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 **Information for Authors**.

Journal Name

Chemical Science

ARTICLE TYPE

Ligand-directed Dibromophenyl Benzoate Chemistry for Rapid and Selective Acylation of Intracellular Natural Proteins

Yousuke Takaoka,[a,b] **Yuki Nishikawa,**[a] **Yuki Hashimoto,**[a] **Kenta Sasaki,**[a] **and Itaru Hamachi***[a,c]

A rapid and selective ligand-directed chemical reaction was developed for acylation of proteins in living cells on the basis of ligand-directed chemistry. By fine tuning the reactivity and stability of the phenyl ester derivatives, we successfully identified *ortho*-dibromophenyl benzoate as the optimal reactive motif. It was sufficiently stable in aqueous buffer, hydrolyzing less than 10% after 13 h incubation, but reactive
10 enough for efficient and selective protein labeling in living mammalian cells, as well as *in vitro* (so-called, ligand-directed dibromophenyl benzoate (LDBB) chemistry). Using this chemistry, various fluorophores can be tethered to the target protein directly, which allows fluorescent visualization of the labeled protein in live cells with different colored fluorophore groups (including coumarin, fluorescein and rhodamine). Furthermore, this labeling is applicable to not only an overexpressed protein (*E. coli*
15 dihydrofolate reductase) but also endogenous human carbonic anhydrase II and XII under living cell conditions. LDBB chemistry is a new entry of ligand-directed protein labeling methods, and should be particularly useful for imaging of natural proteins in living cells.

Introduction

Natural protein labeling by synthetic fluorophores in living cells
20 is powerful for investigating protein localization, quantification or function under native conditions, owing to the small size, colorfulness or micro-environment sensitivity of fluorophores, which are distinct from fluorescent proteins.[1,2,3] We have recently proposed ligand-directed (LD) chemistry as a new strategy for
25 non-invasive and specific labeling of endogenous proteins in live cells. Among them, ligand-directed tosyl (LDT) chemistry allowed for covalent labeling of natural proteins with synthetic fluorophores, by which we constructed a semisynthetic biosensor inside living cells.[4] Fluorophore tethering to natural membrane-
30 bound proteins was also carried out by recently developed ligand-directed acyl imidazole (LDAI) chemistry.[5] Fenical and co-workers independently reported another type of LD chemistry using a phenyl-ester (LDPE) for identification and visualization of a natural product target protein.[6] Despite their high selectivity,
35 these methods often suffer from slow reaction rates (about 10 h of reaction time) and low labeling efficiency, which hampers the fluorescent imaging of intracellular proteins.[4,5] On the other hand, a simple increment in the reactivity of the labeling reagents to accelerate the reaction would result in non-specific reactions
40 toward non-target proteins and/or nonproductive hydrolysis of the reagents. Clearly, for selective imaging of "intracellular" protein by LD chemistry, a balanced reactive motif equipped with the two conflicting capabilities; selectivity and fast kinetics, is required.[7]
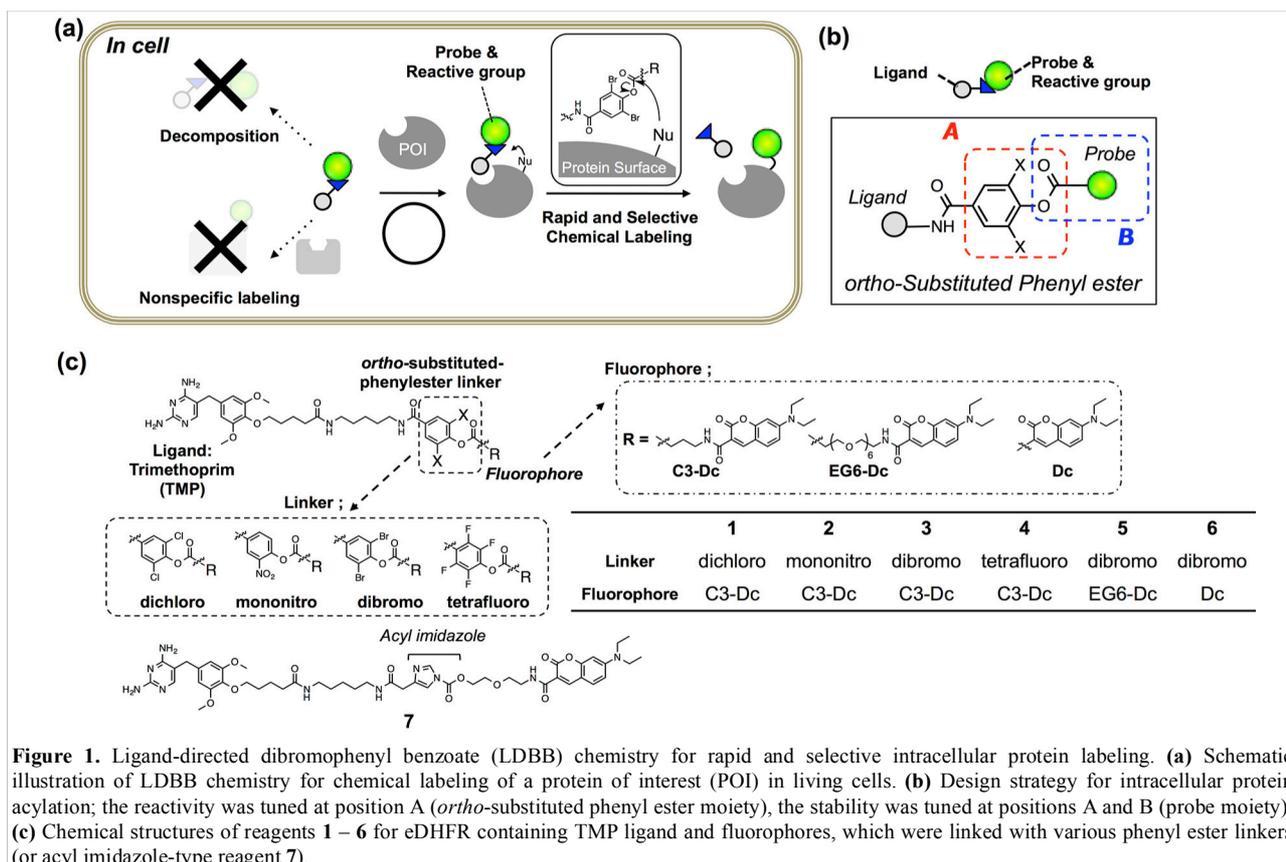
Toward this goal, we focused on protein acylation using
45 Fenical's LDPE chemistry because the phenyl esters can

potentially modify Lys residues that are often abundant on protein surfaces[8] and sufficiently reactive under the physiological conditions.[9] We sought to tune both the reactivity and stability of the LDPE reagents, in order to obtain optimal PE derivatives for
50 "selective but rapid" intracellular protein labeling (**Figure 1a**). The reactivity of the substituted PE group was initially controlled by changing the substituent on the phenol ring (**Figure 1b**, part A). The stability was optimized by using the sterically hindered PE esters, composed of an "*ortho*"-substituted phenol and an
55 alkyl/benzoic acid, both of which are expected to be effective for minimizing nonspecific reactions and enzymatic decomposition (**Figure 1b**, part B). From the screening of several LDPE reagents, *ortho*-dibromophenyl benzoate derivatives were found to exhibit the moderate reactivity required for selective and rapid
60 protein acylation of dihydrofolate reductase (eDHFR) in *E. coli* cell lysates, as well as *in vitro*. By using this *L*igand-*D*irected di*B*romophenyl *B*enzoate (LDBB) chemistry, we successfully labeled and imaged not only overexpressed eDHFR as a model protein, but also endogenously expressed human carbonic
65 anhydrase II and XII (hCAII, XII) in living mammalian cells.

Results and Discussion

Molecular design of *ortho*-substituted phenyl ester reagents

Escherichia coli dihydrofolate reductase (eDHFR) and trimethoprim (TMP) were chosen as the initial target-ligand pair
70 for screening of the reaction moieties because this ligand-protein binding is known to be specific even in mammalian cells (the K_d value was about 10 nM), and a reactive Lys is located near the ligand binding pocket of eDHFR.[10-12] Accordingly, we designed

Figure 1. Ligand-directed dibromophenyl benzoate (LDBB) chemistry for rapid and selective intracellular protein labeling. **(a)** Schematic illustration of LDBB chemistry for chemical labeling of a protein of interest (POI) in living cells. **(b)** Design strategy for intracellular protein acylation; the reactivity was tuned at position A (*ortho*-substituted phenyl ester moiety), the stability was tuned at positions A and B (probe moiety). **(c)** Chemical structures of reagents **1 – 6** for eDHFR containing TMP ligand and fluorophores, which were linked with various phenyl ester linkers (or acyl imidazole-type reagent **7**).

four types of new reagents **1 – 4** containing a TMP ligand and a synthetic fluorophore (7-diethylaminocoumarin (Dc)), connected with *ortho*-substituted phenyl ester linkers, such as 3,5-dichloro-, 3-nitro-, 3,5-dibromo- and 2,3,5,6-tetrafluoro-4-hydroxy-benzoic acid (**Figure 1c**). The theoretical pK_a values of these leaving groups are 5.72, 5.68, 5.59 and 4.17, and therefore the expected order of the reactivity was **4 > 3 > 2 > 1**. To evaluate the effect of steric hindrance, we also prepared **5**, having a flexible-type ethyleneglycol linker between the reactive group and Dc, or **6**, having no linker. Additionally, acyl imidazole-type control compound **7**, having a TMP ligand and Dc fluorophore was prepared.[5] The compounds were synthesized according to the synthetic protocols shown in the supporting information. Unfortunately, we could not isolate the tetrafluoro-type reagent **4**, because of the lower stability of this reactive group under the final condensation conditions.

Optimization of the reactivity of PEL reagents in a test tube

In vitro experiments were first conducted to examine which is the best reactive group for rapid and selective labeling of eDHFR. The labeling reactions of purified eDHFR (10 μM) with Dc-appended PEL reagents **1 – 3** (20 μM) were analyzed by SDS-PAGE and the fluorescence gel imaged. As shown in **Figure 2a**, fluorescence was clearly observed from the band for eDHFR after a 3-h incubation with **2** or **3**, whereas no fluorescence was observed upon incubation with **1**. In addition, no significant bands were observed when eDHFR was incubated with **2** or **3** in

the presence of methotrexate (MTX, 100 μM), a strong inhibitor of eDHFR, indicating that these labeling reactions were driven by a specific ligand-protein interaction. The detailed labeling kinetics, which were evaluated by MALDI-TOF MS analyses, clearly showed that the initial rate of the labeling reaction with **3** was 3.4-fold faster than that with **2** (**Figures 2b** and **S1**). These results showed that the *ortho*-dibromophenyl ester (di-Br)-type of reactive group was useful for LD chemistry. Next, the spacer effect between Dc and the di-Br-type reactive group was evaluated using **3**, **5** and **6** having an alkyl chain, oligoethylene glycol and no spacer, respectively. In all cases, the molecular mass corresponding to Dc-labeled eDHFR was observed as shown in **Figures 2b** and **S1**. The kinetic analyses of the labeling using these reagents revealed that the initial rates of these reactions were almost the same for all three reagents (0.14, 0.16 and 0.17 μM^{-1} min^{-1} for **3**, **5** and **6**, respectively, **Figure 2b**). These were over 7-fold faster than that obtained using acyl imidazole reagent **7** (0.02 μM^{-1} min^{-1} for **7**), indicating that this new LD chemistry is more rapid than LDAI chemistry, at least *in vitro*. We also confirmed that the reaction was site specific using **6** by subjecting the Dc-labeled eDHFR to proteolytic digestion. Conventional peptide mapping analysis revealed that Lys32, located near the ligand-binding site, was predominantly modified (over 94%) by the acylation reaction (**Figure S2**).

Evaluation of the stability of LDBB reagents *in vitro* and the reaction specificity under crude conditions.
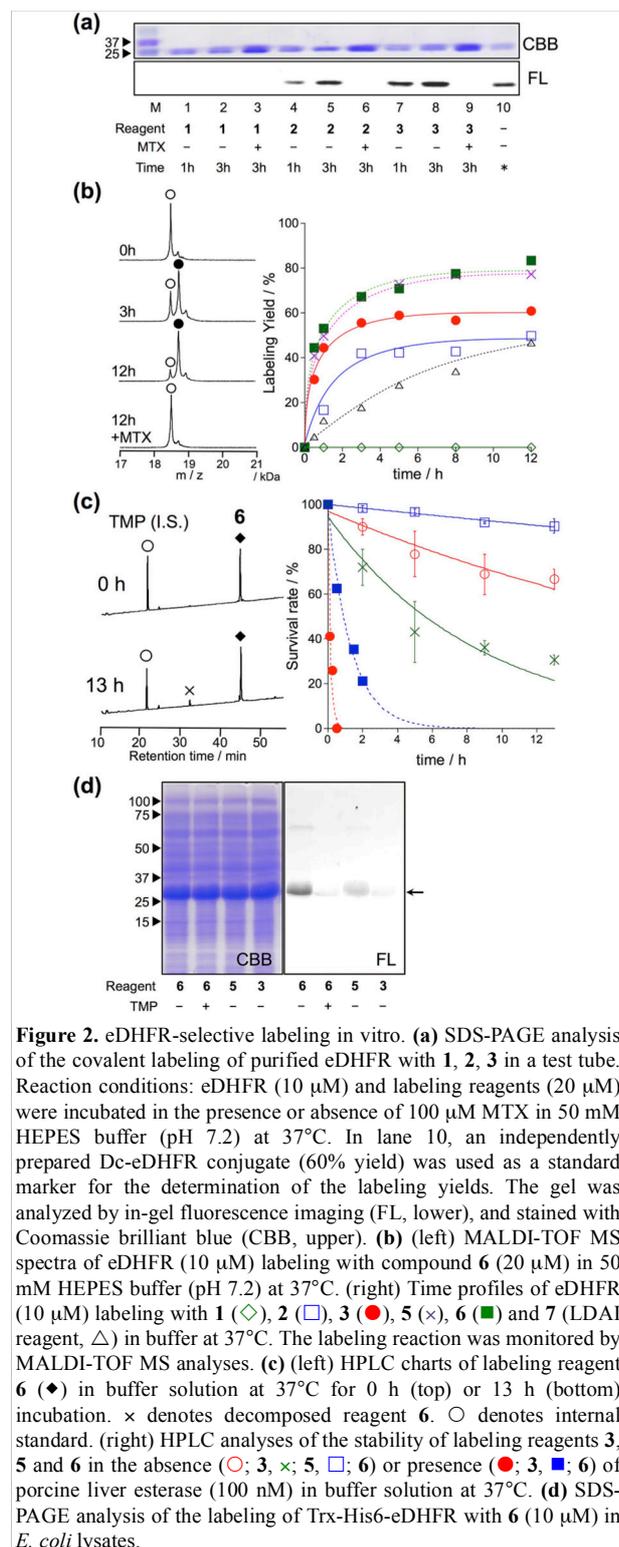
Figure 2. eDHFR-selective labeling in vitro. **(a)** SDS-PAGE analysis of the covalent labeling of purified eDHFR with **1**, **2**, **3** in a test tube. Reaction conditions: eDHFR (10 μM) and labeling reagents (20 μM) were incubated in the presence or absence of 100 μM MTX in 50 mM HEPES buffer (pH 7.2) at 37°C. In lane 10, an independently prepared Dc-eDHFR conjugate (60% yield) was used as a standard marker for the determination of the labeling yields. The gel was analyzed by in-gel fluorescence imaging (FL, lower), and stained with Coomassie brilliant blue (CBB, upper). **(b)** (left) MALDI-TOF MS spectra of eDHFR (10 μM) labeling with compound **6** (20 μM) in 50 mM HEPES buffer (pH 7.2) at 37°C. (right) Time profiles of eDHFR (10 μM) labeling with **1** (◇), **2** (□), **3** (●), **5** (×), **6** (■) and **7** (LDAI reagent, △) in buffer at 37°C. The labeling reaction was monitored by MALDI-TOF MS analyses. **(c)** (left) HPLC charts of labeling reagent **6** (◆) in buffer solution at 37°C for 0 h (top) or 13 h (bottom) incubation. × denotes decomposed reagent **6**. ○ denotes internal standard. (right) HPLC analyses of the stability of labeling reagents **3**, **5** and **6** in the absence (○; **3**, ×; **5**, □; **6**) or presence (●; **3**, ■; **6**) of porcine liver esterase (100 nM) in buffer solution at 37°C. **(d)** SDS-PAGE analysis of the labeling of Trx-His6-eDHFR with **6** (10 μM) in *E. coli* lysates.

Next, we evaluated the stabilities of the reagents under the labeling conditions by HPLC analyses. The half-lives of the reagents in aqueous buffer were 19.5 h and 5.6 h for **3** and **5**, whereas only 10% of **6** was hydrolyzed even after incubation for
5 13 h at 37°C (**Figures 2c** and **S3**). Moreover, upon addition of an esterase, **3** was completely decomposed after 30 min incubation,
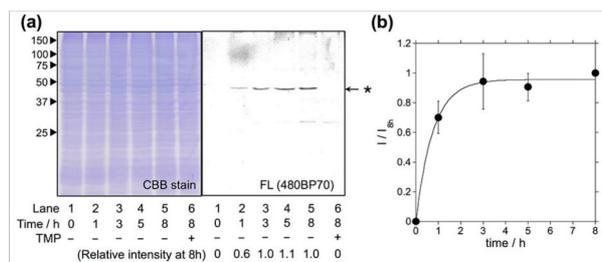

Figure 3. Intracellular eDHFR-EGFP labeling with optimized reagent **6** in HeLa-DG cells. (a) SDS-PAGE analyses of the labeling reaction of **6** (1 μM) in HeLa-DG cells. The gel was analyzed by in-gel fluorescence imaging (right) and stained with CBB (left). Lane 6 shows the reaction in the presence of TMP (10 μM). (b) Time course of eDHFR-EGFP labeling in HeLa-DG cells. The fluorescent intensity (I) was normalized with the intensity for 8 h incubation (I_{8h}). The experiments were performed in triplicate to obtain mean and standard deviation values (shown as error bars)

whereas 20% of **6** remained after 2 hours (**Figures 2c** and **S3**). These results clearly indicated that **6** was resistant toward both autolysis in aqueous buffer, and also catalytic hydrolysis with
10 esterase, maybe because of the steric hindrance provided by the *ortho*-dibromophenyl benzoate (the CPK model is shown in **Figure S4**). Overall, the results clearly demonstrated that ligand-directed dibromophenyl benzoate (LDBB) reagent **6**, which is directly connected with Dc, has moderate reactivity (derived from
15 the leaving-group ability of dibromophenol) and stability (derived from the sterically hindered benzoate structure) for use in a rapid and specific eDHFR labeling reaction *in vitro*. To investigate the labeling selectivity of these reagents under crude conditions, the reactions were carried out using cell lysates of *E. coli* that
20 overexpressed thioredoxin-tag fused eDHFR. A clear single fluorescent band corresponding to tag-fused eDHFR appeared when the cells were incubated with **6** (**Figure 2d**). On the other hand, insensible fluorescence was observed upon addition of an excess amount of TMP, indicating that this labeling reaction was
25 driven by a specific ligand-protein interaction. By using **3** or **5**, the labeling yield was apparently lower than using **6**. Taken together, these results showed that **6** is sufficiently stable and can label eDHFR specifically even in cell lysates containing various reactive biomolecules.

30 **eDHFR labeling with Dc-type reagents in living mammalian cells.**

Having optimized the Dc-type of LDBB labeling reagent, we next attempted to modify intracellular eDHFR in mammalian cells. To assess the labeling specificity of the reagents, we used
35 human epithelial HeLa cells stably expressing eDHFR fused with green fluorescent protein (GFP) (HeLa-DG). The HeLa-DG cells were incubated in a medium containing reagent **6** at 37°C, and washed three times with the fresh medium. The cells were then lysed and analyzed by SDS-PAGE, fluorescence gel imager or
40 western blotting (**Figure 3** and **S5**). Notably, a single fluorescence band corresponding to Dc-labeled eDHFR-GFP was detected, despite the presence of various other proteins as shown in the CBB staining image. The addition of TMP ligand completely blocked the labeling, which indicates that this
45 modification occurs via a specific ligand-eDHFR interaction even inside cells. The labeling yield was determined to be 85 ± 5% of the total eDHFR-GFP, which was calculated from the

fluorescence intensity of Dc and the chemiluminescence intensity of anti-GFP antibody (**Figure S5c**). The time-dependence experiments revealed that the new reagent **6** labeled 50% of eDHFR-GFP with a $t_{1/2}$ of 30 min (**Figure 3b**). Conversely,
5 several proteins other than eDHFR-GFP were also labeled nonspecifically by using **5**, whereas no eDHFR label was observed by using **3** (**Figure S6a**). These results may be explained by the lower selectivity or reactivity of reagents **3** and **5** compared with **6**, similar to results obtained for the *in vitro*
10 experiments. In addition, LDAI reagent **7** did not label intracellular eDHFR-GFP (**Figure S6b**), perhaps because of the lower stability or cellular permeability of this reagent. Together, these results clearly demonstrated that the rapid and quantitative labeling of intracellular natural protein could be achieved using
15 LDBB chemistry in living mammalian cells.

Live-cell imaging of native eDHFR by covalent modification of a synthetic fluorophore.

The in-cell eDHFR-GFP labeling results led us to investigate whether LDBB chemistry can be applied to intracellular protein
20 imaging by covalent modification of a synthetic fluorophore. HeLa-DG cells were incubated in a medium containing reagent **6** (50 nM) for 3 h at 37°C, and then washed three times with fresh medium, and the cells were observed by confocal laser scanning microscopy (CLSM). From the CLSM images of HeLa-DG cells
25 using the GFP channel, eDHFR-GFP was found to be distributed throughout the intracellular region including the cytosol and nuclei. As shown in **Figure 4a**, the fluorescence of Dc was clearly detected inside the cells and merged with the GFP image. However, before the labeling reaction, the addition of TMP
30 resulted in the disappearance of Dc fluorescence (**Figure 4b**). These results clearly showed that the new LDBB reagent **6** can be used to specifically label and visualize intracellular eDHFR-GFP using a synthetic fluorophore.

To assess the applicability of fluorophores other than Dc with
35 LDBB chemistry, we prepared two more fluorophores-appended reagents for labeling of intracellular eDHFR (**Figure 5a**, tetramethylrhodamine (TMR, **8**) and acetylfluorescein (AcFl, **9**)), according to the molecular design principle of LDBB chemistry as discussed above.[13,14] **Figure 5b** shows that the TMR-type of
40 reagent **8** was useful to visualize the localization of eDHFR-GFP in live HeLa-DG cells to a similar degree as **6**. Reagent **9** can also label and image cyan fluorescent protein-fused eDHFR (eDHFR-CFP), which was transiently expressed in live HeLa cells (**Figure 5c**) (we chose CFP as an alternative marker to GFP because the
45 emissions of fluorescein and GFP overlap). These results clearly showed that LDBB chemistry is suitable for fluorescent live-cell imaging of intracellular natural protein with blue, green and red emission, and this imaging tool may be applicable for concomitant use with various other imaging technologies, such as
50 fluorescent protein tags or organelle staining probes.

Endogenous intracellular protein labeling and imaging using a LDBB reagent.

Finally, we attempted to label "endogenous" protein in live cells by altering the affinity ligand, owing to the modular design
55 of LDBB reagents. Human carbonic anhydrase II (hCAII), a cytosolic protein, and XII (hCAXII), a membrane-bound protein, were chosen as the target protein endogenously expressed in
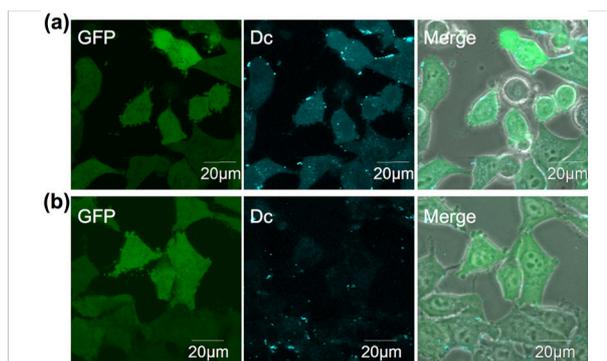

Figure 4. (a, b) Live-cell imaging of eDHFR-GFP by **6** (50 nM) in the absence (**a**) or presence (**b**) of TMP (50 μM) in HeLa-DG cells. Fluorescence images of GFP, Dc and their merges on phase contrast images are displayed in the left, middle and right of each panel, respectively. Images of Dc were obtained at an excitation at 405 nm by using a 420-480-nm emission filter, those of GFP were obtained at an excitation at 473 nm by using a 490-520-nm emission filter. Scale bar: 20 μm.
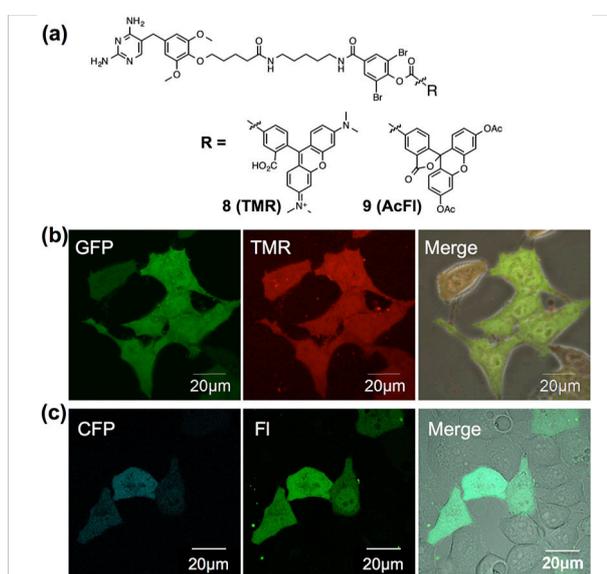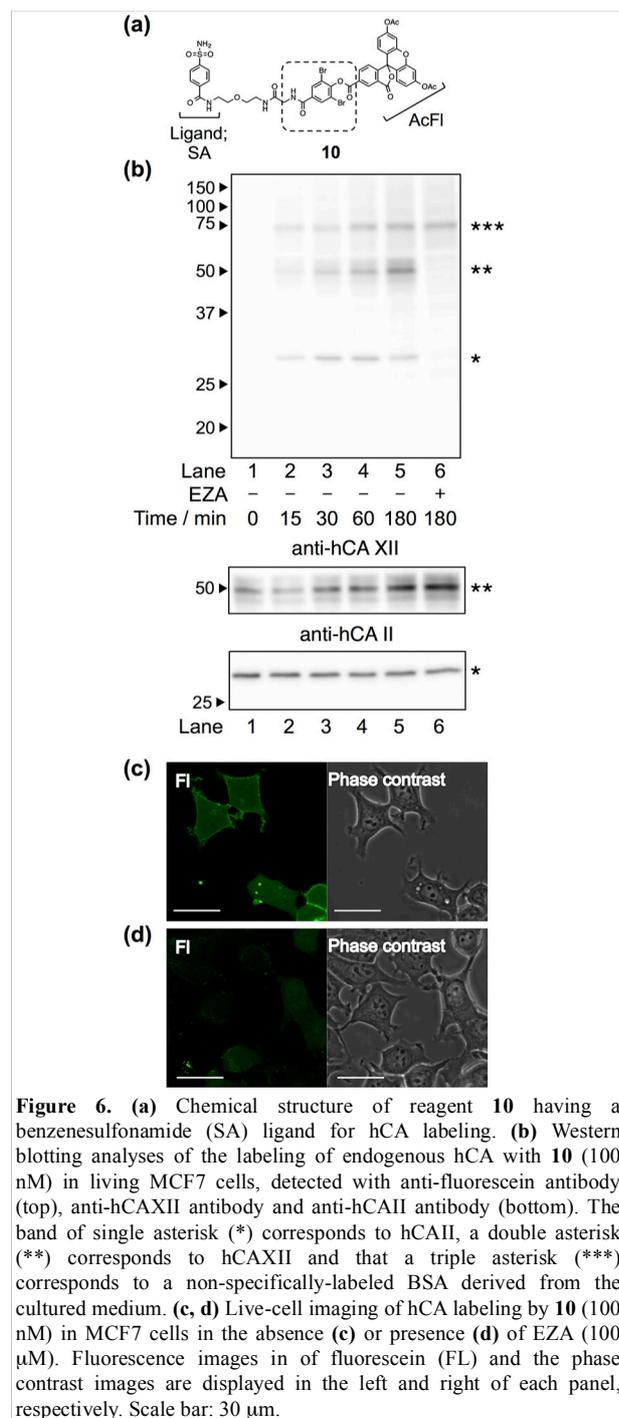

Figure 5. (a) Chemical structures of LD-BB reagents **8** and **9**. **(b)** Live-cell imaging of eDHFR-GFP by **8** (**b**, 20 nM) in HeLa-DG cells. Fluorescence images of GFP, TMR and their merged phase contrast images are displayed in the left, middle and right of each panel, respectively. Images of GFP were obtained at an excitation at 473 nm by using a 490-520-nm emission filter and those of TMR were obtained at an excitation at 559 nm by using a 570-670-nm emission filter, respectively. **(c)** Live-cell imaging of eDHFR-CFP by **9** (50 nM) in HeLa cells expressing eDHFR-CFP. Fluorescence images in of CFP, fluorescein (Fl) and their merged differential interference contrast images are displayed in the left, middle and right of each panel, respectively. Images of CFP were obtained at an excitation at 457 nm by using a 470-490-nm emission filter and those of Fluorescein (Fl) were obtained at an excitation at 515 nm by using a 520-620-nm emission filter. Scale bar: 20 μm.

MCF7 cells (breast cancer cell line). Accordingly, we newly synthesized AcFl-type LDBB reagent **10** containing
60 benzenesulfonamide as the ligand selective for hCA family (**Figure 6a**, K_i of 260 nM).[15,16] MCF7 cells were incubated with

Figure 6. **(a)** Chemical structure of reagent **10** having a benzenesulfonamide (SA) ligand for hCA labeling. **(b)** Western blotting analyses of the labeling of endogenous hCA with **10** (100 nM) in living MCF7 cells, detected with anti-fluorescein antibody (top), anti-hCAXII antibody and anti-hCAII antibody (bottom). The band of single asterisk (*) corresponds to hCAII, a double asterisk (**) corresponds to hCAXII and that a triple asterisk (***) corresponds to a non-specifically-labeled BSA derived from the cultured medium. **(c, d)** Live-cell imaging of hCA labeling by **10** (100 nM) in MCF7 cells in the absence **(c)** or presence **(d)** of EZA (100 μM). Fluorescence images in of fluorescein (FL) and the phase contrast images are displayed in the left and right of each panel, respectively. Scale bar: 30 μm.

10 at 37°C, and then the labeling of hCA was evaluated by western blotting analyses. **Figure 6b** showed that after only 15 to 30 min of incubation, two distinct bands could be detected at ~30 kDa and ~50 kDa by anti-fluorescein antibody, corresponding to the band obtained from western blotting with the anti-hCAII antibody (~30 kDa) and anti-hCAXII antibody (~50 kDa). No band was observed in the presence of ethoxzolamide (EZA, a strong inhibitor for the hCA family of enzymes), which clearly demonstrated that **10** can covalently label not only intracellular hCAII but also cell-surface hCAXII with strict specificity in live

cells.[17] Using the SA-type LDBB reagent *in vitro*, the labeling sites for hCAI or II were identified as Lys171 or Lys169/His3, respectively (Figure S9 and S10). These amino acids were again located near the binding site of each target protein (Figure S9e and S10e), demonstrating that the LDBB reagents were able to react not only lysine, but also (at least) histidine, based on the proximity effect. These results also implied the general applicability of the LDBB chemistry for efficient labeling of various proteins having nucleophilic amino acids appropriately located near the ligand binding site. CLSM images also showed that strong fluorescence was clearly observed from both the cell membrane and cytosolic part of MCF7 cells. In contrast, no fluorescence was detected in the presence of EZA. Taken together, these results clearly demonstrated that the LDBB chemistry is sufficiently rapid, efficient and specific labeling technique for endogenous protein imaging both on the cell surface and inside the cell.

Conclusion

Herein, we demonstrated that LDBB chemistry is useful for selective and rapid labeling of not only overexpressed protein, but also endogenous proteins in mammalian cells. By finely tuning the reactivity and stability of the phenyl ester moiety, it was found that *ortho*-dibromophenyl benzoate was sufficiently stable in aqueous buffer (decomposed less than 10% after 13 h incubation), yet highly reactive in living cells and *in vitro* (over 85% labeling yield was achieved for overexpressed eDHFR in mammalian cells). Moreover, owing to the reactive amino acid preferences and the modular design principle, the LDBB reagents would label various natural (or engineered) proteins having a nucleophilic amino acids (including at least lysine and histidine) located near the ligand binding site, as shown in the cases of the labeling of eDHFR and hCAs.[11c] The present study also showed that LDBB chemistry is a powerful tool for natural protein imaging in living mammalian cells, because of the fact that various benzoate-type fluorophores (including Dc, FL and TMR) can be easily installed on the target proteins with sufficient selectivity and rapidity. We envision that extension of this ligand-directed chemistry should allow for multi-color imaging of various endogenous proteins and analyses of protein-protein interactions in living cells in the future.

Experimental procedures

Synthesis. All synthetic procedures and compound characterization are described in the Supporting Information.

General materials and methods. All chemical reagents and solvents were obtained from commercial suppliers (Aldrich, Tokyo Chemical Industry (TCI), Wako Pure Chemical Industries, Acros Organics, Sasaki Chemical, or Watanabe Chemical Industries) and used without further purification. UV-visible spectra were recorded on a Shimadzu UV-2550 spectrometer. All reactions were carried out under an atmosphere of argon or nitrogen unless otherwise noted. Thin layer chromatography (TLC) was performed on silica gel 60 F254 precoated aluminum sheets (Merck) and visualized by fluorescence quenching or ninhydrin staining. Chromatographic purification was accomplished using flash column chromatography on silica gel

60 N (neutral, 40–50 mm, Kanto Chemical). ^{1}H NMR spectra of samples were recorded on Varian Mercury 400 (400 MHz) spectrometers. ^{13}C NMR spectra of samples were recorded on Jeol ECX-400P (100 MHz) spectrometers. IR spectra were recorded on a JASCO FT/IR-410. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) spectra were recorded on an Autoflex III instrument (Bruker Daltonics) using α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix. High-resolution mass spectra were measured on an Exactive (Thermo Scientific, CA, USA) equipped with electron spray ionization (ESI). SDS-PAGE and western blotting were performed using a Bio-Rad Mini-Protean III electrophoresis apparatus. Fluorescence and chemical luminescent signals were detected with Imagequant LAS 4000 (GE Healthcare).

In vitro **labeling of eDHFR.** *E. coli* DHFR (eDHFR) was prepared and purified as previously reported.[18] As briefly, thioredoxin- and His-tag fused eDHFR (Trx-His6-eDHFR) was purified with TALON metal affinity resin, and the purified proteins were subjected to site-specific cleavage with thrombin, and then purified with TALON metal affinity resin and benzamidine sepharose to remove Trx-His6 and thrombin. The concentration of eDHFR was determined by absorbance at 280 nm using a molar extinction coefficient of 31,100 $M^{-1}cm^{-1}$. The solution of eDHFR (10 μM) was incubated with labeling reagent (20 μM) in the absence or presence of MTX (100 μM) in 50 mM HEPES buffer (pH 7.2) and incubated at 37°C. The aliquots at different time points were taken, mixed with MTX (100 μM), and subjected to SDS-PAGE analyses. Fluorescence gel images were aquired using a Bio-Rad ChemiDoc XRS system with a 480BP80 filter, and analyzed with Quantity One 1-D Analysis Software (Bio-Rad Laboratories). In separate experiments using MALDI-TOF MS analyses (matrix, sinapinic acid), the labeling yield was estimated by determining the relative MS peak intensity of labeled eDHFR to parental eDHFR. The aliquots at different time points were quenched by using Zip-tip purification with 0.1% TFA aq., and spotted onto a MALDI plate with the matrix solution.

In cell labeling of eDHFR. Before labeling, HeLa-DG cells (6 × 10^5 cells) were seeded on 35-mm dish, and incubated in a DMEM (10% fetal bovine serum (FBS)) for 24 h at 37°C, and washed three times with phosphate buffered saline (PBS). The cells were treated with LDBB reagent (100 or 500 nM) at 37°C in DMEM (FBS free, 1 mL) and incubated. As a control experiment, the labeling was conducted in the presence of TMP (10 μM, pre-incubated for 5 min). After labeling, the cells were washed three times with PBS, and then RIPA buffer (pH 7.4, 25 mM Tris-HCl, 150 mM NaCl, 0.1% SDS, 1% Nonidet P-40, 0.25% deoxycholic acid) was added containing 1% protease inhibitor cocktail set III (Calbiochem®) and 10 μM MTX. The lysed sample was collected and centrifuged (10,000 ×g, 10 min at 4°C). The supernatant was mixed with the same volume of 2× sampling buffer (pH 6.8, 125 mM Tris-HCl, 20% glycerol, 4% SDS, 0.01% bromophenol blue, 8% 2-mercaptoethanol) and incubated for 5 min at 95°C. The samples were subjected to SDS-PAGE (12.5%) and in-gel fluorescence images were obtained. In the live cell imaging experiments, the cells (2 × 10^5 cells, pre-cultured for 24 h in a 35-mm glass bottomed dish) were washed three times with

PBS, and then treated with LDBB reagent **6** (50 nM) at 37°C in DMEM (FBS free, 1 mL) and incubated for 8 h. After removing the medium, the cells were treated with DMEM (10% FBS) for 3 h, and then subjected to imaging analysis using a CLSM (Olympus, FLUOVIEW FV10i). Images of Dc were obtained at an excitation of 405 nm by using a 420-480-nm emission filter, those of GFP were obtained at an excitation of 473 nm by using a 490-520-nm emission filter. For the TMR-type reagent **8**, the HeLa-DG cells were labeled with **8** (20 nM) at 37°C in DMEM (FBS free, 1 mL) and incubated for 3 h. After removing the medium, the cells were treated with DMEM (10% FBS) for 5 h, and then subjected to imaging analysis using a CLSM (FV10i). Images of GFP were obtained at an excitation of 473 nm by using a 490-520-nm emission filter and those of TMR were obtained at an excitation of 558 nm by using a 570-670-nm emission filter, respectively. For the AcFl-type reagent **9**, the normal HeLa cells (2 × 10^5 cells) were cultured in DMEM (10% FBS) in a 35-mm glass-bottomed dish at 37°C for 24 h. The cells were transiently transfected with peDHFR-ECFP plasmid using FuGENE HD (Promega). After 24 h of transfection, cells were labeled with **9** (50 nM) at 37°C in DMEM (FBS free, 1 mL) and incubated for 3 h. After removing the medium, the cells were treated with DMEM (10% FBS) for 5 h, and then subjected to imaging analysis using a CLSM (Olympus, FLUOVIEW FV1000). Images of CFP was obtained at an excitation of 457 nm by using a 470-490-nm emission filter and those of fluorescein (Fl) were obtained at an excitation of 515 nm by using a 520-620-nm emission filter.

In vitro **labeling of hCAI and II.** Human carbonic anhydrase I and II (hCAI, II) were purchased from SIGMA-Aldrich, and used without further purification. The concentration of hCAI and II were determined by absorbance at 280 nm using a molar extinction coefficient of 49,000 $M^{-1}cm^{-1}$ for hCAI and 54,000 $M^{-1}cm^{-1}$ for hCAII, respectively in 50 mM HEPES buffer (pH 7.4, 100 mM NaCl). The solution of hCAI or II (20 μM) was incubated with labeling reagent **11** (40 μM) at 37°C. After 10 h, the labeling reaction was confirmed with MALDI-TOF MS analyses (matrix, sinapinic acid), and the labeling yield was determined to be 46% for hCAI and 38% for hCAII, respectively. After labeling, the protein was purified with gel filtration (TOYO PEARL). To this solution was added urea (at a final concentration of 2 M) and lysyl endopeptidase (LEP) (LEP/substrate ratio = 1/30 (w/w)). After incubation at 37°C over night, the digested samples were applied to RP-HPLC. Collected fractions were analysed by MALDI-TOF MS using CHCA as the matrix and the labeled fragment was further characterized by MALDI TOF-TOF MS/MS (Bruker Daltonics) or MALDI-LTQ-Orbitrap-MS/MS (Thermo Scientific, CA, USA).

In cell Labeling of hCAII and XII. Before labeling, MCF7 cells (2 × 10^5 cells) were seeded on a 35-mm dish, and incubated in DMEM (10% fetal bovine serum (FBS)) for 24 h at 37°C, and then washed three times with DMEM (FBS free). The cells were treated with LDBB reagent **10** (100 nM) at 37°C in a DMEM (FBS free, 1 mL) and incubated. As a control experiment, the labeling was conducted in the presence of EZA (100 μM, pre-incubated for 5 min). After labeling, the cells were washed three times with PBS, and then RIPA buffer was added containing 1% protease inhibitor cocktail set III (Calbiochem) and 100 μM EZA.

The lysed sample was collected and centrifuged (10,000 ×g, 10 min at 4°C). The supernatant was mixed with the same volume of 2× sampling buffer (pH 6.8, 125 mM Tris-HCl, 20% glycerol, 4% SDS, 0.01% bromophenol blue, 100 mM DTT) and vortexed for 1 h at room temperature. The samples were resolved by 12.5% SDS-PAGE and electrotransferred onto an Immun-Blot PVDF membrane (Bio-Rad). The labeled products were detected with anti-Fluorescein antibody (Abcam, × 3000) and anti-rabbit IgG antibody-HRP conjugate (GE Healthcare, × 5000). The immunodetection of hCAII used an anti-hCAII antibody (Abcam, × 2000) and anti-rabbit IgG antibody-HRP conjugate (GE Healthcare, ×5000). The immunodetection of hCAXII used an anti-hCAXII antibody (Cellsignaling, × 2000) and anti-rabbit IgG antibody-HRP conjugate (GE Healthcare, × 5000). The HRP signal was detected with a LAS 4000 imaging system (FujiFilm) using ECL plus western blotting detection reagents (GE Healthcare). In the live cell imaging experiments, the MCF7 cells (2×10^5 cells, pre-cultured for 24 h in a 35-mm glass-bottomed dish) were washed three times with DMEM (FBS free), and then treated with LDBB reagent **10** (100 nM) at 37°C in a DMEM (FBS free, 1 mL) and incubated for 1 h. After removing the medium, the cells were washed three times with DMEM (10% FBS), and then subjected to imaging analysis using a CLSM (Olympus, FLUOVIEW FV10i).

Acknowledgement

Notes and references

[a] Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering Kyoto University, Katsura, Kyoto 615-8510, Japan. Fax: +81-22-795-6557; Tel: +81-22-795-6557; E-mail: ihamachi@sbchem.kyoto-u.ac.jp
[b] Present address: Department of Chemistry, Graduate School of Science, Tohoku University, Aramaki-aza Aoba 6-3, Aoba-ku, Sendai 980-8578, Japan.
[c] Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency, 5 Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan.
† Electronic Supplementary Information (ESI) available: [Figures for eDHFR labeling *in vitro* and in cells, HPLC analyses of the stabilities of the reagents, CPK models of the reagents, and synthetic methods]. See DOI: 10.1039/b000000x/

1 (a) E. M. Sletten and C. R. Bertozzi, *Angew. Chem. Int. Ed.* 2009, **48**, 6974-6998. (b) N. Stephanopoulos and M. B. Francis, *Nat. Chem. Biol.* 2011, **7**, 876-884. (c) H. Wang, E. Nakata and I. Hamachi, *ChemBioChem*, 2009, **10**, 2560-2577. (d) Y. Takaoka, A. Ojida and I. Hamachi, *Angew. Chem. Int. Ed.* 2013, **52**, 4088-4106.
2 (a) D. Leung, C. Hardouin, D. L. Boger and B. F. Cravatt, *Nat. Biotechnol.* 2003, **21**, 687-691. (b) A. E. Speers and B. F. Cravatt, *Chem. Biol.* 2004, **11**, 535-546. (c) M. J. Evans and B. F. Cravatt, *Chem. Rev.* 2006, **106**, 3279-3301. (d) A. Adivekian, B. R. Martin, C. Wang, K.-L. Hsu and D. A. Bachovchin, S. Niessen, H. Hoover and B. F. Cravatt, *Nat. Chem. Biol.* 2011, **7**, 469-478.
3 (a) D. Greenbaum, K. F. Medzihradszky, A. Burlingame and M. Bogyo, *Chem. Biol.* 2000, **7**, 569-581. (b) G. Blum, S. R. Mullins, K. Keren, M. Fonovic, C. Jedeszko, M. J. Rice, B. F. Sloane and M. Bogyo, *Nat. Chem. Biol.* 2005, **1**, 203-209. (c) A. M. Sadaghiani, S. H. L. Verhelst and M. Bogyo, *Curr. Opin. Chem. Biol.* 2007, **11**, 20-28. (d) L. E. Edgington, A. B. Berger, G. Blum, V. E. Albrow, M. G. Paulick, N. Lineberry and M. Bogyo, *Nat. Med.* 2009, **15**, 967-973.
4 (a) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura and I. Hamachi, *Nat. Chem. Biol.* 2009, **5**, 341-343. (b) S. Tsukiji, H. Wang, M. Miyagawa, T. Tamura, Y. Takaoka, I. Hamachi, *J. Am. Chem. Soc.* 2009, **131**, 9046-9054. (c) T. Tamura, S. Tsukiji, I. Hamachi, *J. Am. Chem. Soc.* 2012, **134**, 2216-2226. (d) Y. Takaoka, Y. Kioi, A. Morito, J. Otani, K. Arita, E. Ashihara, M. Ariyoshi, H. Tochio, M. Shirakawa and I. Hamachi, *Chem. Commun.* 2013, **49**, 2801-2803. (e) T. Tamura, Y. Kioi, T. Miki, S. Tsukiji and I. Hamachi, *J. Am. Chem. Soc.* 2013, **135**, 6782-6785.
5 (a) S. Fujishima, R. Yasui, T. Miki, A. Ojida and I. Hamachi, *J. Am. Chem. Soc.* 2012, **134**, 3961-3964. (b) K. Matsuo, Y. Kioi, R. Yasui, Y. Takaoka, T. Miki, S. Fujishima and I. Hamachi, *Chem. Sci.* 2013, **4**, 2573-2580. (c) T. Miki, S. Fujishima, K. Komatsu, K. Kuwata, S. Kiyonaka and I. Hamachi, *Chem. Biol.* 2014, **21**, 1013-1022.
6 C. C. Hughes, Y.-L. Yang, W.-T. Liu, P. C. Dorrestein, J. J. L. Clair and W. Fenical, *J. Am. Chem. Soc.* 2009, **131**, 12094-12096.
7 Recently, enzyme- or peptide-tag fused technology have been reported as powerful tools for intracellular protein imaging by chemical labeling, however, these cannot be applied to endogenously expressed natural proteins: (a) B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269-271. (b) G. Gaietta, T. J. Deerinck, S. R. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien and M. H. Ellisman, *Science*, 2002, **269**, 503-507. (c) A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat. Biotechnol.* 2003, **21**, 86-89. (d) A. Gautier, A. Juillerat, C. Heinis, I. R. Corrêa, Jr., M. Kindermann, F. Beaufils and K. Johnsson, *Chem. Biol.* 2008, **15**, 128-136. (e) G. V. Los *et al.*, *ACS Chem. Biol.* 2008, **3**, 373-382. (f) C. Uttamapinant, K. A. White, H. Baruah, S. Thompson, M. Fernández-Suárez, S. Puthenveetil and A. Y. Ting, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 10914-10919. (g) S. Watanabe, S. Mizukami, Y. Akimoto, Y. Hori and K. Kikuchi, *Chem. Eur. J.* 2011, **17**, 8342-8349. (h) Y. Hori, T. Norinobu, M. Sato, K. Arita, M. Shirakawa, K. Kikuchi, *J. Am. Chem. Soc.* 2013, **135**, 12360-12365.
8 (a) G. T. Hermanson, *Bioconjugate Techniques,* 1st ed.; Academic Press: San Diego, CA, 1966. (b) Haugland, R. P. The handbook: a guide to fluorescent probes and labeling technologies, 10th ed.; Invitrogen: Carlsbad, CA, **2005**.
9 (a) B. K. Shin, H. Wang, A. M. Yim, F. L. Naour, F. Brichory, J. H. Jang, R. Zhao, E. Puravs, J. Tra, C. W. Michael, D. E. Misek, and S. M. Hanash, *J. Biol. Chem.* 2003, **278**, 7607-7616. (b) G. Elia, *Proteomics*, 2008, **8**, 4012-4024. (c) J. M. McFarland and M. B. Francis, *J. Am. Chem. Soc.* 2005, **127**, 13490-13491. (d) R. A. Scheck and M. B. Francis, *ACS Chem. Biol.* 2007, **2**, 247-251. (e) K. Tanaka, T. Masuyama, K. Hasegawa, T. Tahara, H. Mizuma, Y. Wada, Y. Watanabe, K. Fukase, *Angew. Chem. Int. Ed.* 2008, **47**, 102-105.
10 M. Ishida, H. Watanabe, K. Takigawa, Y. Kurishita, C. Oki, A. Nakamura, I. Hamachi and S. Tsukiji, *J. Am. Chem. Soc.* 2013, **135**, 12684-12689.
11 (a) S. S. Gallagher, J. E. Sable, M. P. Sheetz and V. W. Cornish, *ACS Chem. Biol.* 2009, **4**, 547-556. (b) Z. Chen, C. Jing, S. S. Gallagher, M. P. Sheetz and V. W. Cornish, *J. Am. Chem. Soc.* 2012, **134**, 13692-13699. (c) C. Jing and V. W. Cornish, *ACS Chem. Biol.* 2013, **8**, 1704-1712. (d) W. Liu, F. Li, X. Chen, J. Hou, L. Yi and Y.-W. Wu, *J. Am. Chem. Soc.* 2014, **136**, 4468-4471.
12 D. A. Mattews, R. A. Alden, J. T. Bolin, S. T. Freer, R. Hamlin, N. Xuong, J, Kraut, M. Poe, M. Williams and K. Hoogsteen, *Science* 1977, **197**, 452-455.
13 **Figure S7** showed that the ligand-recognition-driven selective and efficient labeling was also observed by using purified eDHFR and TMR- or fluorescein-appended LDBB reagents **8** or **9-2** (deacetylated derivative of compound **9**), confirmed by MALDI-TOF MS analyses in test tube settings.
14 The SDS-PAGE and in-gel fluorescence analyses revealed that the TMR-type and AcFl-type of LDBB reagents (**8** and **9**) could label eDHFR-GFP specifically in live HeLa-DG cells (**Figure S8**).
15 A. Casini, F. Abbate, A. Scozzafava and C. T. Supuran, *Bioorg. Med. Chem. Lett.* 2003, **13**, 2759-2763.

16 C. T. Supuran, *Nat. Rev. Drug Discov.* 2008, **7**, 168-181.

17 On the basis of the antigen-specificity, anti-fluorescein antibody can react only the fluorescein structure (not acetyl fluorescein). Our results indicated that hCAXII was labeled either by deacetylated **10** directly, or
5 by **10** and the subsequent deacetylation of the fluorescein moiety. It is well known that acetyl fluorescein was gradually hydrolyzed to generate fluorescein in an aqueous buffer (J. Hofmann and M. Sernetz, *Anal. Biochem.* 1983, **131**, 180-186), but it is hard to declare which or both possibilities from the current results.

10 18 T. Tanaka, N. Kamiya and T. Nagamune, *FEBS Lett.* 2005, **579**, 2092-2096.