtained peptides were desalted with a homemade spintip packed with a C18 disk and dried in a speedvac. The obtained peptides were resuspended in 15 μ L of 0.1% (v/v) FA and 5 μ L of samples were used for MS analysis. The peptides were separated by an Easy-nLC 1000 and analyzed by an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). The full MS scans were performed with the m/z range of 350–1550 and the mass resolution of 120000. Peptides were selected in a quadrupole mass analyzer with a 1.6 Da isolation window and fragmented by HCD. The raw data was loaded into the MaxQuant software and searched against the Uniprot Carbonic anhydrase 2 P00921 (CAH2_BOVIN).

Labeling of recombinant CA-II in HeLa cell lysates

HeLa cell lysate proteins (10 μ g), recombinant CA-II (1 μ g), **3** (10 μ M) and **CBS** (100 μ M) were mixed in a total volume of 20 μ L solution in PBS buffer, which was then incubated under darkness for 1 hour at room temperature. The samples were irradiated at 430–490 nm at a distance of 10 cm for 1 h using a blue LED lamp. The reaction was treated with SDS sample

buffer then separated by 12% SDS-PAGE. In-gel fluorescence was visualized by Typhoon FLA 9000 Gel Imaging Scanner (GE Healthcare) using 473 nm blue LD laser and detected by emission filter of ≥ 510 nm.

Labeling of endogenous CA-II in living cells

MCF7 or HCT116 cells were seeded in 6-well plates at 10^5 cells/mL density 24 h prior to the experiment. Cells were then treated with 1 mL of the DMEM containing **3** (10 μ M), with or without **CBS** (100 μ M). After 6 h of incubation at 37 °C/5% CO₂, the medium was aspirated, and cells were washed gently once with HBSS to remove the excessive probe, followed by the blue LED lamp irradiation for 30 min in HBSS. The cell lysate proteins were collected and were separated by SDS-PAGE. In-gel fluorescence was visualized by Typhoon FLA 9000 Gel Imaging Scanner (GE Healthcare) using 473 nm blue LD laser and detected by emission filter of ≥ 510 nm, the resolved proteins were then transferred PVDF membranes. Western blot was performed with anti-CA-II antibody (Abcam, #ab124687, 1:1000, Rb) and anti-Rb antibody-HRP (Cell Signaling Technology, #7074P2, 1:2000).

Labeling in CA-II knock-down MCF7 cells

MCF7 cells were seeded in 24-well plates at 10^4 cells/mL density 24 h prior to the experiment. Knock-down of CA-II was performed by first premixing CA-II siRNA (SMARTpool siGENOME, Dharmcon, 5 μ M, 2.5 μ L) and transfection reagent (DharmaFECT 1, Dharmcon, 0.25 μ L) in Opti-MEM reduced serum medium at room temperature, followed by addition to 400 μ L DMEM to culture cells for 48 h. Cells were then treated with medium containing **3** (10 μ M), with or without **CBS** (100 μ M). After 6 h of incubation at 37 °C/5% CO₂, the medium was aspirated, and cells were washed gently once with HBSS to remove the excessive probe, followed by the blue LED lamp irradiation for 30 min in HBSS. The cell lysate proteins were collected and were separated by SDS-PAGE. In-gel fluorescence was visualized by Typhoon FLA 9000 Gel Imaging Scanner (GE Healthcare) using 473 nm blue LD laser and detected by emission filter of ≥ 510 nm, the resolved proteins were then transferred PVDF membranes for western blot. Western blot was performed with anti-CA-II antibody (Abcam, #ab124687, 1:1000, Rb), anti-GAPDH antibody (Santa Cruz, #sc-25778, 1:1000, Rb) and anti-Rb antibody-HRP (Cell Signaling Technology, #7074P2, 1:2000).

Live-cell imaging of endogenous CA-II

Cells were seeded in glass-bottom dishes at 5×10^4 cells/mL density 24 h prior to the experiment. The cells were treated with 1 mL of DMEM with **3** (5 μ M), with or without **CBS** (100 μ M). After incubation for 2 h at 37 °C/5% CO₂, the medium was removed and cells were

gently washed once with HBSS, followed by irradiation with blue LED lamp for 30 min in HBSS. NucRed™ Live 647 ReadyProbes™ Reagent (Thermo Scientific) was added to the incubation solution for 30 min. All imaging data were collected on a Carl Zeiss LSM 780 confocal microscope system equipped with 40x/1.3 Oil and a photomultiplier tube (PMT) detector. Coumarin was excited by 405 nm diode laser and detected by bandpass of 420–550 nm. NucRed was excited by 633 nm laser and detected by bandpass of 634–750 nm.

Immunofluorescence staining of CA-II of fixed MCF7 cells

MCF7 cells were seeded in glass-bottom dishes at 5×10^4 cells/mL density 24 h prior to the experiment. The cells were treated with 1 mL of DMEM with **3** (5 μ M) in the presence or absence of **CBS** (100 μ M). After incubation for 2 h at 37 °C/5% CO₂, the medium was removed and cells were gently washed once with HBSS, followed by irradiation with blue LED lamp for 30 min in HBSS. The cells were then fixed and permeabilized for 10 min at –20°C with ice-cold methanol, then washed twice with PBS. Cells were then blocked with 2% BSA in PBS for 1 h, and incubated with the anti-CA-II antibody (Novus, NB600-919, 1:20, Rb) for 1 h at room temperature, washed twice with TBST, and then incubated with Alexa Fluor 488 goat anti-rabbit (Thermofisher, # A27034, 1:1000) for 1 h, followed by washing with TBST before imaging. All imaging data were collected on a Carl Zeiss LSM 780 confocal microscope system equipped with 40x/1.3 Oil and a photomultiplier tube (PMT) detector. Coumarin was excited by 405 nm diode laser and detected by bandpass of 420–550 nm. NucRed was excited by 633 nm laser and detected by bandpass of 634–750 nm. The Pearson's coefficient *r* was determined to be 0.92 using colocalization analysis (Coloc 2) of ImageJ software (Fiji).

Flow cytometry analysis of endogenous CA-II

Cells were seeded into a 150-mm culture dish at 5×10^5 cells/mL density 24 h prior to the experiment. The cells were treated with 1 mL of DMEM with **3** (5 μ M) in the presence or absence of **CBS** (100 μ M). After incubation for 2 h at 37 °C/5% CO₂, the medium was removed and cells were gently washed with HBSS, followed by irradiation with blue LED lamp (430–490 nm) for 30 min at room temperature. Cells were then harvested by trypsinization, and collected by centrifugation. After discarding the supernatant, cell pellets were re-suspended gently into single cells with 1 mL HBSS. The fluorescence intensity was measured in Pacific Blue channel (405 nm) by using BD LSR Fortessa Analyzer. The data were analyzed by FlowJo software.

Live-cell imaging of CA-IX under hypoxia-mimetic condition

HeLa cells were seeded in glass-bottom dishes at 1×10^4 cells/mL density 24 h prior to the experiment. To induce CA-IX expression, cells were treatment with 1 mL of DMEM with DFO (100 μ M) and incubate for 24 h at 37 °C/5% CO₂. The cells were then treated with 1 mL of DMEM with **3** (5 μ M), with or without **CBS** (100 μ M). After incubation for 1 h at 37 °C/5% CO₂, the medium was removed, followed by irradiation with blue LED lamp for 30 min in HBSS. All imaging data were collected on a Carl Zeiss LSM 780 confocal microscope system equipped with 40x/1.3 Oil and a photomultiplier tube (PMT) detector. Coumarin was excited by 405 nm diode laser and detected by bandpass of 420–550 nm. To confirm the CA-IX expression, the cell lysate proteins were collected and were separated by SDS-PAGE. The resolved proteins were then transferred PVDF membranes for western blot. Western blot was performed with anti-CA-IX antibody (Abcam, #ab15086, 1:1000, Rb), anti-HIF-1 α antibody (Santa Cruz, #sc-10790, 1:1000, Rb), anti-GAPDH antibody (Santa Cruz, #sc-25778, 1:1000, Rb) and anti-Rb antibody-HRP (Cell Signaling Technology, #7074P2, 1:2000).

In situ two-photon labeling and imaging of CA-II in MCF7 cells

Cells were seeded in glass-bottom dishes at 5×10^4 cells/mL density 24 h prior to the experiment. The cells were treated with 1 mL of DMEM with **3** (5 μ M) in the presence or absence of **CBS** (100 μ M). After incubation for 2 h at 37 °C/5% CO₂, the medium was removed and cells were gently washed once with HBSS. NucRed™ Live 647 ReadyProbes™ Reagent (Thermo Scientific) was added to the incubation solution for 30 min. Dishes were then placed on Carl Zeiss LSM 780 confocal microscope system equipped with 40x/1.3 Oil and a photomultiplier tube (PMT) detector. Cells were first tracked with NucRed excited by 633 nm laser. Region of interest was selected and two-photon activation was performed with bleaching mode using 800 nm Coherent Chameleon Vision II laser (1% laser power, 30 cycles). Subsequently, cells were fixed and washed with MeOH for 10 min on stage, then washed with PBS once. One-photon imaging was then performed with a 405 nm diode laser and detected by a bandpass of 420–550 nm.

Live-cell imaging of HeLa cells expressing mCherry-eDHFR

HeLa cells were seeded in glass-bottom dishes at 10^4 cells/mL density 24 h prior to the experiment. mCherry-eDHFR plasmid DNA (Addgene #61500) (2 μ g) was dissolved in 100 μ L Opti-MEM (Gibco) and incubated with 3 μ L X-treme Gene HP DNA transfection agent (Roche) for 15 min. The cells were then treated with the transfection complex in 2 mL DMEM. After 48 h, the cells were treated with 1 mL of DMEM with **4** (5 μ M) in the presence or absence

of **TMP** (100 μ M). After incubation for 2 h at 37 °C/5% CO₂, the medium was removed and cells were gently washed once with HBSS, followed by irradiation with blue LED lamp for 30 min in HBSS. All imaging data were collected on a Carl Zeiss LSM 780 confocal microscope system equipped with 40x/1.3 Oil and a photomultiplier tube (PMT) detector. Coumarin was excited by 405 nm diode laser and detected by bandpass of 420–550 nm. mCherry was excited by 561 nm laser and detected by bandpass of 567–781 nm.

Chemical synthesis

Preparation of **1**

Scheme S1

1 was synthesized according to reported procedures in literature^[1]. ¹H NMR (400 MHz, CDCl₃) δ 7.07 (d, J = 9.0 Hz, 1H), 6.47 (dd, J = 9.0, 2.2 Hz, 1H), 6.38 (d, J = 2.2 Hz, 1H), 5.45 (s, 1H), 5.31 (s, 1H), 3.33 (q, J = 7.0 Hz, 4H), 1.13 (t, J = 7.0 Hz, 6H); ¹³C NMR (100 MHz, CDCl₃) δ 161.8, 155.7, 150.8, 146.6, 123.7, 108.3, 104.9, 97.6, 94.8, 45.4, 44.6, 12.4; HRMS (ESI): m/z calcd for C₁₄H₁₆N₃O₂ [M+H]⁺ 258.1243, found 258.1239.

Preparation of **2**

Scheme S2

To a solution of **S2** (56 mg, 0.23 mmol) in DMF (3 mL) were added TMSCF₃ (41 μ L, 0.28 mmol) and K₂CO₃ (3 mg, 0.02 mmol). The resulting mixture was stirred at room temperature for 1 h. The mixture was then concentrated *in vacuo*. The residue was dissolved in THF (10

mL), and 3 *N* HCl (5 mL) was added to the solution. The resulting mixture was stirred at room temperature for 1 h. After that, the mixture was diluted with Et₂O. The organic layer was separated, and the aqueous layer was extracted three times with Et₂O. The combined organic layer was washed with the saturated aqueous solution of NaHCO₃ and brine subsequently, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (20% EA/Hexane) to afford **S4** (62 mg, 81%) as a yellow oil. ¹H NMR (400 MHz, CDCl₃) δ 7.70–7.43 (m, 1H), 6.57 (dd, *J* = 9.1, 2.4 Hz, 1H), 6.42 (d, *J* = 2.4 Hz, 1H), 6.36 (s, 1H), 5.39 (q, *J*_{H,F} = 6.1 Hz, 1H), 5.17–4.81 (m, 1H), 3.38 (q, *J* = 7.0 Hz, 4H), 1.18 (t, *J* = 7.0 Hz, 6H); ¹³C NMR (100 MHz, CDCl₃) δ 163.0, 156.4, 150.9, 149.6, 126.0, 109.4 (q, ¹*J*_{C,F} = 288.1 Hz), 109.2, 108.5, 97.6, 78.2, 68.4 (q, ²*J*_{C,F} = 33.5 Hz), 44.8, 12.5; ¹⁹F NMR (376.5 MHz, CDCl₃) δ –76.1 (d, *J* = 6.0 Hz); HRMS (ESI): *m/z* calcd for C₁₅H₁₇F₃NO₃ [M+H]⁺ 316.1161, found 316.1139.

To a solution of **S4** (62 mg, 0.2 mmol) in CH₂Cl₂ (2 mL) was added freshly prepared Dess-Martin periodinane (170 mg, 0.4 mmol). The resulting suspension was stirred at room temperature for 2 h. The reaction was quenched by aqueous solution of sodium thiosulfate. The organic layer was separated, and the aqueous layer was extracted three times with CH₂Cl₂. The combined organic layer was washed with saturated aqueous solution of NaHCO₃ and brine subsequently, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (40% EA/Hexane) to give **S5** (50 mg, 80%) as a red oil. ¹H NMR (400 MHz, CD₃OD) δ (hydrate form) 8.29–8.09 (m, 1H), 6.69–6.64 (m, 1H), 6.51–6.48 (m, 1H), 6.47–6.40 (m, 1H), 3.45–3.39 (m, 4H), 1.24–1.13 (m, 6H); ¹³C NMR (100 MHz, CD₃OD) δ (hydrate form) 164.2, 158.3, 152.2, 150.7, 130.11, 124.0 (q, ¹*J*_{C,F} = 288.1 Hz), 112.3, 110.4, 107.1, 98.29 (q, ²*J*_{C,F} = 33.5 Hz), 98.25, 45.7, 12.9; ¹⁹F NMR (376.5 MHz, CD₃OD) δ –83.6 (s); HRMS (ESI): *m/z* calcd for C₁₅H₁₅F₃NO₃ [M+H]⁺ 314.1004, found 314.0982.

To a solution of **S5** (75 mg, 0.24 mmol) in freshly distilled CH₂Cl₂ (3 mL) were added NEt₃ (0.10 mL, 0.73 mmol), CbzHNNH₂ (73 mg, 0.29 mmol) and TiCl₄ (13 μL, 0.12 mmol) at 0 °C under argon atmosphere. The reaction was allowed to warm to room temperature and react for 12 h. The reaction was quenched by the saturated aqueous NH₄Cl. The organic layer was separated, and the aqueous layer then extracted three times with EA. The combined organic layer was washed with brine, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (20% EA/Hexane) to give **S6** (114 mg, 87%) as a yellow solid. ¹H NMR (500 MHz, DMSO-*d*₆) mixture of two isomers in 2.3:1 ratio δ 11.70 (br s, 0.3H), 11.16 (br s, 0.7H), 7.63 (d, *J* = 7.2 Hz, 0.3H), 7.47–7.28 (m, 5H), 6.91 (d, *J* = 8.9 Hz, 0.7H), 6.69 (d, *J* = 9.0 Hz, 1H), 6.60 (d, *J* = 10.8 Hz, 1H), 6.16 (s, 0.7H), 6.11 (s, 0.3H), 5.23–5.15 (m, 2H), 3.42 (q, *J* = 7.0 Hz, 4H), 1.11 (t, *J* = 7.0 Hz, 6H); ¹³C NMR (125 MHz, DMSO-*d*₆) δ 160.1 (160.0), 156.6 (156.2), 153.0 (153.2), 150.94

(150.86), 141.3, 135.6 (135.7), 132.4 (q, $^2J_{C,F} = 33.5$ Hz), 128.7, 128.4, 128.0, 126.6, 120.2 (117.4) (q, $^1J_{C,F} = 288.1$ Hz), 110.8, 109.2 (109.3), 104.8, 97.3 (97.1), 67.1 (67.2), 44.1, 12.3; ^{19}F NMR (470.5 MHz, DMSO- d_6) δ -66.3 (s); HRMS (ESI): m/z calcd for $\text{C}_{23}\text{H}_{23}\text{F}_3\text{N}_3\text{O}_4$ $[\text{M}+\text{H}]^+$ 462.1641, found 462.1609.

To a solution of **S6** (38 mg, 0.05 mmol) in methanol (10 mL) was added Pd/C (4 mg). The reaction was stirred at room temperature for 1 h under H_2 atmosphere. After that, the mixture was filtered through celite and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (20% EA/Hexane) to give **S7** (13 mg, 79%) as a yellow oil. ^1H NMR (400 MHz, CDCl_3) δ 7.06 (d, $J = 9.0$ Hz, 1H), 6.58 (dd, $J = 9.0, 2.5$ Hz, 1H), 6.53 (d, $J = 2.5$ Hz, 1H), 6.03–6.05 (m, 3H), 3.42 (q, $J = 7.0$ Hz, 4H), 1.21 (t, $J = 7.14$, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 161.2, 156.9, 151.7, 150.9, 142.5, 129.3 (q, $^2J_{C,F} = 33.5$ Hz), 126.9, 120.9 (q, $^1J_{C,F} = 272.8$ Hz), 110.2, 109.5, 97.9, 45.0, 12.5; ^{19}F NMR (376.5 MHz, CDCl_3) δ -66.9 (s); HRMS (ESI): m/z calcd for $\text{C}_{15}\text{H}_{17}\text{F}_3\text{N}_3\text{O}_2$ $[\text{M}+\text{H}]^+$ 328.1273, found 328.1250.

To a stirring solution of **S7** (48 mg, 0.13 mmol) in DMF (1 mL) was added TsNIK^[2] (42 mg, 0.14 mmol). The reaction was stirred at room temperature for 1 h. Then the mixture was concentrated *in vacuo*, purified by silica gel column chromatography (10% EA/Hexane) to afford **2** (41 mg, 96%) as a yellow solid. ^1H NMR (500 MHz, CDCl_3) δ 7.30 (d, $J = 9.1$ Hz, 1H), 6.58 (dd, $J = 9.1, 2.6$ Hz, 1H), 6.53 (d, $J = 2.6$ Hz, 1H), 6.05 (s, 1H), 3.42 (q, $J = 7.1$ Hz, 4H), 1.22 (t, $J = 7.1$ Hz, 6H); ^{13}C NMR (125 MHz, CDCl_3) δ 161.5, 157.4, 151.94, 139.1, 126.9, 125.8 (q, $^1J_{C,F} = 270.9$ Hz), 109.7, 105.8, 102.6 (q, $^2J_{C,F} = 38.1$ Hz), 99.3, 45.8, 30.7, 13.4; ^{19}F NMR (470.5 MHz, CDCl_3) δ -56.2 (s); HRMS (ESI): m/z calcd for $\text{C}_{15}\text{H}_{15}\text{F}_3\text{N}_3\text{O}_2$ $[\text{M}+\text{H}]^+$ 326.1116, found 326.1094.

Preparation of **3** and **4**

Scheme S3

S8 was prepared according to procedures reported in literature^[3]. ¹H NMR (300 MHz, CDCl₃) δ 7.32 (d, *J* = 8.6 Hz, 1H), 6.51 (dd, *J* = 8.6, 2.4 Hz, 1H), 6.43 (d, *J* = 2.4 Hz, 1H), 5.94 (d, *J* = 0.9 Hz, 1H), 4.48 (s, 1H), 3.31–3.08 (m, 2H), 2.32 (d, *J* = 1.1 Hz, 3H), 1.28 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (75 MHz, CDCl₃) δ 162.2, 156.0, 153.3, 151.8, 125.5, 110.3, 110.2, 108.9, 97.8, 38.0, 18.6, 14.4.

To a solution of **S8** (7.0 g, 34.4 mmol) in MeCN (150 mL) were added *tert*-butyl bromoacetate (25 g, 170 mmol), K₂CO₃ (6.18 g, 44.7 mmol) and KI (5.7 g, 34.4 mmol). The reaction was stirred under reflux for 24 h and was filtered after cooling down. The solvent was

evaporated, and the residue was then dissolved with EA. The organic layer was washed with water and brine subsequently, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (20% EA/Hexane) to give **S9** (8.89 g, 82%) as a yellow crystal. ¹H NMR (500 MHz, CDCl₃) δ 7.37 (d, *J* = 8.9 Hz, 1H), 6.53 (dd, *J* = 8.9, 2.6 Hz, 1H), 6.43 (d, *J* = 2.6 Hz, 1H), 5.94 (d, *J* = 1.1 Hz, 1H), 3.95 (s, 2H), 3.48 (q, *J* = 7.1 Hz, 2H), 2.31 (d, *J* = 0.9 Hz, 3H), 1.44 (s, 9H), 1.27–1.17 (m, 3H); ¹³C NMR (125 MHz, CDCl₃) δ 169.3, 162.0, 155.7, 152.9, 150.9, 125.6, 110.2, 109.5, 108.7, 98.3, 82.2, 53.1, 46.7, 28.1, 18.5, 12.4; HRMS (ESI): *m/z* calcd for C₁₈H₂₄NO₄ [M+H]⁺ 318.1705, found 318.1679.

To a suspension of **S9** (8.89 g, 28.0 mmol) in xylene (280 mL) was added SeO₂ (4.67 g, 42 mmol). The mixture was heated under reflux for 24 h. While hot, the solution was filtered through celite and washed thoroughly with EA. The filtrate was then concentrated *in vacuo*. The residue was purified by silica gel column chromatography (40% EA/Hexane) to give **S10** (4.78 g, 51%) as a red solid. ¹H NMR (500 MHz, CDCl₃) δ 9.97 (s, 1H), 8.26 (dd, *J* = 9.1, 1.6 Hz, 1H), 6.54 (d, *J* = 9.2 Hz, 1H), 6.47–6.42 (m, 2H), 3.95 (s, 2H), 3.48 (q, *J* = 7.2 Hz, 2H), 1.43 (s, 9H), 1.22 (t, *J* = 7.2 Hz, 3H); ¹³C NMR (125 MHz, CDCl₃) δ 192.4, 168.9, 161.6, 157.0, 151.4, 143.8, 127.1, 118.5, 109.7, 104.7, 98.3, 82.4, 53.0, 46.8, 28.1, 12.3; HRMS (ESI): *m/z* calcd for C₁₈H₂₂NO₅ [M+H]⁺ 331.1414, found 331.1415.

To a solution of **S10** (4.78 mg, 14.4 mmol) in DMF (100 mL) were added TMSCF₃ (2.77 mL, 18.8 mmol) and K₂CO₃ (20 mg, 0.14 mmol). The resulting mixture was stirred at room temperature for 2 h and then concentrated *in vacuo*. The residue was dissolved in THF (60 mL) and 1 *N* HCl (30 mL) was added to the solution. The resulting mixture was stirred at room temperature for 3 h. The mixture was then diluted with Et₂O. The organic layer was separated, and the aqueous layer was extracted three times with Et₂O. The combined organic layer was washed with saturated aqueous solution of NaHCO₃ and brine subsequently, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (30% EA/Hexane) to afford **S11** (5.38 g, 93%) as a yellow oil. ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.80 (d, *J* = 9.0 Hz, 1H), 7.23 (d, *J* = 5.3 Hz, 1H), 6.64 (dd, *J* = 9.0, 2.3 Hz, 1H), 6.51 (d, *J* = 2.3 Hz, 1H), 6.27 (s, 1H), 5.74–5.61 (m, 1H), 4.18 (s, 2H), 3.48 (q, *J* = 7.4 Hz, 2H), 1.41 (s, 9H), 1.12 (t, *J* = 7.4 Hz, 3H); ¹³C NMR (100 MHz, DMSO-*d*₆) δ 169.3, 160.4, 155.7, 151.1, 150.2, 126.8, 124.4 (q, ¹*J*_{C,F} = 283.4 Hz), 109.0, 106.5, 97.5, 81.1, 66.8 (q, ²*J*_{C,F} = 31.0 Hz), 52.3, 45.9, 27.7, 12.2; ¹⁹F NMR (376.5 MHz, DMSO-*d*₆) δ –75.1 (d, *J* = 6.6 Hz); HRMS (ESI): *m/z* calcd for C₁₉H₂₃F₃NO₅ [M+H]⁺ 402.1528, found 402.1504.

To a solution of **S11** (5.38 mg, 13.4 mmol) in CH₂Cl₂ (100 mL) was added freshly prepared Dess-Martin periodinane (6.2 g, 14.6 mmol). The resulting suspension was stirred at room temperature for 1 h. The reaction was quenched by aqueous solution of sodium thiosulfate. The

organic layer was separated, and the aqueous layer was extracted three times with CH₂Cl₂. The combined organic layer was washed with saturated aqueous solution of NaHCO₃ and brine subsequently, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (40% EA/Hexane) to give **S12** (4.46 g, 83%) as a red oil. ¹H NMR (500 MHz, DMSO-*d*₆) δ (hydrate form) 8.22 (s, 2H), 8.12 (d, *J* = 9.3 Hz, 1H), 6.70–6.56 (m, 1H), 6.47 (d, *J* = 2.4 Hz, 1H), 6.34 (s, 1H), 4.16 (s, 2H), 3.56–3.42 (m, 2H), 1.41 (s, 9H), 1.12 (t, *J* = 7.0 Hz, 3H); ¹³C NMR (125 MHz, DMSO-*d*₆) δ (hydrate form) 169.2, 160.5, 156.0, 151.6, 150.6, 129.4, 123.3 (q, ¹*J*_{C,F} = 290.3 Hz), 110.3, 108.7, 106.1, 97.3, 93.1 (q, ²*J*_{C,F} = 33.2 Hz), 81.0, 52.2, 45.7, 27.7, 12.2; ¹⁹F NMR (470.5 MHz, DMSO-*d*₆) δ –81.6 (s); HRMS (ESI): *m/z* calcd for C₁₉H₂₀F₃NNaO₅ [M+Na]⁺ 422.1191, found 422.2913.

To a stirring solution of **S12** (4.46 g, 11.1 mmol) in freshly distilled CH₂Cl₂ (110 mL) were added NEt₃ (7.74 mL, 55.5 mmol), CbzHNNH₂ (2.41 g, 14.5 mmol) and TiCl₄ (1.46 mL, 13.3 mmol) at 0 °C under argon atmosphere. The reaction was allowed to warm to room temperature and react for 12 h. The reaction was quenched by saturated aqueous NH₄Cl. The organic layer was separated, and the aqueous layer then extracted three times with EA. The combined organic layer was washed with brine, dried over anhydrous magnesium sulfate and concentrated *in vacuo*. The residue was purified by silica gel column chromatography (20% EA/Hexane) to give **S13** (3.96 g, 65%) as an orange solid. ¹H NMR (500 MHz, CDCl₃) δ 8.04 (br s, 1H), 7.38–7.29 (m, 5H), 7.00 (d, *J* = 9.0 Hz, 1H), 6.59–6.45 (m, 2H), 6.05 (s, 1H), 5.33–5.12 (m, 2H), 3.97 (s, 2H), 3.51 (q, *J* = 7.2 Hz, 2H), 1.46 (s, 9H), 1.25 (t, *J* = 7.2 Hz, 3H); ¹³C NMR (125 MHz, CDCl₃) δ 168.7, 160.0, 156.8, 152.5, 140.8, 134.8, 129.1, 129.0, 128.92, 128.86, 128.85 (q, ¹*J*_{C,F} = 275.6 Hz), 126.4, 119.9 (q, ²*J*_{C,F} = 32.4 Hz), 111.3, 110.1, 104.7, 99.0, 82.9, 68.9, 53.1, 47.1, 28.2, 12.4; ¹⁹F NMR (470.5 MHz, CDCl₃) δ –67.6 (s); HRMS (ESI): *m/z* calcd for C₂₇H₂₉F₃N₃O₆ [M+H]⁺ 548.2008, found 548.1963.

To a flask containing **S13** (3.96 g, 7.23 mmol) under argon atmosphere were added Pd/C (400 mg) and MeOH (50 mL) sequentially. The reaction atmosphere was then purged with hydrogen gas via a balloon, allowed to stir for 2 h at room temperature. The reaction mixture was then filtered through celite and concentrated *in vacuo* to afford **S14** as a yellow oil. Subsequently, to the solution of **S14** in MeCN (30 mL) and water (0.6 mL) was added iodine (549 mg, 2.16 mmol). The reaction was warmed to gentle reflux for 2 h and concentrated *in vacuo* to afford **S15**, which was used for next step without further purification. To a stirring solution of **S15** in DMF (20 mL) was added TsNIK (1.90 g, 5.54 mmol). The reaction was stirred at room temperature for 1 h. Then the mixture was concentrated *in vacuo*. The residue was dissolved in water, washed with EA three times. The aqueous layer was acidified to pH 3 with 1 *N* HCl and extracted with EA three times. The combined organic layer was concentrated *in vacuo* to afford **S16** (1.68 g, 65% for 3 steps) as a yellow solid. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 9.11 (br s, 1H), 7.28 (d, *J* = 9.0 Hz, 1H), 6.56 (dd, *J* = 9.0, 2.7 Hz, 1H), 6.44 (d, *J* = 2.2 Hz, 1H), 6.02

(s, 1H), 4.08 (s, 2H), 3.47 (q, $J = 7.0$ Hz, 2H), 1.20 (t, $J = 7.1$ Hz, 3H); ^{13}C NMR (100 MHz, CDCl_3) δ 174.7, 160.9, 156.0, 151.4, 138.6, 124.8 (q, $^1J_{\text{C,F}} = 271.3$ Hz), 124.7, 109.3, 105.7, 105.1, 99.2, 58.4 (q, $^2J_{\text{C,F}} = 44.0$ Hz), 52.0, 46.7, 12.2; ^{19}F NMR (376.5 MHz, CDCl_3) δ -72.3 (s); HRMS (ESI): m/z calcd for $\text{C}_{15}\text{H}_{13}\text{F}_3\text{N}_3\text{O}_4$ $[\text{M}-\text{N}_2+\text{H}]^+$ 328.0797, found 328.0891.

S17 was synthesized from **CBS** according to procedures reported in literature^[4]. To a solution of **S16** (40 mg, 0.1 mmol) in DMF (10 mL) were added EDCI (38 mg, 0.2 mmol) and NEt_3 (40 μL , 0.2 mmol) under argon atmosphere. The mixture was stirred for 15 min, and **S18** (60 mg, 0.2 mmol) was added. The resulting mixture was stirred for further 12 h at room temperature. The reaction was then concentrated *in vacuo* and was purified by silica gel column chromatography (2% MeOH/ CH_2Cl_2) to afford **3** (60 mg, 80%) as a yellow oil. ^1H NMR (400 MHz, CD_3OD) δ 7.32–7.11 (m, 4H), 6.78 (d, $J = 9.2$ Hz, 1H), 6.00 (dd, $J = 9.2, 2.5$ Hz, 1H), 5.81 (d, $J = 2.5$ Hz, 1H), 5.24 (d, $J = 1.0$ Hz, 1H), 2.84 (q, $J = 7.0$ Hz, 2H), 2.68–2.47 (m, 6H), 2.47–2.38 (m, 2H), 1.14–1.10 (m, 2H), 0.98–0.74 (m, 4H), 0.73–0.56 (m, 2H), 0.56–0.42 (m, 3H); ^{13}C NMR (100 MHz, CD_3OD) δ 172.0, 169.0, 162.8, 157.5, 153.6, 147.9, 141.2, 139.4, 129.2, 127.6, 126.8 (q, $^1J_{\text{C,F}} = 270.3$ Hz), 126.5, 111.3, 107.0, 105.3, 100.3, 59.9 (q, $^2J_{\text{C,F}} = 44.8$ Hz), 54.9, 47.7, 41.3, 40.5, 30.4, 27.6, 25.6, 12.4; ^{19}F NMR (376.5 MHz, CD_3OD) δ -57.2 (s); HRMS (ESI): m/z calcd for $\text{C}_{27}\text{H}_{30}\text{F}_3\text{N}_6\text{O}_6\text{S}$ $[\text{M}+\text{H}]^+$ 623.1900, found 623.1857.

S18 was synthesized from **TMP** according to procedures reported in literature^[5]. To a solution of **S16** (14 mg, 0.04 mmol) in DMF (4 mL) were added EDCI (15 mg, 0.08 mmol) and NEt_3 (16 μL , 0.08 mmol) under argon atmosphere. The mixture was stirred for 15 min, and **S17** (46 mg, 0.08 mmol) was added. The resulting mixture was stirred for further 12 h at room temperature. The reaction was then concentrated *in vacuo* and was purified by silica gel column chromatography (10% MeOH/ CH_2Cl_2) to afford **4** (19 mg, 58%) as a yellow oil. ^1H NMR (500 MHz, CD_3OD) δ 7.53 (d, $J = 9.2$ Hz, 1H), 7.21 (d, $J = 5.1$ Hz, 1H), 6.75 (dd, $J = 9.2, 2.4$ Hz, 1H), 6.58–6.52 (m, 3H), 5.99 (s, 1H), 4.05 (s, 2H), 3.91 (t, $J = 6.1$ Hz, 2H), 3.79 (s, 6H), 3.64 (s, 2H), 3.58 (q, $J = 7.4$ Hz, 2H), 3.21 (t, $J = 6.9$ Hz, 2H), 3.14 (t, $J = 6.9$ Hz, 2H), 2.24 (t, $J = 7.3$ Hz, 2H), 1.83–1.75 (m, 2H), 1.66–1.73 (m, 2H), 1.55–1.44 (m, 4H), 1.36–1.26 (m, 2H), 1.22 (t, $J = 7.4$ Hz, 3H); ^{13}C NMR (125 MHz, CD_3OD) δ 176.0, 171.7, 166.3, 162.4, 157.3, 156.1, 155.1, 153.3, 140.9, 140.3, 137.2, 133.7, 133.6, 127.6, 126.9, 126.2 (q, $^1J_{\text{C,F}} = 270.3$ Hz), 111.1, 111.0, 107.3, 106.7, 105.0, 100.0, 99.5, 74.0, 56.6, 54.7 (q, $^2J_{\text{C,F}} = 44.8$ Hz), 54.6, 47.5, 40.3, 40.2, 36.8, 33.9, 30.6, 30.08, 30.06, 25.2, 23.7, 12.1; ^{19}F NMR (470.6 MHz, CD_3OD) δ -55.2 (s); HRMS (ESI): m/z calcd for $\text{C}_{38}\text{H}_{47}\text{F}_3\text{N}_9\text{O}_7$ $[\text{M}+\text{H}]^+$ 798.3551, found 798.3546.

Appendix

MS/MS spectra of BSA modification by 1

Raw File Scan Method Score m/z
1 40083 FTMS; HCD 192.14 602.63

Sequence PSMs Fragments Properties
- L K P D P N Q L C D E F K -

Raw File Scan Method Score m/z
1 41766 FTMS; HCD 303.33 989.42

Sequence PSMs Fragments Properties
- Q N G V F Q E C C Q A E D K -

Raw File Scan Method Score m/z
1 40087 FTMS; HCD 141.97 906.42

Sequence	PSMs	Fragments	Properties
- Y	Y20	Y19	Y18
	N	G	V
	b2	b3	b4
	b3	b4	b5
	Y17	Y16	Y15
	Q	E	C
	C	C	Q
	A	E	D
	K	G	A
	C	L	L
	P	K	-

Raw File Scan Method Score m/z
1 39941 FTMS; HCD 234.65 989.49

Sequence	PSMs	Fragments	Properties
- Y	Y13	Y12	Y11
	N	G	V
	b2	b3	b4
	b3	b4	b5
	Y10	Y9	Y8
	Q	E	C
	C	C	Q
	A	E	D
	K	-	-

Raw File Scan Method Score m/z
1 30032 FTMS; HCD 122.5 525.29

Raw File Scan Method Score m/z
1 28376 FTMS; HCD 164 538.81

Sequence	PSMs	Fragments	Properties
- L	S Q K F P K	b2 b3 b4 b5	

Raw File Scan Method Score m/z
1 22350 FTMS; HCD 204.78 782.01

Sequence	PSMs	Fragments	Properties
- V	H K E C C H G D L L E C A D D R	b2 b3 b4 b5 b6 b7 b8 b9 b10 b11 b12 b13 b14 b15 b16	

Raw File Scan Method Score m/z
1 37045 FTMS; HCD 397.31 826.69

Sequence	PSMs	Fragments	Properties
- E	Y16 C	Y17 C	Y15 H
	Y14 D	Y13 L	Y12 L
	Y11 E	Y10 C	Y9 A
	Y8 D	Y7 R	Y6 A
	Y5 D	Y4 L	Y3 A
	Y2 K	Y1 K	

Raw File Scan Method Score m/z
1 38397 FTMS; HCD 355.02 837.38

Sequence	PSMs	Fragments	Properties
- Y	Y11 D	Y10 C	Y9 Y
	Y8 D	Y7 N	Y6 Q
	Y5 D	Y4 T	Y3 I
	Y2 S	Y1 S	

Raw File Scan Method Score m/z
1 26839 FTMS; HCD 226.46 651.82

Sequence	PSMs	Fragment	Properties
- S	<div> <div>y8</div> <div>y7</div> <div>y6</div> <div>y5</div> <div>y4</div> <div>y3</div> <div>y2</div> <div>y1</div> </div>	<div> <div>b2</div> <div>b2</div> <div>b4</div> <div>b5</div> <div>b6</div> <div>b7</div> <div>b8</div> </div>	-

Raw File Scan Method Score m/z
1 32066 FTMS; HCD 223.88 651.31

Sequence	PSMs	Fragment	Properties
- S	<div> <div>y8</div> <div>y7</div> <div>y6</div> <div>y5</div> <div>y4</div> <div>y3</div> <div>y2</div> <div>y1</div> </div>	<div> <div>b2</div> <div>b2</div> <div>b4</div> <div>b5</div> <div>b6</div> <div>b7</div> <div>b8</div> </div>	-

Raw File Scan Method Score m/z
1 33944 FTMS; HCD 135.94 1022.98

Sequence	PSMs	Fragment	Properties
- L A K E Y E A T L E E C C A K -			

Raw File Scan Method Score m/z
1 39770 FTMS; HCD 124.78 623.69

Sequence	PSMs	Fragment	Properties
- K V P Q V S T P L V E V S R -

MS/MS spectra of BSA modification by 2

Raw File Scan Method Score m/z
2 36993 FTMS; HCD 189.77 826.69

Sequence	PSMs	Fragment	Properties
- E	Y18 C	Y17 C	Y16 H
	b2	b3	b4
			b5
			b6
			b7
			b8
			b9
			b10
			b11
			b12
			b13
			b14
			b15
			b16
			b17
			b18
			b19
			b20
			b21
			b22
			b23
			b24
			b25
			b26
			b27
			b28
			b29
			b30
			b31
			b32
			b33
			b34
			b35
			b36
			b37
			b38
			b39
			b40
			b41
			b42
			b43
			b44
			b45
			b46
			b47
			b48
			b49
			b50
			b51
			b52
			b53
			b54
			b55
			b56
			b57
			b58
			b59
			b60
			b61
			b62
			b63
			b64
			b65
			b66
			b67
			b68
			b69
			b70
			b71
			b72
			b73
			b74
			b75
			b76
			b77
			b78
			b79
			b80
			b81
			b82
			b83
			b84
			b85
			b86
			b87
			b88
			b89
			b90
			b91
			b92
			b93
			b94
			b95
			b96
			b97
			b98
			b99
			b100
			b101
			b102
			b103
			b104
			b105
			b106
			b107
			b108
			b109
			b110
			b111
			b112
			b113
			b114
			b115
			b116
			b117
			b118
			b119
			b120
			b121
			b122
			b123
			b124
			b125
			b126
			b127
			b128
			b129
			b130
			b131
			b132
			b133
			b134
			b135
			b136
			b137
			b138
			b139
			b140
			b141
			b142
			b143
			b144
			b145
			b146
			b147
			b148
			b149
			b150
			b151
			b152
			b153
			b154
			b155
			b156
			b157
			b158
			b159
			b160
			b161
			b162
			b163
			b164
			b165
			b166
			b167
			b168
			b169
			b170
			b171
			b172
			b173
			b174
			b175
			b176
			b177
			b178
			b179
			b180
			b181
			b182
			b183
			b184
			b185
			b186
			b187
			b188
			b189
			b190
			b191
			b192
			b193
			b194
			b195
			b196
			b197
			b198
			b199
			b200
			b201
			b202
			b203
			b204
			b205
			b206
			b207
			b208
			b209
			b210
			b211
			b212
			b213
			b214
			b215
			b216
			b217
			b218
			b219
			b220
			b221
			b222
			b223
			b224
			b225
			b226
			b227
			b228
			b229
			b230
			b231
			b232
			b233
			b234
			b235
			b236
			b237
			b238
			b239
			b240
			b241
			b242
			b243
			b244
			b245
			b246
			b247
			b248
			b249
			b250
			b251
			b252
			b253
			b254
			b255
			b256
			b257
			b258
			b259
			b260
			b261
			b262
			b263
			b264
			b265
			b266
			b267
			b268
			b269
			b270
			b271
			b272
			b273
			b274
			b275
			b276
			b277
			b278
			b279
			b280
			b281
			b282
			b283
			b284
			b285
			b286
			b287
			b288
			b289
			b290
			b291
			b292
			b293
			b294
			b295
			b296
			b297
			b298
			b299
			b300
			b301
			b302
			b303
			b304
			b305
			b306
			b307
			b308
			b309
			b310
			b311
			b312
			b313
			b314
			b315
			b316
			b317
			b318
			b319
			b320
			b321
			b322
			b323
			b324
			b325
			b326
			b327
			b328
			b329
			b330
			b331
			b332
			b333
			b334
			b335
			b336
			b337
			b338
			b339
			b340
			b341
			b342
			b343
			b344
			b345
			b346
			b347
			b348
			b349
			b350
			b351
			b352
			b353
			b354
			b355
			b356
			b357
			b358
			b359
			b360
			b361
			b362
			b363
			b364
			b365
			b366
			b367
			b368
			b369
			b370
			b371
			b372
			b373
			b374
			b375
			b376
			b377
			b378
			b379
			b380
			b381
			b382
			b383
			b384
			b385
			b386
			b387
			b388
			b389
			b390
			b391
			b392
			b393
			b394
			b395
			b396
			b397
			b398
			b399
			b400
			b401
			b402
			b403
			b404
			b405
			b406
			b407
			b408
			b409
			b410
			b411
			b412
			b413
			b414
			b415
			b416
			b417
			b418
			b419
			b420
			b421
			b422
			b423
			b424
			b425
			b426
			b427
			b428
			b429
			b430
			b431
			b432
			b433
			b434
			b435
			b436
			b437
			b438
			b439
			b440
			b441
			b442
			b443
			b444
			b445
			b446
			b447
			b448
			b449
			b450
			b451
			b452
			b453
			b454
			b455
			b456
			b457
			b458
			b459
			b460
			b461
			b462
			b463
			b464
			b465
			b466
			b467

Raw File Scan Method Score m/z
2 38211 FTMS; HCD 269.33 837.38

Raw File Scan Method Score m/z
2 38244 FTMS; HCD 201.49 684.31

Sequence	PSMs	Fragments	Properties
- C	<div> <div>y6</div> <div>y7</div> <div>y6</div> <div>y5</div> <div>y4</div> <div>y3</div> <div>y2</div> <div>y1</div> </div>	<div> <div>b2</div> <div>b3</div> <div>b4</div> <div>b6</div> </div>	<div> <div>CS</div> <div>VS</div> <div>NS</div> <div>RS</div> </div>

NMR spectra of compounds

Compound S2

Compound S3

Compound 1

dsy-2-153.001.001.1r.esp

dsy-2-153.002.001.1r.esp

Compound S4

dsy-3-10B-1.001.001.1r.esp

dsy-3-10B-1.002.001.1r.esp

Compound S5

Chemical structure: CCN(CC)c1ccc2c(c1)c(=O)oc(C(=O)NNC(=O)c3ccccc3)c2

¹H NMR spectrum (CDCl₃) showing peaks from 0 to 12 ppm. Integration values are provided below the baseline, and chemical shifts are listed above the peaks.

Chemical Shift (ppm)	Integration
1.120, 1.114, 1.106, 1.093	6.68
3.446, 3.432, 3.418, 3.404, 3.371, 2.500	4.53
5.154, 5.180, 5.200, 5.225, 5.265, 6.110, 6.156, 6.592, 6.613, 6.681, 6.999, 6.903, 6.920, 7.281, 7.299, 7.312, 7.327, 7.321	2.18
7.344, 7.361, 7.396, 7.413, 7.429, 7.443, 7.458, 7.621, 7.635	0.42
8.043	0.68
11.698, 11.161	0.79

Compound S7

dsy-3-141.001.001.1f.esp

dsy-3-141.003.001.1f.esp

Compound 2

dsy-3-30.001.001.1r.esp

Compound 1a

dsy-3-37.007.001.1r.esp

Compound 1b

dsy-3-37.005.001.1r.esp

Compound 1d

dsy-3-37.003.001.1r.esp

Compound 1e

dsy-3-37.001.001.1r.esp

Compound 2c

Compound 2d

DSY-3-46-1.001.001.1R.esp

Compound S8

Compound S9

dsy-4-117-2.001.001.1r

dsy-4-117-2.002.001.1r.esp

Compound S10

Compound S11

Compound S12

Compound S13

Compound S16

Compound 3

Compound 4

References

- [1] K. Ito, J. Maruyama, *Chem. Pharm. Bull.* **1983**, *31*, 3012-3014.
- [2] S. M. Nicolle, C. J. Moody, *Chem. Eur. J.* **2014**, *20*, 4420-4425.
- [3] Q. Lin, Q. Huang, C. Li, C. Bao, Z. Liu, F. Li, L. Zhu, *J. Am. Chem. Soc.* **2010**, *132*, 10645-10647.
- [4] K. Matsuo, Y. Kioi, R. Yasui, Y. Takaoka, T. Miki, S.-h. Fujishima, I. Hamachi, *Chem. Sci.* **2013**, *4*, 2573-2580.
- [5] Y. Takaoka, Y. Nishikawa, Y. Hashimoto, K. Sasaki, I. Hamachi, *Chem. Sci.* **2015**, *6*, 3217-3224.