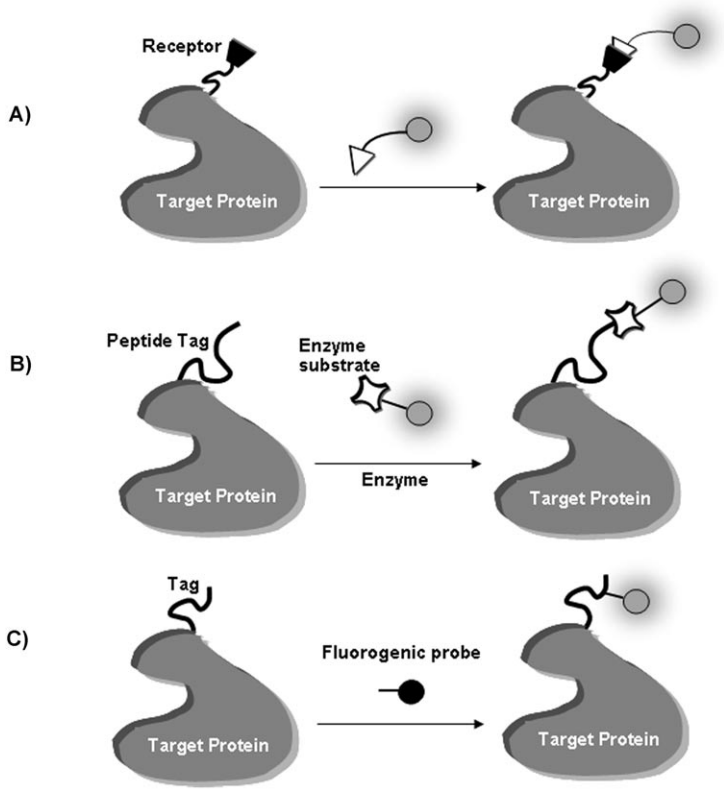
DOI: 10.1002/cbic.200700158

Chemical Labeling of Protein in Living Cells

Anca Dragulescu-Andrasi and Jianghong Rao*[a]

Fluorescence imaging of fluorophore-labeled proteins opens a window into cellular protein biochemistry and enables direct visualization of protein dynamics, localization, and interactions in single living cells. Green fluorescent protein (GFP) and its color variants are popular fluorescent tags because they can be genetically fused to any protein of interest with great specificity.[1,2] However, their relatively large size (~238 amino acids) can potentially alter the structure and/or function of the host proteins, such as mislocalization or misexpression.[3,4] Alternatively, chemoselective labeling of a protein in the context of the native cellular environment with small-molecule fluorescent probes could offer an exciting opportunity to expand the utility of in vivo protein imaging. Several small-molecule-based chemical labeling methods have been developed to address this great challenge.[5]

A commonly used chemical-labeling method takes advantage of the specific interaction between a receptor protein and its small-molecule ligand. Just like GFP, the receptor protein is genetically fused to the protein of interest, and the ligand is synthetically conjugated to a fluorophore (Scheme 1 A). The binding of the ligand to the receptor protein directs the fluorophore label to the fusion protein. In this way, any fluorophore with desirable optical properties can, in principle, be introduced to the fusion protein. An incopmlete list of receptor proteins includes *E. coli* dihydrofolate reductase (DHFR),[6,7] FK506-binding protein (FKBP12),[8] a mutant of human O^6-alkylguanine-DNA transferase (hAGT),[9,10] a mutated prokaryotic dehalogenase (HaloTag™ protein),[11] small dye-binding peptides (in this case, the ligand itself is the fluorescent tag),[12] and lanthanide-binding tags.[13] Depending on the nature of the interaction between the receptor protein and the ligand, the labeling can be covalent (e.g., in the case of hAGT and HaloTag protein) or noncovalent (e.g., DHFR, FKBP12). The fusion protein may be located intracellularly or at the membrane.



Scheme 1. Chemical methods for in vivo protein labeling. A) Receptor-protein-based method. The fluorescent probe is linked to an affinity ligand that specifically binds to its receptor. B) Enzyme-mediated ligation. The probe is linked to a substrate that can be ligated onto a peptide tag genetically fused to the host protein by a "third-party" enzyme. C) A fluorogenic probe binds to a peptide tag–protein fusion, which activates its fluorescence.

This method allows the incorporation of a variety of small-molecule fluorophores with diverse spectral properties into the fusion protein, particularly fluorophores with infrared or near-infrared emission, which are not available with GFP mutants. Tags for other imaging modalities, such as MRI, may be incorporated as well with this method. However, this labeling flexibility comes with several drawbacks. The size of the receptor proteins varies, but is often comparable to that of GFPs. The fluorescent emission of the tag is not modulated by binding, and a washing procedure is generally required to remove unbound ligands and minimize the background. If an endogenous equivalent of the receptor protein

[a] *Dr. A. Dragulescu-Andrasi, Prof. J. Rao*
Molecular Imaging Program at Stanford
Department of Radiology & Bio-X Program
Cancer Biology Program
Stanford University School of Medicine
1201 Welch Road
Stanford, CA 94305-5484 (USA)
Fax: (+1) 650-736-7925
E-mail: jrao@stanford.edu

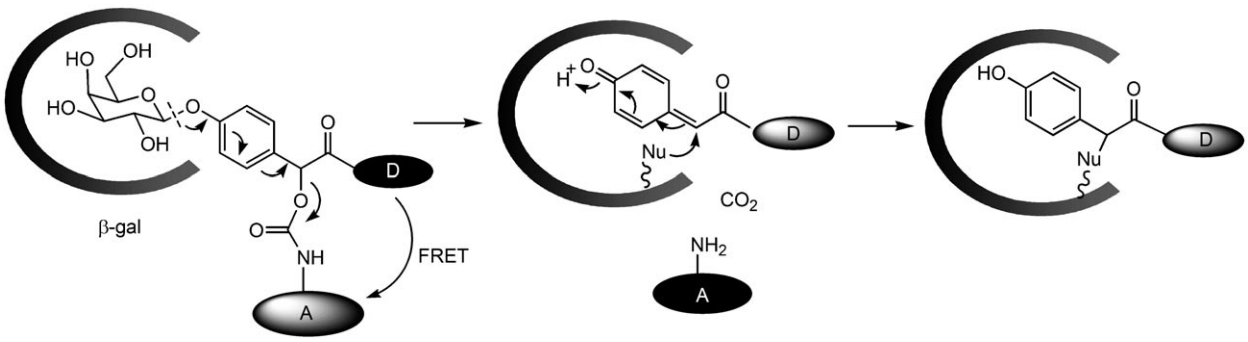
exists, such as for hAGT, mislabeling might occur and lead to a high background. When the labeling is noncovalent, there is also a concern that the tag could "fall off" the fusion protein due to the dissociation of the ligand.

An alternative chemical-labeling method utilizes the substrate of an enzyme instead of a receptor protein as an accepting group for the tag (Scheme 1B). For example, a 15-amino-acid peptide substrate of the *E. coli* biotin ligase was fused to a target protein, thus allowing biotin ligase to specifically add the biotin tag or its ketone analogue to the lysine on the substrate.[14] Acyl carrier protein (ACP; 77 amino acids) from *E. coli*, a substrate of phosphopantetheine transferase (PPTase), has been shown to be a useful tag in labeling fusion proteins on cell surfaces.[15] In this case, PPTase transfers a 4′-phosphopantetheine-conjugated reporter probe from coenzyme A to a serine residue on ACP. Both examples exploit an external enzyme to transfer the small-molecule tag to an acceptor domain linked to the target protein. Although the size of the tag in this method is small and the enzyme-catalyzed ligation offers high specificity, the relatively complicated labeling procedure currently limits the approach essentially to labeling proteins residing at the outer membrane of the cell. Moreover, as with the first method, the emission of the fluorescent tag does not change upon ligation, so again excess unreacted tags have to be washed out.

A fluorogenic labeling probe—that is, one whose fluorescent properties (emission intensity or spectrum) change upon labeling of the target (Scheme 1C)—would be preferable. The induced fluorescent signal immediately reports when the labeling reaction is proceeding, thus avoiding the washout procedure and any delay in imaging. Successful examples of such a designed probe are the biarsenical compounds—fluorescein arsenical hairpin binder (FlAsH) and resorufin arsenical hairpin binder (ReAsH), developed by Tsien et al.[16,17] Both probes are cell-permeable, essentially nonfluorescent, and bind a short peptide sequence (CCPGCC) with high affinity (K_d ~ a few pM) and a remarkable increase in fluorescence emission. This method offers several nice features including a much smaller tagging group and fluorogenic labeling, but the intrinsic affinity of arsenic compounds for thiol groups presents concerns for specificity and toxicity. Lengthy and complex washouts are necessary to minimize the nonspecific arsenic–monothiol interactions, which compromise the power of fluorogenic labeling.

Now Nagano and colleagues[18] have reported a new fluorogenic probe that exhibits fluorescence spectral change upon labeling of the target enzyme, β-galacotosidase (β-gal), in living cells. The probe (CMFβ-gal) is a substrate of β-gal and is conjugated to fluorophores 7-hydroxycoumarin and fluorescein via a cleverly designed linker that contains a quinone methide-forming molecule and a carbamate leaving group (Scheme 2). Due to the efficient fluorescence resonance energy transfer (FRET) between 7-hydroxycoumarin (energy donor) and fluorescein (energy acceptor), the probe emitted strongly at 515 nm when excited at 400 nm. Hydrolysis by β-gal triggers fragmentation of the linker to release the FRET acceptor fluorescein through the expulsion and subsequent decomposition of the carbamate group and generate reactive quinone methide species. The authors suggested that nucleophilic amino acids on β-gal (probably at or close to the active site) were able to trap quinone methide and form covalent adducts in achieving β-gal labeling, though specific labeling sites have not been determined. The resultant labeled β-gal thus displayed a different fluorescence spectrum—a shift in maximum emission from 515 to 460 nm. This spectral change can report in real time the formation of labeled products in the presence of unlabeled proteins and probes through ratiometric measurement of the emissions at 460 and 515 nm. Therefore, this activity-based fluorogenic probe allows real-time imaging of protein labeling in living cells without the need for washout of unreacted probes.

This work has taken an important step towards developing a real-time fluorogenic probe for in vivo protein labeling; however, there are several issues that warrant further research. First, kinetic measurements suggested that, unlike other fluorescently labeled β-D-galactopyranoside analogues,[19] this probe did not act as a suicide inhibitor but appeared to be a substrate (with $k_{cat} = 6.1\ s^{-1}$ and $K_m = 0.15$ mM). It is thus unclear whether all enzymatically produced quinone methide species are confined to



Scheme 2. In vivo labeling of β-gal by a fluorogenic probe with spectral change. The labeling is proposed to take place in two steps. The first step involves *O*-galactoside bond cleavage, which generates an active intermediate quinone methide. This intermediate is susceptible to nucleophilic attack by a nearby amino acid residue, which leads to covalent attachment of the FRET donor (D) to the enzyme and displacement of the acceptor (A).

β-gal or just some. The detailed labeling mechanism requires further investigation and the possibility of cross-labeling of other cellular proteins needs careful evaluation. Secondly, the large size of β-gal (a tetrameric enzyme with a molecular weight of 540 kDa) makes it unsuitable as a fusion tag for other proteins of interest. It remains to be seen how far this strategy can be extended to other smaller enzyme targets, because it requires nucleophilic amino acids on the protein targets to be appropriately located for labeling. In principle, genetic engineering would be able to introduce such nucleophilic amino acids if they are absent in the native structure. Thirdly, the probe has to be microinjected into the cell due to its high polarity; this severely hampers its in vivo applicability. This limitation might be overcome by incorporating a small membrane translocation unit that could enable efficient delivery in living cells.[20]

Keywords: fluorescent probes · in vivo labeling · proteins · ratiometric imaging

[1] R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509 – 544.
[2] A. Chiesa, E. Rapizzi , V. Tosello, P. Pinton, M. E. De Virgilio, K. Fogarty , R. Rizzuto, *Biochem. J.* **2001**, *355*, 1 – 12.
[3] G.-W. Tian, A. Mohanty, S. N. Chary, S. Li, B. Paap, G. Drakakaki, C. D. Kopec, J. Li, D. Ehrhardt, D. Jackson, S. Y. Rhee, N. V. Raikhel, V. Citovsky, *Plant Physiol.* **2004**, *135*, 25 – 38.
[4] M. Andersen, R. Schmitz-Salue, S. Jakobs, *Mol. Biol. Cell* **2004**, *15*, 5616 – 5622.
[5] I. Chen, A. Y. Ting, *Curr. Opin. Biotechnol.* **2005**, *16*, 35 – 40.
[6] L. W. Miller, J. Sable, P. Goelet, M. P. Sheetz, V. W. Cornish, *Angew. Chem.* **2004**, *116*, 1704 – 1707; *Angew. Chem. Int. Ed.* **2004**, *43*, 1672 – 1675.
[7] L. W. Miller, Y. Cai, M. P. Sheetz, V. W. Cornish, *Nat. Methods* **2005**, *2*, 255 – 257.
[8] K. M. Marks, M. Rosinov, G. P. Nolan, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9982 – 9987.
[9] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86 – 89.
[10] A. Keppler, H. Pick, C. Arrivoli, H. Vogel, K. Johnsson, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9955 – 9959.
[11] Technical Manual, Promega, "HaloTag Interchangeable Labeling Technology"; http://www.promega.com, **2006**.
[12] K. M. Marks, M. Rosinov, G. P. Nolan, *Chem. Biol.* **2004**, *11*, 347 – 356.
[13] B. R. Sculimbrene, B. Imperiali, *J. Am. Chem. Soc.* **2006**, *128*, 7346 – 7352.
[14] I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, *2*, 99 – 104.
[15] L. Vivero-Pol, N. George, H. Krumm, K. Johnsson, N. Johnsson, *J. Am. Chem. Soc.* **2005**, *127*, 12770 – 12771.
[16] B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269 – 272.
[17] G. Gaietta, T. Deerinck, S. R. Adams, J. Bouwer, O. Tour, D. Laird, G. Sosinsky, R. Y. Tsien, M. Ellisman, *Science* **2002**, *296*, 503 – 507.
[18] T. Komatsu, K. Kikuchi, H. Takakusa, K. Hanaoka, T. Ueno, M. Kamiya, Y. Urano, T. Nagano, *J. Am. Chem. Soc.* **2006**, *128*, 15946 – 15947.
[19] M. Kurogochi, S.-I. Nishimura, Y. C. Lee, *J. Biol. Chem.* **2004**, *279*, 44704 – 44712.
[20] J. S. Wadia, S. F. Dowdy, *Adv. Drug Delivery Rev.* **2005**, *57*, 579 – 596.

Received: March 29, 2007
Published online on May 10, 2007

