COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Cell-permeable small molecule probes for site-specific labeling of proteins†

Dawn S. Y. Yeo,[a] Rajavel Srinivasan,[b] Mahesh Uttamchandani,[a] Grace Y. J. Chen,[ab] Qing Zhu[b] and Shao Q. Yao*[ab]

[a] *Department of Biological Sciences, National University of Singapore, 3 Science Drive 3, Singapore 117543, Republic of Singapore. E-mail: chmyaosq@nus.edu.sg; Fax: 65 6779 1691; Tel: 65 6874 1683*

[b] *Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore 117543, Republic of Singapore. E-mail: chmyaosq@nus.edu.sg; Fax: 65 6779 1691; Tel: 65 6874 1683*

Received (in Cambridge, UK) 5th August 2003, Accepted 14th October 2003
First published as an Advance Article on the web 29th October 2003

We have successfully synthesized a number of small molecule probes designed for site-specific labeling of N-terminal cysteine-containing proteins expressed in live cells. Their utility for site-specific, covalent modifications of proteins was successfully demonstrated with purified proteins *in vitro*, and with live bacterial cells *in vivo*.

Studying the dynamic movement and interactions of proteins inside living cells is critical for a better understanding of cellular mechanisms and functions. Traditionally this has been done by *in vitro* labeling of proteins with fluorescent and other molecular probes, followed by transferring them into a live cell and monitoring them, in real time and using advanced imaging techniques such as confocal microscopy, and others.[1] Recent advances in genetic engineering have made it possible to directly generate fluorescently labeled proteins in living cells, or even live animals, by fusion of fluorescent proteins such as GFP (green fluorescent protein) to the protein of interest.[2] This strategy, although extremely powerful, has several problems. Firstly, the introduction of GFP and other fluorescent proteins to a target protein may affect its biological and cellular activities, due to the relatively large size of the fusion (*i.e.* 27 KDa for GFP). Secondly, there are currently few fluorescent proteins available, thereby limiting the number of colors that can be used to "tag" a protein *in vivo*. Lastly, the strategy is limited to protein labeling with fluorophores but not other molecular probes. In order to address some of these problems, Tsien *et al.* recently described a novel method which allows efficient labeling of proteins *in vivo* using cell-permeable organoarsenic compounds.[3] Johnsson *et al.* described enzyme-catalyzed, *in vivo* labeling of proteins fused to the human DNA repair protein hAGT.[4] We are interested in developing a novel strategy for site-specific covalent labeling of proteins, *in vivo*, by taking advantage of the chemoselective reaction between thioester-containing small molecules and proteins expressing N-terminal cysteines under physiological environments. Previously, N-terminal cysteine proteins had mostly been used in the semi-synthesis of proteins, as well as site-specific labeling of proteins *in vitro*.[5] More recently, the Muir group has successfully demonstrated, in living cells, the protein semi-synthesis between two protein fragments fused to the trans-splicing domains of the *Ssp* DnaE intein.[6]

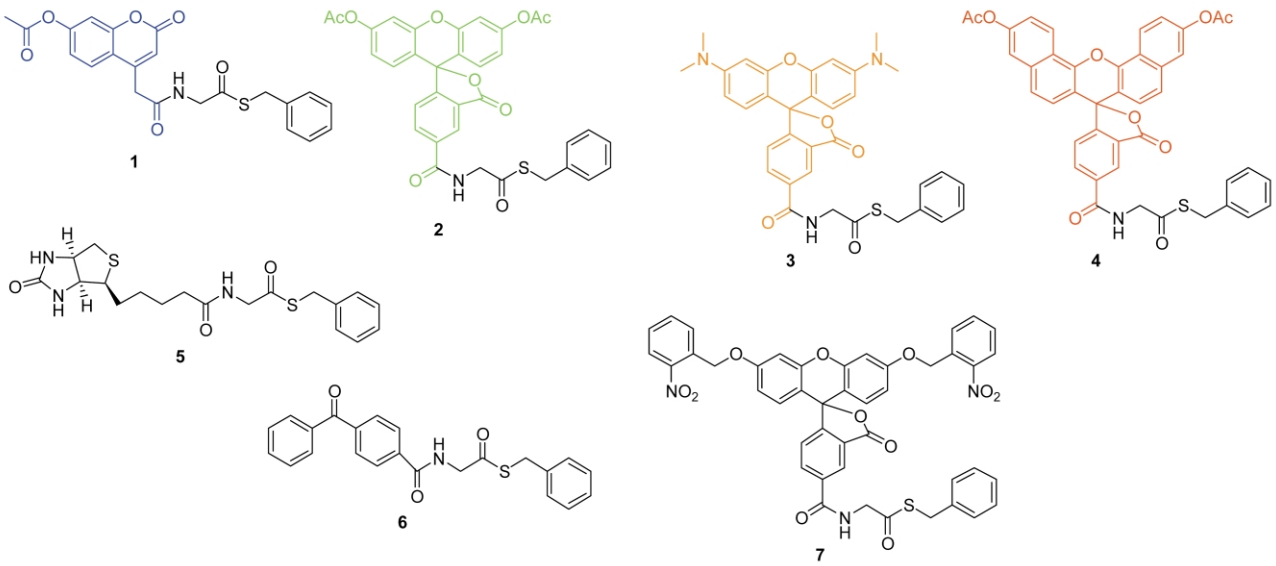
In our strategy (Supplementary Information), a protein of interest having an N-terminal cysteine is expressed inside a live cell, by either intein-mediated protein splicing, or ubiquitin fusion.[7] Incubation of the cell with a thioester-containing, cell-permeable molecule probe allows the probe to efficiently penetrate through the cell membrane into the cell, where the chemoselective reaction occurs predominantly between the thioester and the N-terminal cysteine of the protein. This is true since endogenous N-terminal cysteine proteins are rare. Other endogenous molecules, such as cysteine and cystamine, are present in the cell and will also react with the probe. However, their reaction products are also small molecules in nature, and could be easily removed, together with any excessive unreacted probe, by extensive washing of the cells after labeling. Herein, we report the detailed synthesis of different cell-permeable probes, the *in vitro* labeling studies of selected probes, as well as their utility in the site-specific labeling of proteins expressed inside live bacterial cells.

A total of 7 different probes have been synthesized (Supplementary Information), of which probes **1** to **4** are fluorophore-containing thioesters (Scheme 1). **5** and **6** are biotin- and benzophenone-containing probes, respectively. **7** is a "caged" molecule of **2**, in which the fluorescence is designed to be "turned on" selectively upon photolysis. All probes were designed to be cell-permeable, in that acetates of different fluorophores were incorporated in **1, 2** and **4** to increase their cell permeability. The fluorophore in **3**, tetramethylrhodamine (TMR), as well as the biotin and benzophenone moieties in **5** and **6**, respectively, were previously shown to be cell-permeable.[8] Addition of the hydrophobic, benzyl-based thioester in all probes should further increase their cell permeability. Probes **1** to **4**, containing different fluorophores (*e.g.* coumarin (CM), fluorescein (FL), TMR and carboxynaphthofluorescein (CF), respectively) that emit in different colors (*e.g.* blue, green, orange/yellow and red, respectively; see Fig. 1), were designed for potential multicolor cell labeling and imaging. Proteins labeled with probes **5** and **6** may be used to study protein–protein interactions by *in vivo* experiments utilizing biotin–avidin binding and protein crosslinking, respectively. Probe **7** may be used for protein labeling in a live cell where temporal and/or confined fluorescence activation is needed.[9]

In vitro labeling of proteins expressing an N-terminal cysteine was then carried out with the probes. A model protein, EGFP (enhanced green fluorescent protein), engineered to contain an N-terminal cysteine, was incubated with probes **2, 3, 4** and **5** individually, and the extent of protein labeling was monitored over 24 hours by SDS-PAGE and Western blotting. 8 μM of each probe, with or without 1 mM of DTT, was added to the pure protein in 1 X PBS buffer. The reaction was quenched at specified time intervals with 10 mM of cysteine. Following protein separation on a 12% SDS-PAGE gel, the labeled protein was visualized and quantitated, either by a fluorescence gel scanner (in the cases of probes **2, 3** and **4**) or Western blotting using anti-biotin HRP conjugate and Amersham's ECL kit (in the case of probe **5**). Results are summarized in Fig. 2a and 2b and in the Supplementary Information. In all cases, the labeling was shown to reach near completion (> 75% labeling) within the first 3 hours of the reaction. In addition, more than 50% labeling occurred within the first 30 min of the reaction, making this strategy suitable for potential real-time bioimaging experiments in live cells. The site-specific nature of the labeling reaction was confirmed by repeating the experiment under identical conditions with control proteins which either do not have cysteine residues at all, or have only internal cysteines.

† Electronic supplementary information (ESI) available: experimental details and characterization of compounds. See http://www.rsc.org/suppdata/cc/b3/b309196a/

DOI: 10.1039/b309196a



Scheme 1 Small molecule probes synthesized and used in the studies.

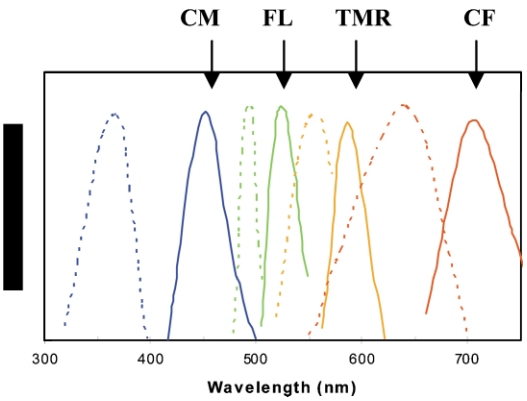


Fig. 1 Fluorescence spectra of probes **1** (blue), **2** (green), **3** (orange) and **4** (red). Excitation and emission spectra are shown by dashed and solid lines, respectively. All spectra were taken with 10 μM of dyes in potassium carbonate buffer. See Supplementary Information for details.

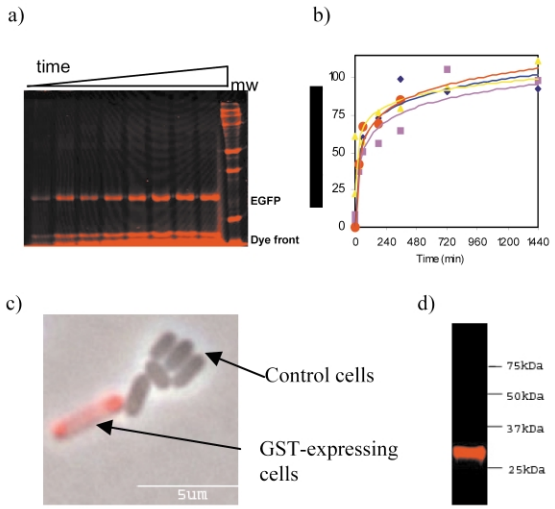


Fig. 2 Site-specific labeling of N-terminal cysteine proteins with the small molecule probes. (a) SDS-PAGE of EGFP labeling with probe **4** for 1 min, 10 min, 30 min, 1 h, 3 h, 6 h, 12 h, 24 h (left to right). For other probes, see Supplementary Information. (b) % completion of EGFP labeling over 24-hour intervals with probes **2** (blue diamonds), **3** (purple squares), **4** (yellow triangles) and **5** (red diamonds). (c) Fluorescence microscope images of *E. coli* expressing N-terminal cysteine GST after labeling with probe **3** for 24 h. The fluorescence image was overlapped with the phase contrast image for easy visualization. (d) 12% SDS-PAGE protein analysis of cells from (c).

In all cases, labeling occurred ONLY with proteins possessing an N-terminal cysteine (Supplementary Information), thereby unambiguously supporting our design principle, in which exclusive labeling should only occur at the N-terminal cysteine of the target protein. We next applied the strategy to the labeling of N-terminal cysteine proteins expressed inside live bacterial cells. *E. coli* cells overexpressing an N-terminal cysteine GST (glutathione-*S*-transferase) protein were treated with probe **3**, and the site-specific labeling of the probe inside the cells was assessed by fluorescence microscopy and SDS-PAGE (Fig. 2c and 2d, respectively): exclusive labeling occurred only with *E. coli* cells overexpressing N-terminal cysteine GST. Minimal background labeling was unambiguously confirmed by SDS-PAGE protein analysis of the labeled cells, indicating that the labeling occurred predominantly with the N-terminal cysteine GST.

In conclusion, we have successfully synthesized a number of cell-permeable, small molecule probes capable of site-specific labeling of N-terminal cysteine-containing proteins expressed in live bacterial cells. The strategy should make feasible future bioimaging experiments where other types of live cells are involved.

Funding support was provided by the National University of Singapore (NUS) and the Agency for Science, Technology and Research (A*STAR) of Singapore.

Notes and references

1 G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, San Diego, CA, 1996.
2 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **57**, 509.
3 B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269.
4 A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel and K. Johnsson, *Nat. Biotechnol.*, 2003, **21**, 86.
5 (*a*) P. E. Dawson and S. B. Kent, *Annu. Rev. Biochem.*, 2000, **69**, 923; (*b*) T. J. Tolbert and C. H. Wong, *Angew. Chem., Int. Ed.*, 2002, **41**, 2171; (*c*) B. Schuler and L. K. Pannell, *Bioconj. Chem.*, 2002, **13**, 1039.
6 I. Giriat and T. W. Muir, *J. Am. Chem. Soc.*, 2003, **125**, 7180.
7 (*a*) M. Q. Xu and T. C. Evans, *Methods*, 2001, **24**, 257; (*b*) R. T. Baker, *Curr. Opin. Biotechnol.*, 1996, **7**, 541.
8 R. P. Haugland, *Handbook of fluorescent probes and research products*; 9th Edn., Molecular Probes: Eugene, OR, 2002.
9 J. P. Schwartz and G. H. Patterson, *Science*, 2003, **300**, 87.