nature chemical biology

Ligand-directed tosyl chemistry for protein labeling *in vivo*

Shinya Tsukiji[1], Masayoshi Miyagawa[1], Yousuke Takaoka[1], Tomonori Tamura[1] & Itaru Hamachi[1,2]

Here we describe a method for the site-selective attachment of synthetic molecules into specific 'endogenous' proteins *in vivo* using ligand-directed tosyl (LDT) chemistry. This approach was applied not only for chemically labeling proteins in living cells, tissues and mice but also for constructing a biosensor directly inside cells without genetic engineering. These data establish LDT chemistry as a new tool for the study and manipulation of biological systems.

Introducing nongenetically encoded, synthetic probes into specific cellular proteins is a key component in chemical biology. Recently, various strategies for chemical protein labeling have been developed, most of which are based on peptide or protein tags that are selectively labeled with designed probes in living cells[1,2]. A nonsense suppression technique can now also be used for site-specific incorporation of several types of unnatural amino acids into proteins in mammalian cells[3,4]. Although undoubtedly powerful, these approaches have limitations. In particular, they require genetic modification of proteins of interest and subsequent (over)expression of the proteins, which inevitably perturbs the physiological condition of cells. Additionally, these techniques cannot be applied to cells or tissues that are not amenable to exogenous gene expression. Therefore, the development of a strategy for modifying specific endogenous proteins within their native habitat—that is, living cells and whole organisms—without gene manipulations is of fundamental importance[5,6].

Toward this end, we focused on a ligand-directed (affinity-based) protein labeling method that uses synthetic compounds consisting of a protein ligand and a reactive group[7–15]. In this approach, the ligand binds to the target protein with high affinity, driving a chemical reaction of the reactive group with an amino acid located on the protein surface through the proximity effect. Affinity labeling can thus satisfy the requirements of target selectivity and site specificity. A crucial disadvantage, however, is that the labeled product loses its function owing to the covalent attachment of the ligand, which permanently occupies the active site[9–13]. To overcome this problem, we designed new reagents that react with target proteins without yielding covalent protein-ligand adducts. We reasoned that, by connecting an affinity ligand and a probe through an electrophilic phenylsulfonate ester group[13–15], the surface of the target protein would be specifically labeled by an S_N2-type reaction with the concomitant release of the ligand molecule: so-called tosyl chemistry (**Fig. 1a**). We chose carbonic anhydrase (CA) as the initial target protein to test our strategy. Accordingly, we synthesized LDT reagents containing benzenesulfonamide as the specific ligand[8,10,12], and also synthetic probes, including fluorophore (SA-Dc, **1**), biotin-tag (SA-Bt, **2**) and ^{19}F NMR probe (SA-Fb, **3**) (**Fig. 1b** and **Supplementary Methods** online).

SA-Dc was stable in aqueous buffer solution, and only about 10% of the starting material was hydrolyzed after the incubation at 37 °C for 48 h (data not shown). We first investigated the reactivity of SA-Dc

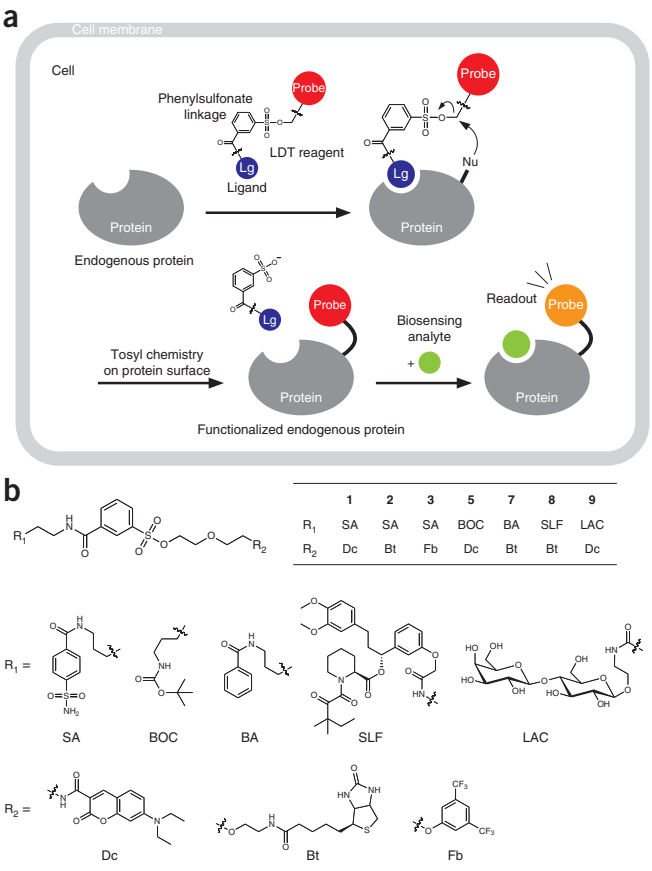

Figure 1 LDT chemistry for labeling endogenous proteins in living native cells. (**a**) Schematic illustration of the strategy. The LDT-based protein labeling method allows for the conversion of an endogenous protein to a semisynthetic functional one—for example, a biosensor as demonstrated in this work. Lg, protein ligand; Nu, nucleophilic amino acid. (**b**) LDT reagents and control compounds used in this study: **1**, **2** and **3** for carbonic anhydrase (CA), **8** for FK506-binding protein 12 (FKBP12), **9** for congerin and **5** and **7** for control experiments.

[1]Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto, Japan. [2]Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency, Sanbancho, Chiyoda-ku, Tokyo, Japan. Correspondence should be addressed to I.H. (ihamachi@sbchem.kyoto-u.ac.jp).

Received 16 December 2008; accepted 26 February 2009; published online 29 March 2009; doi:10.1038/nchembio.157

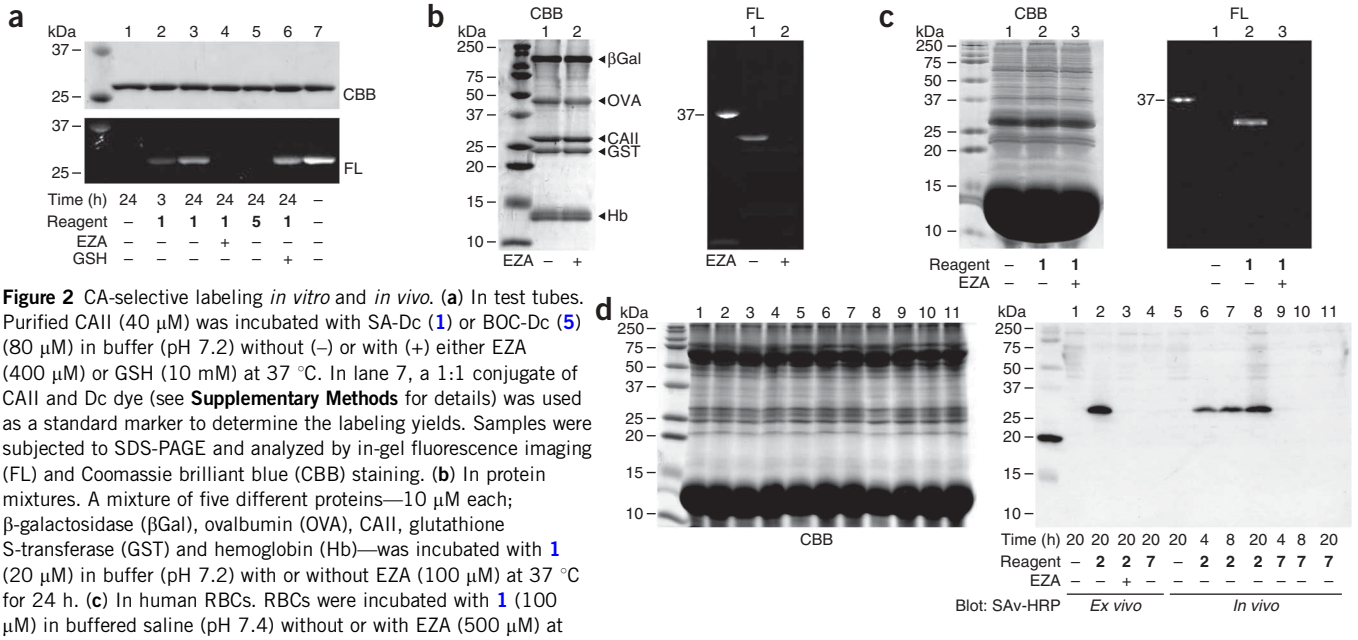

Figure 2 CA-selective labeling *in vitro* and *in vivo*. (**a**) In test tubes. Purified CAII (40 μM) was incubated with SA-Dc (**1**) or BOC-Dc (**5**) (80 μM) in buffer (pH 7.2) without (–) or with (+) either EZA (400 μM) or GSH (10 mM) at 37 °C. In lane 7, a 1:1 conjugate of CAII and Dc dye (see **Supplementary Methods** for details) was used as a standard marker to determine the labeling yields. Samples were subjected to SDS-PAGE and analyzed by in-gel fluorescence imaging (FL) and Coomassie brilliant blue (CBB) staining. (**b**) In protein mixtures. A mixture of five different proteins—10 μM each; β-galactosidase (βGal), ovalbumin (OVA), CAII, glutathione S-transferase (GST) and hemoglobin (Hb)—was incubated with **1** (20 μM) in buffer (pH 7.2) with or without EZA (100 μM) at 37 °C for 24 h. (**c**) In human RBCs. RBCs were incubated with **1** (100 μM) in buffered saline (pH 7.4) without or with EZA (500 μM) at 25 °C for 48 h. (**d**) In living mice and *ex vivo*. For *ex vivo* labeling, blood taken from a mouse was incubated with SA-Bt (**2**) or BA-Bt (**7**) (5 μM) in buffered saline (pH 7.4) with or without EZA (500 μM) at 35 °C for 20 h. For *in vivo* labeling, mice were intravenously injected with **2** or **7** (100 μM in 0.5 ml of buffered saline). Blood was taken from the tail vein and analyzed by western blot using SAv-HRP (right) and anti-mouse CAII antibody (**Supplementary Fig. 3**).

with purified human CA II (CAII) in test tubes. CAII was covalently labeled with 7-diethylaminocoumarin (Dc) dye, as shown by SDS-PAGE and fluorescence gel imaging (**Fig. 2a**). The labeling yield was determined to be approximately 75% after 48 h. To examine the possibility of nonspecific labeling, we incubated CAII with SA-Dc in the presence of ethoxzolamide (EZA, **4**)[16], a strong inhibitor, or with BOC-Dc (**5**; **Fig. 1b**) lacking the sulfonamide ligand. In neither case did the labeling occur. We observed no noticeable decrease in the labeling yield in the presence of a high concentration (10 mM) of reduced glutathione (GSH, **6**). This demonstrates the sufficient stability of the LDT reagent even under a thiol-rich environment. Furthermore, SA-Dc reacted only with CAII in a protein mixture containing five different proteins (**Fig. 2b**), which indicates that SA-Dc is essentially incapable of random protein modification under these conditions, whereas the labeling reaction is facilitated by the proximity effect. We also confirmed the site specificity of this modification by subjecting the labeled CAII to proteolytic digestion. Conventional peptide mapping analysis revealed that His3, located near the N terminus, was predominantly modified by the proposed S_N2-type reaction (**Supplementary Fig. 1** online). In accordance with this labeling scheme, after removing small reagents, the labeled CAII displayed the same catalytic activity as native CAII (**Supplementary Fig. 2** and **Supplementary Table 1** online). These results verify that LDT chemistry can specifically label the target protein without covalently (irreversibly) masking the active site.

We next attempted to modify endogenously expressed CA (eCA) in intact cells. Human red blood cells (RBCs), which express CAI and CAII (ref. 17), were incubated with SA-Dc for 48 h. No hemolysis occurred during this period. The cells were washed, lysed and analyzed by SDS-PAGE and fluorescence gel imaging (**Fig. 2c**). Notably, only a single fluorescent band corresponding to labeled eCA was detected, despite the presence of a large amount of hemoglobin. The addition of EZA completely blocked the labeling, which indicates that this modification occurs via a ligand-eCA interaction inside cells. This intracellular labeling also revealed the cell-membrane permeability of SA-Dc. Owing to high target specificity, the LDT chemistry also worked for modifying eCA in blood cells inside living animals. After SA-Bt carrying a biotin tag was intravenously injected into Slc:ICR mice, blood was taken from the tail vein and analyzed by western blot. A single band was detected by both streptavidin–horseradish peroxidase conjugate (SAv-HRP) and anti-mouse CAII antibody, which indicates that eCA-specific biotin labeling proceeded *in vivo* (**Fig. 2d** and **Supplementary Fig. 3** online). No labeling was detected when BA-Bt (**7**; **Fig. 1b**), which has no sulfonamide ligand, was used. A considerable amount of biotin-modified eCA was produced even within 4 h. No adverse physiological effects on mice were observed upon exposure to the LDT reagents. These results demonstrate the exquisite orthogonality and biocompatibility of LDT-based protein labeling.

This technique was also applicable to other endogenous proteins by altering the affinity ligand. When SLF-Bt (**8**; **Fig. 1b**) containing a synthetic analog of FK506 (ref. 18) was used, endogenous FKBP12 in T lymphocyte Jurkat cells was labeled by biotin in a highly specific manner with no cytotoxicity (**Supplementary Fig. 4** online). By incubating mucus tissue derived from conger eel skin with LAC-Dc (**9**; **Fig. 1b**) bearing a lactose ligand[19], endogenously expressed congerin (an animal lectin) was selectively fluorescently labeled (**Supplementary Fig. 5** online).

The chief advantage of this method is the ability to attach various probes at the vicinity of the active site without compromising protein function. This feature allowed us to directly convert an endogenous protein into a semisynthetic biosensor inside cells. As proof of principle, here we constructed a ^{19}F NMR–based biosensor in cells (for general information on the advantages of ^{19}F NMR, see refs. 20,21). SA-Fb was designed for installing the ^{19}F probe, bis(trifluoromethyl)benzene derivative (Fb), to eCA. After adding SA-Fb to a human RBC suspension, we first monitored the time-dependent change of the ^{19}F signal by in-cell NMR. With the benefit of no background signal, only the extrinsic ^{19}F signals were clearly observed even in the cellular environment. An original peak corresponding to SA-Fb at –62.3 p.p.m. gradually disappeared, and a new single peak appeared at

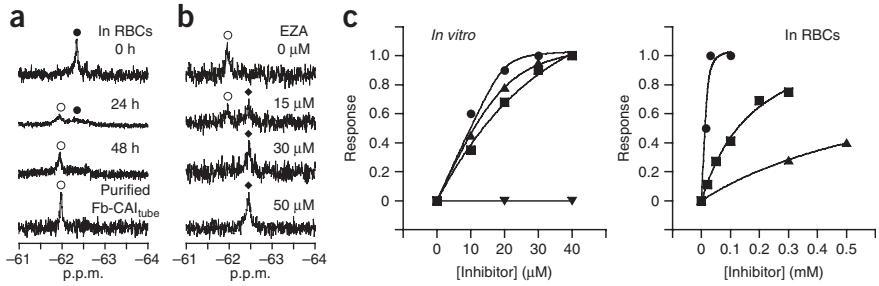

Figure 3 Functional conversion of endogenous CA into a ^{19}F NMR–based biosensor in RBCs. (**a**) Monitoring of ^{19}F labeling processes by in-cell NMR. RBCs were incubated with SA-Fb (100 μM) in buffered saline (pH 7.4) at 25 °C. The cells were washed, resuspended in HEPES-buffered saline containing 20% v/v D_2O and 167 μM trifluoroacetic acid (internal standard at –75.6 p.p.m.) (HBS^{DT}), and subjected to ^{19}F NMR analysis. Fb-labeled CAI (Fb-CAI$_{tube}$) was prepared with purified human CAI in a test tube (see **Supplementary Methods** for details) and used for the ^{19}F NMR measurement (20 μM). Peak ●, –62.3 p.p.m.; peak ○, –62.0 p.p.m. Note that all the NMR spectra in each panel are displayed on the same intensity scale. (**b**) In-cell detection of CA inhibitors by ^{19}F NMR. After installing the ^{19}F probe into CA as described in **a**, the cells were washed, resuspended in HBS^{DT} containing various concentrations of inhibitors and analyzed by ^{19}F NMR. Peak ◆, –62.5 p.p.m. (**c**) Quantitative NMR detection of CA inhibitors with Fb-labeled CAI. The response on the *y* axis is defined as the relative area of peak ◆ divided by the sum of the relative areas of peak ○ and ◆ in **b** and was plotted against the concentration of inhibitors. ●, EZA; ▲, SBA; ■, BS; ▼, TA. Left, *in vitro* system with Fb-CAI$_{tube}$ (20 μM) (**Supplementary Fig. 7**); right, in-cell system.

–62.0 p.p.m. over 48 h (**Fig. 3a**). The chemical shift of the new peak was identical to that of Fb-labeled CAI independently prepared in a test tube (Fb-CAI$_{tube}$) (**Fig. 3a**). These data indicate that ^{19}F labeling proceeds smoothly inside RBCs, and that this process can be monitored by in-cell NMR in real time. Peptide mapping analysis of eCA roughly purified from SA-Fb–treated RBCs indicated that a single Fb probe was attached specifically at His67 of eCAI, which is located at the entrance of the active site (150 μM CAI is expressed more predominantly than 20 μM CAII in human RBCs[17]) (**Supplementary Fig. 6** online). After intracellular labeling, the cells were washed and incubated with several CA inhibitors including EZA, 4-sulfamoylbenzoic acid (SBA, **10**) and benzenesulfonamide (BS, **11**), whose K_i values for human CAI were reported to be 25 nM, 3.4 μM and 3.3 μM, respectively[16,22]. By increasing the extracellular EZA concentration, a new distinct peak appeared at –62.5 p.p.m. with the concurrent disappearance of the original peak at –62.0 p.p.m. (**Fig. 3b**). This change can be ascribed to the binding of the inhibitor to Fb-labeled eCAI because *in vitro* NMR analysis using purified Fb-CAI$_{tube}$ showed identical behavior (**Supplementary Fig. 7** online). No NMR change was observed toward nonligand *p*-toluenesulfonic acid (TA, **12**) (**Fig. 3c**). These results clearly demonstrate that Fb-labeled eCAI acts as a biosensor for CA ligands with the ^{19}F NMR signal as a readout. Notably, whereas Fb-CAI$_{tube}$ showed indistinguishable concentration dependence with each of the inhibitors, the response profiles in the RBC-based system varied substantially depending on the type of inhibitor (**Fig. 3c**). As an example, the binding ability of SBA to Fb-labeled eCAI was much lower than that of BS in cells, despite the fact that these inhibitors have almost identical K_i values toward CAI *in vitro*. This is reasonably explained by the lower cell permeability of anionic SBA compared to neutral BS. The susceptibility of this cell-based biosensor system toward both the affinity and cell permeability of ligand molecules should prove useful in many applications, including drug discovery.

In this work, we have demonstrated the feasibility of site-specific chemical modification and engineering of endogenous proteins within cells and animals using LDT chemistry. Benefiting from the modular nature of LDT reagents, this strategy is generally applicable to various proteins (not only endogenous proteins but also exogenously expressed proteins) by using appropriate ligands as shown in the labeling of CA, FKBP12 and congerin. Given that the target selectivity is determined by the ligand molecule, and that diverse synthetic probes can be incorporated, it may also be possible to label several distinct proteins with different probes in the same cells. The ability to confer new functions on endogenous proteins in their native settings without genetic manipulations should provide powerful tools for the study and engineering of biological processes. More importantly, the design of new chemical reagents that are applicable to natural protein modification in a cellular and whole-organism context is part of the ongoing challenge of expanding the power of organic chemistry in living systems[5,6,19,23–25].

Note: Supplementary information and chemical compound information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

We thank M. Shirakawa and H. Tochio (Kyoto University) for helpful discussions and assistance with ^{19}F NMR measurements. We are grateful to Y. Mori and all members of his group (Kyoto University), in particular S. Kiyonaka, Y. Uriu and S. Yamamoto, for generous assistance on animal experiments. We thank E. Ashihara (Kyoto University Hospital) for his help in taking blood samples. We thank M. Yamada (Shimadzu Corporation) for measuring tandem mass spectra (MALDI-QIT-TOF MS/MS). We also thank all members of the Hamachi laboratory, particularly H. Nonaka, for discussions and experimental assistance. Y.T. acknowledges the Japan Society for the Promotion of Science Research Fellowships for Young Scientists.

AUTHOR CONTRIBUTIONS

S.T. and I.H. conceived and designed the project. S.T. performed the eCA labeling in mice. M.M. performed the synthesis and the *in vitro* CAII labeling. Y.T. performed the synthesis, the eCA labeling in RBCs, including the ^{19}F NMR measurements, and the congerin labeling. T.T. performed the synthesis and the FKBP12 labeling. The manuscript was written by S.T. and I.H.

Published online at http://www.nature.com/naturechemicalbiology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Marks, K.M. & Nolan, G.P. *Nat. Methods* **3**, 591–596 (2006).
2. Fernández-Suárez, M. & Ting, A.Y. *Nat. Rev. Mol. Cell Biol.* **9**, 929–943 (2008).
3. Liu, W., Brock, A., Chen, S., Chen, S. & Schultz, P.G. *Nat. Methods* **4**, 239–244 (2007).
4. Wang, W. *et al. Nat. Neurosci.* **10**, 1063–1072 (2007).
5. Prescher, J.A. & Bertozzi, C.R. *Nat. Chem. Biol.* **1**, 13–21 (2005).
6. Fortin, D.L. *et al. Nat. Methods* **5**, 331–338 (2008).
7. Wold, F. *Methods Enzymol.* **46**, 3–14 (1977).
8. Chen, G. *et al. J. Am. Chem. Soc.* **125**, 8130–8133 (2003).
9. Hamachi, I., Nagase, T. & Shinkai, S. *J. Am. Chem. Soc.* **122**, 12065–12066 (2000).
10. Takaoka, Y., Tsutsumi, H., Kasagi, N., Nakata, E. & Hamachi, I. *J. Am. Chem. Soc.* **128**, 3273–3280 (2006).
11. Cohen, M.S., Zhang, C., Shokat, K.M. & Taunton, J. *Science* **308**, 1318–1321 (2005).
12. Harvey, J.H. & Trauner, D. *ChemBioChem* **9**, 191–193 (2008).
13. Adam, G.C., Cravatt, B.F. & Sorenson, E.J. *Chem. Biol.* **8**, 81–95 (2001).
14. Adam, G.C., Sorenson, E.J. & Cravatt, B.F. *Nat. Biotechnol.* **20**, 805–809 (2002).
15. Weerapana, E., Simon, G.M. & Cravatt, B.F. *Nat. Chem. Biol.* **4**, 405–407 (2008).
16. Winum, J.-Y. *J. Med. Chem.* **48**, 2121–2125 (2005).
17. Casey, J.R. *et al. J. Med. Chem.* **47**, 2337–2347 (2004).
18. Clackson, T. *et al. Proc. Natl. Acad. Sci. USA* **95**, 10437–10442 (1998).
19. Koshi, Y. *et al. J. Am. Chem. Soc.* **130**, 245–251 (2008).
20. Danielson, M.A. & Falke, J.J. *Annu. Rev. Biophys. Biomol. Struct.* **25**, 163–195 (1996).
21. Yu, J.-X., Kodibagkar, V.D., Cui, W. & Mason, R.P. *Curr. Med. Chem.* **12**, 819–848 (2005).
22. Srivastava, D.K. *et al. J. Am. Chem. Soc.* **129**, 5528–5537 (2007).
23. Saxon, E. & Bertozzi, C.R. *Science* **287**, 2007–2010 (2000).
24. Prescher, J.A., Dube, D.H. & Bertozzi, C.R. *Nature* **430**, 873–877 (2004).
25. Laughlin, S.T., Baskin, J.M., Amacher, S.L. & Bertozzi, C.R. *Science* **320**, 664–667 (2008).