# ChemComm

## Accepted Manuscript

Volume 52 | Number 1 | 4 January 2016 | Pages 1–216

## ChemComm

Chemical Communications
www.rsc.org/chemcomm

ISSN 1359-7345

COMMUNICATION
Marilyn M. Olmstead, Alan L. Balch, Josep M. Poblet, Luis Echegoyen et al.
Reactivity differences of Sc₃N@C₈₀ (Iₕ and D₅ₕ) and Synthesis of the
first methanofullerene derivatives of Sc₃N@Cₛ-Cₛ

175 YEARS

This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the author guidelines.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the ethical guidelines, outlined in our author and reviewer resource centre, still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

ROYAL SOCIETY
OF CHEMISTRY

rsc.li/chemcomm

# Journal Name

## COMMUNICATION

### Specific fluorescence labeling of target protein by using ligand–4-azidophthalimide conjugate

Kosuke Chiba,[a] Miwako Asanuma,[b,c] Minoru Ishikawa,[a] Yuichi Hashimoto,[a] Kosuke Dodo,[b,c] Mikiko Sodeoka,[b,c] and Takao Yamaguchi*[a,d]

**We herein propose a simple affinity-labeling method using ligand–4-azidophthalimide (AzPI) conjugate. As proof of concept, we show that two different ligand–AzPI conjugates enabled highly specific fluorescence labeling of their individual target proteins even in crude cell lysates. The method was also applied to label endogenous target protein inside living cells.**

Target protein identification (target-ID) is extremely important to elucidate the mechanisms of action of bioactive small molecules. During the past few decades, a number of target-ID methods have been developed.[1] Among them, affinity labeling can be reliably used for specific modification and identification of weakly interacting protein targets, membrane-associated protein targets and target-interacting proteins (partner proteins) under native cellular conditions, which are difficult to achieve by conventional pull-down methods.[2] Affinity labeling is also helpful for identification of ligand-binding sites, because labeling generally occurs on adjacent amino acid residue(s). In general, this method utilizes a chemical probe composed of a bioactive small molecule (*i.e.*, ligand unit), a reactive group (*e.g.* electrophile or photoreactive group) and a tag unit (*e.g.* fluorophore, biotin or radioisotope) (Figure 1a). Upon addition of the chemical probe to a protein mixture or to living cells, the target protein can be selectively labeled by covalent capture through interaction between the ligand unit and the target protein. However, a major obstacle to this method is often the design of the highly functionalized chemical probe, because the probe must retain most of the parent ligand properties, such as target-binding affinity, target-specificity, water solubility and membrane permeability.[3] For this reason, small alkyne/azide handles for subsequent conjugation with the tag unit are commonly used[4] despite the inconvenience of the additional conjugation step(s). To address this issue, we recently presented a simple design strategy for chemical probes bearing only a small alkoxy nitrobenzoxadiazole (*O*-NBD, 180 Da)[5] or 2,3-dichloromaleimide (diCMI, 164 Da)[6] unit. The *O*-NBD and diCMI units are electrophiles that selectively react with a nucleophilic lysine residue near the ligand-binding site to generate amino NBD or 2-amino-3-chloromaleimide fluorophore on the target. We have since focused on photoreactive groups (*e.g.* phenyl azide, diazirine and benzophenone), because of their applicability to any amino acid (*i.e.*, any target protein) *via* C–H/X–H bond insertion reaction. Here, we demonstrate the utility of the compact 4-azidophthalimide (AzPI, 189 Da) unit as a photoactivatable profluorophore based on the structures of photoreactive phenyl azide and fluorescent 4-aminophthalimide[7] (Figure 1b).
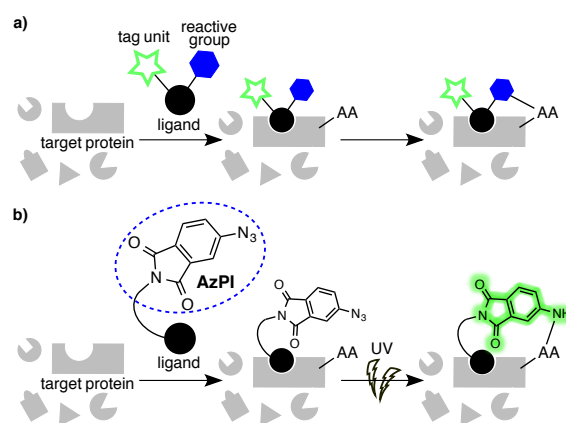


**Figure 1.** Affinity-labeling methods a) using a chemical probe bearing a reactive group and a tag unit (general approach), and b) using a ligand–AzPI conjugate (this work). AA means amino acid.

To test the availability of ligand–AzPI conjugates for affinity labeling, we selected 4-sulfamoylbenzamide and carbonic

a. *Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. E-mail: yamaguchi@iam.u-tokyo.ac.jp*
b. *Synthetic Organic Chemistry Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.*
c. *AMED-CREST, Japan Agency for Medical Research and Development, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.*
d. *Present address: Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: yamaguchi-ta@phs.osaka-u.ac.jp*
† Electronic Supplementary Information (ESI) available: Supplementary results (CA-II-inhibitory activity, photoreaction data, fluorescence spectra and mass spectrometric analysis), synthetic procedure, and ^{1}H and ^{13}C NMR spectra of new compounds. See DOI: 10.1039/x0xx00000x

anhydrase II (CA-II) as a model ligand and target protein.[8] Based on the reported X-ray co-crystal structure of *N*-hexyl-4-sulfamoylbenzamide (**1**) bound to human CA-II (Figure 2a),[9] we designed three chemical probes **P1**–**P3**, bearing the AzPI unit at the end of the linker tail (Figure 2b). These probes were synthesized from 4-sulfamoylbenzoic acid in three steps (see Supporting Information), and their binding abilities to the target protein were evaluated by means of assay of inhibitory activity towards bovine CA-II (see Supporting Information, Figure S1). We found that **P1**–**P3** showed sufficient CA-II-inhibitory activities (IC$_{50}$ = 0.40–0.53 μM), comparable to that of the parent compound **1** (IC$_{50}$ = 0.17 μM). Importantly, this indicated that the small AzPI unit interfered only minimally with the ligand-target interaction.
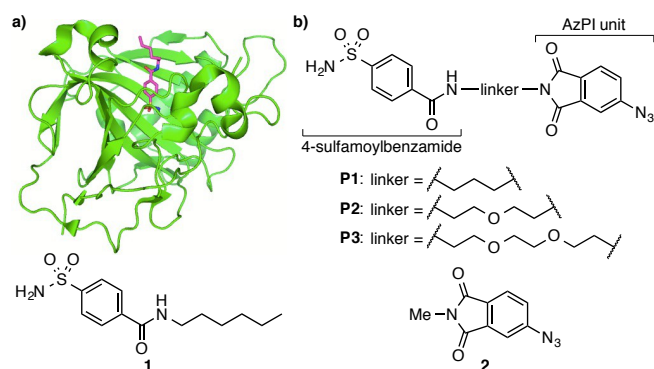


**Figure 2.** Probe design for CA-II labeling. a) X-Ray co-crystal structure of **1** bound to human CA-II (PDB ID: 3RZ8).[9] b) The structures of the designed AzPI probes **P1**–**P3** and control **2**.

Then, labeling of bovine CA-II with **P1**–**P3** was conducted in phosphate buffer (pH 7.0) with or without ultraviolet (UV) irradiation at 365 nm (5 min, 0 °C). After the labeling, the products were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and analysed by in-gel fluorescence imaging (λ$_{ex}$ 440–500 nm, λ$_{em}$ >540 nm) and Coomassie Brilliant Blue (CBB) staining. As shown in Figure 3a, all the chemical probes were found to have labeled the target CA-II, based on fluorescence, in the UV-irradiated condition (lanes 3–5). On the other hand, a control without the ligand unit, 4-azido-*N*-methylphthalimide (**2**), afforded no fluorescence labeling (lane 2). These observations suggested that the ligand unit of **P1**–**P3** was essential for CA-II labeling. To confirm this point, we conducted competition analysis using compound **1** (see Supporting Information, Figure S2). As expected, the CA-II labeling by **P1**, **P2** or **P3** was completely blocked by addition of ten equivalents of **1**, supporting the conclusion that ligand-target interaction was involved in the labeling. In the absence of UV irradiation (Figure 3a, lanes 7–10), no labeling occurred. In this context, the irradiation time and wavelength were briefly optimized using **P3** (see Supporting Information, Figures S3 and S4), and it was found that efficient CA-II labeling proceeded with an irradiation time of ≤5 min (365 nm, 0 °C).

The target-specificity of **P3** was then evaluated by conducting CA-II labeling in the presence of HEK293 cell lysate (Figure

3b). To our delight, the target CA-II was selectively labeled by **P3** even in the presence of a large excess of non-target proteins (lane 2). Addition of compound **1** to this mixture resulted in dose-dependent inhibition of the CA-II labeling (lanes 3 and 4); this again confirmed that the target labeling depends upon ligand-target interaction.
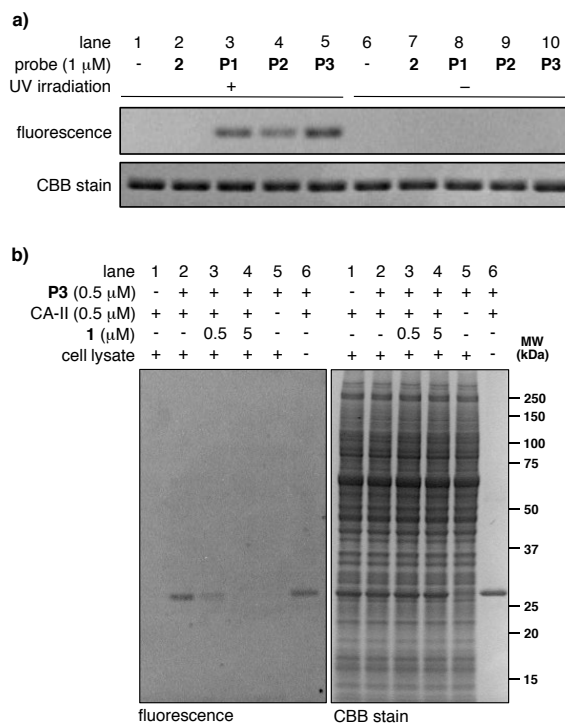


**Figure 3.** Fluorescence labeling of the target CA-II by ligand–AzPI conjugate. a) Labeling of bovine CA-II (1 μM) with or without UV irradiation at 365 nm. b) Specific fluorescence labeling of CA-II by **P3** in the presence of HEK293 cell lysate. Conditions: **P3** (0.5 μM), **1** (0–5 μM), CA-II (0.5 μM = 0.015 mg/mL), HEK293 cell lysate (0.72 mg/mL), phosphate buffer (pH 7.0), and UV irradiation at 365 nm (5 min, 0 °C).

To understand in detail the mechanism of the fluorescence labeling of the target protein by ligand–AzPI conjugate, the photoreaction products were analyzed for simple azidophthalimide compound **2** (see Supporting Information, Table S1). When **2** was irradiated at 365 nm for 1 h in DMSO-$d_6$, dimerized azo compound **3** was obtained as the only isolable product (Table S1, entry 1). In the presence of diethylamine, fluorescent 4-amino-*N*-methylphthalimide (**4**) was generated in 64% yield (Table S1, entry 2). Both the photoreaction products, **3** and **4**, are expected to be generated from **2** *via* triplet nitrene species.[10] It is noteworthy that azepines, which are generally formed by photo-irradiation of phenyl azides, were not isolated. These results support the formation of 4-aminophthalimide fluorophore during the CA-II labeling (see also Supporting Information, Figure S5: UV-Vis absorbance and fluorescence spectra of **4** and its *N*-propyl derivative **5**).

To investigate the generality of the AzPI method and to confirm covalent binding of the ligand–AzPI conjugate to the target protein, we selected biotin and streptavidin as a second model,[5,6] and we designed and prepared probe **P4** and its non-AzPI counterpart **P5** (Figure 4). Labeling of streptavidin was

**ChemComm**

**COMMUNICATION**

examined in phosphate buffer (pH 7.0) with UV irradiation at 365 nm (5 min, 0 ºC). When a mixture of streptavidin and **P4** was irradiated, fluorescence labeling of target streptavidin took place, as expected (Figure 5, lane 2). The covalent binding of **P4** to streptavidin was confirmed by chemiluminescence detection using streptavidin–horseradish peroxidase (streptavidin–HRP) conjugate, which can visualize biotinylated proteins quantitatively. Pretreatment with 50 μM biotin resulted in disappearance of both the fluorescence and chemiluminescence bands (lane 3), confirming affinity-dependent target labeling by **P4**. The fluorescence signals seemed to be well correlated with the chemiluminescence signals, *i.e.*, the amount of covalently captured protein. In contrast, no fluorescence labeling occurred with non-AzPI probe **P5**, despite its covalent attachment to the target streptavidin (lane 4). These results demonstrated the importance of the AzPI structure for the fluorescence labeling of the target protein.
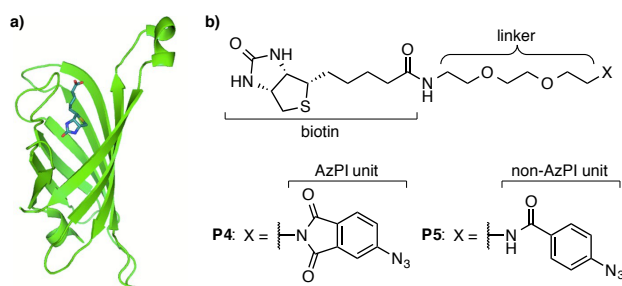


**Figure 4.** Probe design for streptavidin labeling. a) The X-ray co-crystal structure of biotin bound to streptavidin (PDB ID: 1STP).[11] b) The structures of the designed biotin–AzPI conjugate **P4** and its non-AzPI counterpart **P5**.
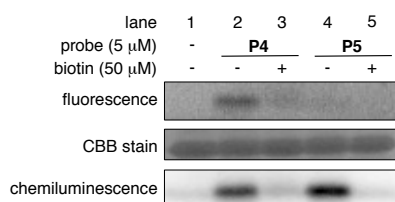


**Figure 5.** Labeling of streptavidin by biotin–AzPI conjugate **P4** and its non-AzPI counterpart **P5**. Conditions: **P4** or **P5** (5 μM), biotin (50 μM), streptavidin (5 μM), phosphate buffer (pH 7.0), and UV irradiation at 365 nm (5 min, 0 ºC).

In the case of the first model, covalent probe-binding to target protein was confirmed by LC-MS/MS analysis.[12] CA-II labeled with **P3** was subjected to tryptic digestion and analyzed by LC-MS/MS (see Supporting Information for the detailed protocol). We identified Ser2–Lys9 [N-terminally acetylated SHHWGYGK], Ala37–Arg57 [AVVQDPALKPLALVYGEATSR], Met59–Lys80 [MVNNGHSFNVEYDDSQDKAVLK] and Val159–Lys169 [VLDALDSIKTK], located close to the entrance of the ligand-binding site, as amino-**P3**-immobilized peptide candidates (Figure 6a, see also Table S2 in the Supporting Information; amino-**P3**-immobilized peptides were identified by the SEQUEST search engine, taking account of the molecular weight gain of **P3** (aminophthalimide form); sequence coverage

was >96%). The Ser2–Lys9 sequence showed the highest ion intensity. Extracted ion chromatography (EIC) of the amino-**P3**-immobilized Ser2–Lys9 indicated the presence of many modification patterns on this sequence (Figure 6b), but MS/MS analysis of the signal at 38.0 min demonstrated that N-terminal Ser2 or His3 had been covalently labeled with **P3** (Figure 6c, see also Figures S8 and S9 in the Supporting Information).[13] These amino acids should be easily accessible, because they are located within about 9–14 Å from the amide nitrogen atom of **P3**, as estimated from the binding model of **1** (Figure 6a). These results show that covalent attachment of amino-**P3** to CA-II occurs near the ligand-binding site. Therefore, the AzPI method should also be helpful for identification of the ligand-binding site.
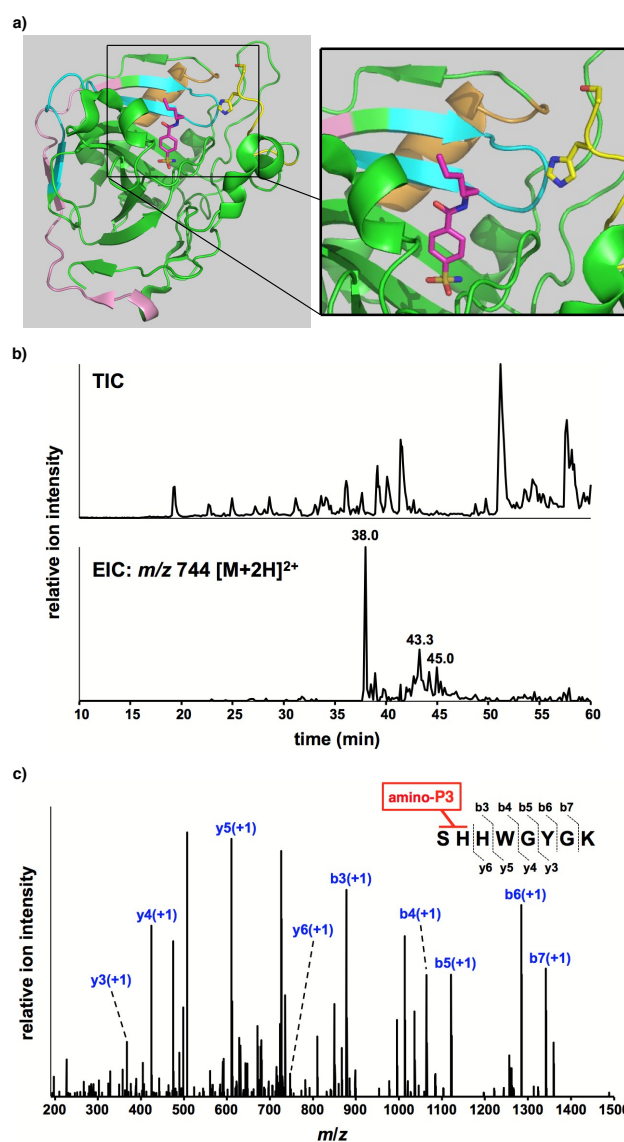


**Figure 6.** Identification of the labeled amino acid by LC-MS/MS analysis. Bovine CA-II was labeled with **P3** under UV irradiation at 365 nm (5 min, 0 ºC), and the products were subjected to tryptic digestion and then analysed by LC-MS/MS. a) Binding model of **1** and bovine CA-II. The model was constructed based on the X-ray co-crystal structure of **1** bound to human CA-II (PDB ID: 3RZ8) and the X-ray structure of bovine CA-II (PDB ID: 1V9E). The sequences corresponding to Ser2–Lys9, Ala37–Arg57, Met59–Lys80 and Val159–Lys169 are shown in

yellow, pink, cyan and orange, respectively. Ser2 and His3 are shown in stick form. b) Chromatograms obtained with total ion current (TIC) (upper) and extracted ion monitoring in the range of *m/z* 744.23–744.34, corresponding to amino-**P3**-immobilized Ser2–Lys9 (lower). c) MS/MS analysis of the amino-**P3**-immobilized Ser–Lys9 peptide (*R*t 38.0 min). The representative MS/MS spectrum of the doubly charged ion at *m/z* 744.29 shows the peptide fragment [Ser2–Lys9] modified with ligand–AzPI conjugate at Ser2 or His3; for details, see Supporting Information.

Finally, we examined CA-II labeling in living cells (Figure 7). Red blood cells (RBCs), which highly express CA-II, were quickly isolated from bovine blood and used immediately. The labeling was performed under UV irradiation at 365 nm, after incubation of RBCs and **P3** for 30 min at room temperature. We found that a 29 kDa protein corresponding to the target CA-II was selectively labeled (lane 3).[14] This indicates that **P3** efficiently permeated through the cell membrane and reached cytosolic CA-II. The fluorescent band disappeared in the presence of **1** (lanes 4 and 5), and no labeling proceeded with compound **2** (lane 2). Thus, the AzPI unit is available for specific fluorescence labeling of target protein inside living cells.
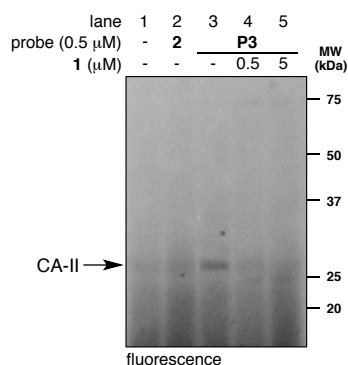


**Figure 7.** Labeling of endogenous CA-II by **P3** in living RBCs. RBCs were quickly isolated from bovine blood and used for the labeling. Conditions: **2** or **P3** (0.5 μM), **1** (0–5 μM), RBCs, HEPES-buffered saline, and UV irradiation at 365 nm (5 min).[12]

In summary, the results of the two case studies (4-sulfamoylbenzamide–CA-II and biotin–streptavidin) confirm that the ligand–AzPI conjugate strategy enables specific fluorescence labeling of the target protein. Not only purified CA-II, but also endogenous CA-II in living cells was selectively labeled and visualized without the need for complex procedures. Since the AzPI unit is compact, it may have relatively little effect on many ligand-target interactions, and thus our simple affinity-labeling method should be available for a wide range of target-ID studies, as well as for determination of ligand-binding sites.

## Notes and references

1　(a) H. Park, J. Ha, J. Y. Koo, J. Park and S. B. Park, *Chem. Sci.*, 2017, **8**, 1127–1133; (b) S. Ziegler, V. Pries, C. Hedberg and H. Waldmann, *Angew. Chem. Int. Ed.*, 2013, **52**, 2744–2792; (c) M. Schenone, V. Dančík, B. K. Wagner and P. A. Clemons, *Nature Chem. Biol.*, 2013, **9**, 232–240; (d) F. Cong, A. K. Cheung and S.-M. A. Huang, *Annu. Rev. Pharmacol. Toxicol.*, 2012, **52**, 57–78; (e) S. Sato, A. Murata, T. Shirakawa and M. Uesugi, *Chem. Biol.*, 2010, **17**, 616–623; (f) G. C. Terstappen, C. Schlüpen, R. Raggiaschi and G. Gaviraghi, *Nature Rev. Drug Discov.*, 2007, **6**, 891–903.

2　(a) S. Pan, H. Zhang, C. Wang, S. C. L. Yao and S. Q. Yao, *Nat. Prod. Rep.*, 2016, **33**, 612–620; (b) J. Park, M. Koh and S. B. Park, *Mol. BioSyst.*, 2013, **9**, 544–550; (c) G. M. Simon, M. J. Niphakis and B. F. Cravatt, *Nature Chem. Biol.*, 2013, **9**, 200–205; (d) J. Lee and M. Bogyo, *Curr. Opin. Chem. Biol.*, 2013, **17**, 118–126.

3　(a) J. Park, M. Koh, J. Y. Koo, S. Lee and S. B. Park, *ACS Chem. Biol.*, 2016, **11**, 44–52; (b) Z. Li, D. Wang, L. Li, S. Pan, Z. Na, C. Y. J. Tan and S. Q. Yao, *J. Am. Chem. Soc.*, 2014, **136**, 9990–9998; (c) Z. Li, P. Hao, L. Li, C. Y. J. Tan, X. Chen, G. Y. J. Chen, S. K. Sze, H.-M. Shen and S. Q. Yao, *Angew. Chem. Int. Ed.*, 2013, **52**, 8551–8556.

4　(a) H. Shi, C.-J. Zhang, G. Y. J. Chen and S. Q. Yao, *J. Am. Chem. Soc.*, 2012, **134**, 3001–3014; (b) T. Hosoya, T. Hiramatsu, T. Ikemoto, M. Nakanishi, H. Aoyama, A. Hosoya, T. Iwata, K. Maruyama, M. Endo and M. Suzuki, *Org. Biomol. Chem.*, 2004, **2**, 637–641.

5　T. Yamaguchi, M. Asanuma, S. Nakanishi, Y. Saito, M. Okazaki, K. Dodo and M. Sodeoka, *Chem. Sci.*, 2014, **5**, 1021–1029.

6　K. Chiba, Y. Hashimoto and T. Yamaguchi, *Chem. Pharm. Bull.*, 2016, **64**, 1647–1653.

7　V. Wintgens and C. Amiel, *J. Photochem. Photobiol. A Chem.*, 2004, **168**, 217–226, and references therein.

8　(a) M. Denda, T. Morisaki, T. Kohiki, J. Yamamoto, K. Sato, I. Sagawa, T. Inokuma, Y. Sato, A. Yamaguchi, A. Shigenaga and A. Otaka, *Org. Biomol. Chem.*, 2016, **14**, 6244–6251; (b) S. Sato and H. Nakamura, *Angew. Chem. Int. Ed.*, 2013, **52**, 8681–8684; (c) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura and I. Hamachi, *Nat. Chem. Biol.*, 2009, **5**, 341–343.

9　J. Mecinović, P. W. Snyder, K. A. Mirica, S. Bai, E. T. Mack, R. L. Kwant, D. T. Moustakas, A. Héroux and G. M. Whitesides, *J. Am. Chem. Soc.*, 2011, **133**, 14017–14026.

10　S. Bräse and K. Banert in *Organic Azides: Synthesis and Applications*, Eds.: N. Gritsan and M. Platz, Wiley-VCH, Weinheim, 2010, Chapter 11.

11　P. C. Weber, D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme, *Science*, 1989, **243**, 85–88.

12　MS analysis of the whole protein showed that the labeling occurred efficiently (estimated labeling yield: 29%) to the target CA-II (see Supporting Information, Figure S6).

13　Since the labeling mainly occurred on the sequence Ser2–Lys9, firm MS/MS data were not obtained for the other labeled peptides (Ala37–Arg57, Met59–Lys80, and Val159–Lys169).

14　Smear in the low molecular weight region seems to be derived from the fluorescence of hemoglobin.

**Table of contents entry**
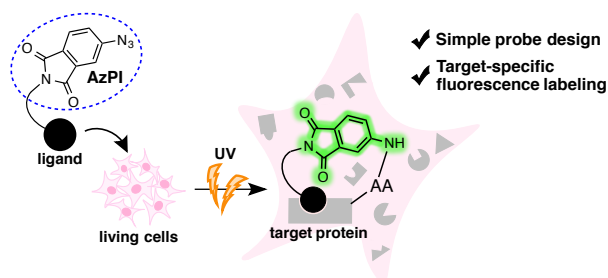


Two distinct model studies demonstrate that the ligand–4-azidophthalimide conjugate strategy is useful for specific fluorescence labeling of target protein.