Chemistry & Biology
Article

An Engineered Protein Tag for Multiprotein Labeling in Living Cells

Arnaud Gautier,[1] Alexandre Juillerat,[1,3] Christian Heinis,[1,4] Ivan Reis Corrêa, Jr.,[1] Maik Kindermann,[2,5] Florent Beaufils,[2] and Kai Johnsson[1,*]
[1]Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[2]Covalys Biosciences AG, CH-4108 Witterswil, Switzerland
[3]Present address: Pasteur Institute, Unité d'Immunologie Structurale, 25–28 rue du Docteur Roux, 75724 Paris Cedex 15, France.
[4]Present address: Division of Protein and Nucleic Acid Chemistry, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom.
[5]Present address: Sanofi-Aventis Deutschland, Chemical Sciences, D-65926 Frankfurt am Main, Germany.
*Correspondence: kai.johnsson@epfl.ch
DOI 10.1016/j.chembiol.2008.01.007

SUMMARY

The visualization of complex cellular processes involving multiple proteins requires the use of spectroscopically distinguishable fluorescent reporters. We have previously introduced the SNAP-tag as a general tool for the specific labeling of SNAP-tag fusion proteins in living cells. The SNAP-tag is derived from the human DNA repair protein O^6-alkylguanine-DNA alkyltransferase (AGT) and can be covalently labeled in living cells using O^6-benzylguanine derivatives bearing a chemical probe. Here we report the generation of an AGT-based tag, named CLIP-tag, which reacts specifically with O^2-benzylcytosine derivatives. Because SNAP-tag and CLIP-tag possess orthogonal substrate specificities, SNAP and CLIP fusion proteins can be labeled simultaneously and specifically with different molecular probes in living cells. We furthermore show simultaneous pulse-chase experiments to visualize different generations of two different proteins in one sample.

INTRODUCTION

Fluorescence microscopy is the method of choice for the visualization of protein function and biochemical activity in the living cell. Such experiments often use autofluorescent proteins (AFPs) to selectively tag individual proteins (Giepmans et al., 2006). Biologists now possess a broad range of spectroscopically distinguishable AFPs that can be used for multicolor imaging, and a variety of AFP-based sensors have also been devised to report on key biochemical activities in living cells (Giepmans et al., 2006; Shaner et al., 2005). As a complement to AFPs, various technologies have been developed for specific protein labeling with synthetic probes (Johnsson and Johnsson, 2007; O'Hare et al., 2007). These methods rely on the fusion of the protein of interest to a tag that can be covalently labeled with a small molecule, thereby combining the simplicity of fusion protein expression with the diversity of molecular probes provided by chemistry. Recent innovations in chemical probes include environmentally sensitive fluorophores that can be specifically targeted to subcellular microenvironments to report on ion concentrations (Tour et al., 2007) and selective crosslinkers that can be used to sense protein-protein interactions inside living cells (Lemercier et al., 2007).

An important further development of selective protein labeling in living cells would be the possibility to label two proteins simultaneously with different molecular probes for multiparameter imaging of cellular functions (Schultz et al., 2005). Currently, three tags are used for covalent labeling inside living cells: the tetracysteine tag (Griffin et al., 1998), SNAP-tag (Keppler et al., 2003), and HaloTag (Los et al., 2005). SNAP-tag is derived from the 20 kDa DNA repair protein O^6-alkylguanine-DNA alkyltransferase (AGT) and is labeled using O^6-benzylguanine derivatives. Compared to the tetracysteine tag, SNAP-tag allows highly specific labeling without restrictions on the cellular compartment (Keppler et al., 2004b) and without the need for additional reagents to suppress background. In the case of HaloTag, there are so far too few published examples to fully evaluate the method. In order to extend the number of tags available for multicolor labeling, we aimed to generate an AGT mutant with a new substrate specificity that could be used in combination with SNAP-tag (or other tags) for specific labeling of multiple proteins in one cell. SNAP-tag is normally covalently labeled with O^6-benzylguanine (BG) derivatives bearing a chemical probe by undergoing an irreversible reaction in which the functionalized benzyl group of the BG derivative is transferred to an active site cysteine to form a covalently modified protein (Keppler et al., 2003, 2004a, 2004b). Recently, we reported the generation of an AGT mutant capable of reacting with O^6-propargylguanine (PG) (Heinis et al., 2006). Although PG derivatives do not react with SNAP-tag, the AGT mutant generated in this earlier work displayed a relatively low reactivity with PG and furthermore retained its activity toward BG. This lack of specificity prohibited simultaneous labeling of this mutant plus SNAP-tag in living cells. Here we describe the generation of an AGT-based tag, dubbed CLIP-tag, which allows the simultaneous and specific covalent labeling of two different SNAP and CLIP fusion proteins in living cells (Figure 1A). The selectivity and speed of the labeling and the ease of synthesis of the substrates should make the CLIP-tag an important addition to existing labeling methods.

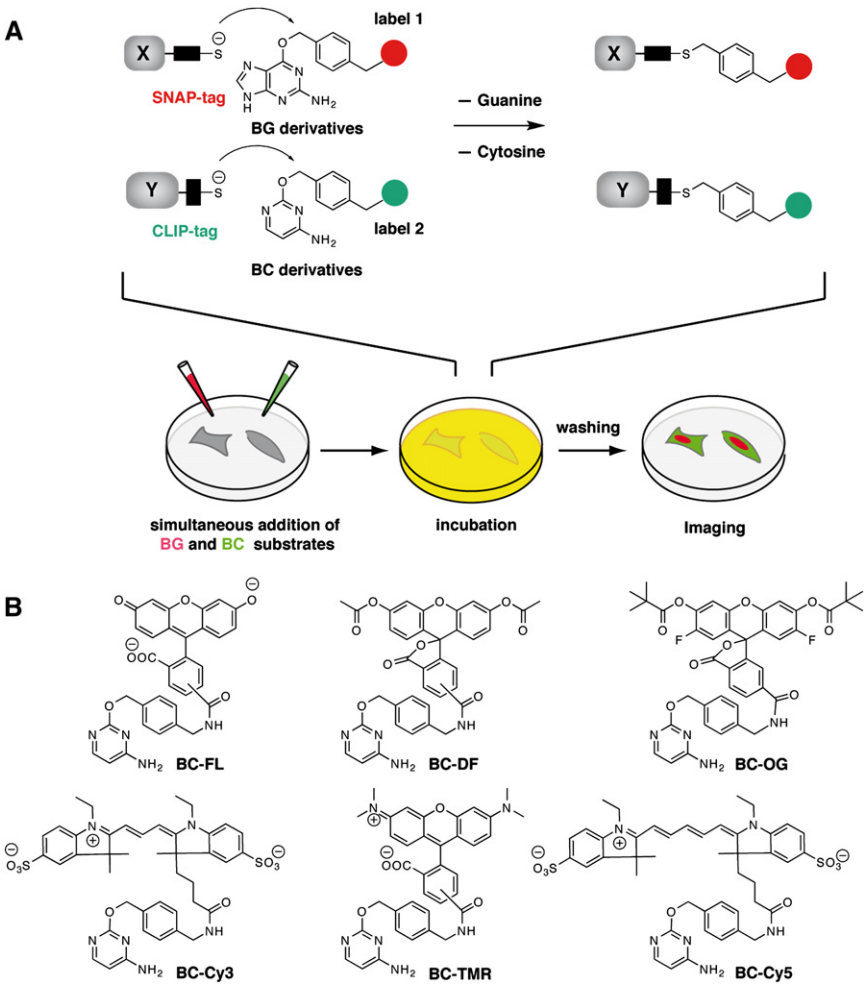

Figure 1. Simultaneous and Specific Labeling of Two Fusion Proteins with Different Molecular Probes

(A) Use of BC derivatives and the AGT-based CLIP-tag (vide infra) together with BG derivatives and SNAP-tag for specific and simultaneous labeling of CLIP and SNAP fusion proteins.

(B) BC derivatives used in this work for labeling with fluorescein (BC-FL), diacetylfluorescein (BC-DF), dipivaloyl Oregon green (BC-OG), tetramethylrhodamine (BC-TMR), Cy3 (BC-Cy3), and Cy5 (BC-Cy5).

RESULTS

Substrate Design, Synthesis, and Properties

The specificity of AGT for alkylguanine derivatives is mainly a result of molecular recognition of the leaving group guanine, so we envisioned the use of substrates with modified leaving groups as potential substrates for a new AGT mutant. Specifically, we focused on O^2-benzylcytosine (BC), in which an alkylated cytosine replaces the alkylated guanine of BG (Figure 1A). We predicted that BG and BC should display similar reactivity in S_N2 reactions because BG and BC differ only in the leaving group and because the pK_As of guanine and cytosine are similar (Fasman, 1975). We also predicted that AGT would not react well with BC, as the specific interactions that facilitate its reaction with BG, notably the hydrogen bonding of Tyr114 to the N^3 of guanine, could not be formed with BC (Daniels et al., 2004). Finally, cytosine possesses a different pattern of hydrogen-bond donors and acceptors and is less bulky than guanine, so that AGT mutants that react efficiently with BC should react poorly with BG.

A straightforward synthesis was developed to generate a series of fluorescent substrates by coupling BC to (1) green fluorescent probes: fluorescein (BC-FL), diacetylfluorescein (BC-DF), and dipivaloyl Oregon green (BC-OG), (2) red fluorescent probes: Cy3 (BC-Cy3) and tetramethylrhodamine (BC-TMR), and (3) far-red fluorescent Cy5 (BC-Cy5) (Figure 1B). The common intermediate in the synthesis of these substrates is available in just two steps from a commercially available precursor (see Figure S1 in the Supplemental Data available with this article online).

First, we characterized the reactivity of BC derivatives toward SNAP-tag by a fluorescence assay using purified SNAP-tag protein. BC-FL was shown to label SNAP-tag with fluorescein about 1000-fold less efficiently than BG-FL, a BG substrate carrying fluorescein: the second-order rate constant of the reaction of SNAP-tag with BC-FL was found to be 26 M^{-1} s^{-1}, whereas the rate constant for the reaction of SNAP-tag with BG-FL was measured to be 2.8 × 10^4 M^{-1} s^{-1} (Table 1).

Next, we assessed the reactivity of BC derivatives with the mammalian proteome by incubating six different cell lines (HEK293T, CHO, BHK, HeLa CCL2, HeLa MZ, and HT29) with BC-DF, followed by in-gel fluorescence analysis (Figure 2). In order to compare the intrinsic reactivity of BC derivatives to those of other existing substrates available for covalent labeling in living cells, we performed parallel experiments with BG-DF and HaloTag-DF, diacetylfluorescein-bearing substrates of SNAP-tag and HaloTag, respectively. No significant labeling of proteins by BC-DF was observed in any of the tested cell lines. Incubating the cell lines with BG-DF did not lead to any significant labeling of proteins in four of the cell lines (CHO, HEK293T, BHK, and HeLa MZ), whereas a protein of 23 kDa, which we assigned to be endogenous AGT, was detected in HT29 and HeLa CCL2 cells at about 0.5–1 pmol per mg soluble protein (pmol/mg). This observation is in agreement with the fact that numerous spontaneous immortalized and virus-transfected cell lines are AGT deficient, whereas the expression level of cell lines that do express AGT has been reported to be around 0.1–1 pmol/mg (Foote and Mitra, 1984; Kaina et al., 2007). Incubation of the six cell lines with HaloTag-DF led to the labeling of an unknown 28 kDa protein labeled at 10–30 pmol/mg. We assume that the relatively high background labeling observed with HaloTag-DF is because of the intrinsic reactivity of primary chlorides toward nucleophiles.

Table 1. Key AGT Mutants Described in This Work

Name	Description	Mutations/SNAP-tag	k_{BCFL} ($M^{-1}s^{-1}$)	k_{BGFL} ($M^{-1}s^{-1}$)	$[Urea]_{1/2}$ (M)
SNAP-tag	Previously described tag with 50-fold enhanced activity toward BG derivatives compared to wild-type AGT (Gronemeyer et al., 2006)	None	26 ± 5	2.8×10^4	6.3 ± 0.1
Mut1	Mutant selected by yeast display from a saturation mutagenesis library based on SNAP-tag	Y114E, K131N, S135D, G157P, E159F	90 ± 15	≤1	4.1 ± 0.2
CLIP-tag	Mutant selected by phage display from a random mutagenesis library based on Mut1	M60I, Y114E, A121V, K131N, S135D, L153S, G157P, E159L	1130 ± 150	≤10	5.1 ± 0.2

For each mutant, mutations relative to parental SNAP-tag, the second-order rate constants of the labeling reactions with BG-FL and BC-FL (k_{BGFL} and k_{BCFL}), and the urea concentrations necessary for 50% inactivation of protein ($[urea]_{1/2}$) are listed. Data represent mean ± SD.

It can thus be concluded that (1) BC-DF and BG-DF show significantly lower background labeling than HaloTag-DF and that (2) BC-DF, in contrast to BG-DF, does not lead to detectable labeling of endogenous AGT. BC derivatives are therefore suitable substrates for the generation of a new self-labeling tag.

Generation of AGT Mutants with Activity toward BC

To generate AGT mutants with altered substrate specificity, we attempted to redesign the active site of SNAP-tag by directed evolution. We generated a library of AGT mutants with random residues at positions 114, 131, 135, 148, 156, 157, and 159. These positions in direct proximity to BG bound in the active site were chosen with the aid of the crystal structure of wild-type AGT (Wibley et al., 2000) (Figure 3) and because of the known role of some of them: Tyr114 forms a hydrogen bond with the N^3 of BG to stabilize the developing negative charge on the leaving group guanine (Daniels et al., 2004) and Glu159 has been proposed to form a hydrogen bond with the N^7 of BG (Juillerat et al., 2003). Selected codons were replaced by the degenerate codon NNK using PCR (Figure S2) and the library was prepared by inserting the randomized gene in plasmid pCTCON for display of the AGT mutants on the yeast cell surface as fusion

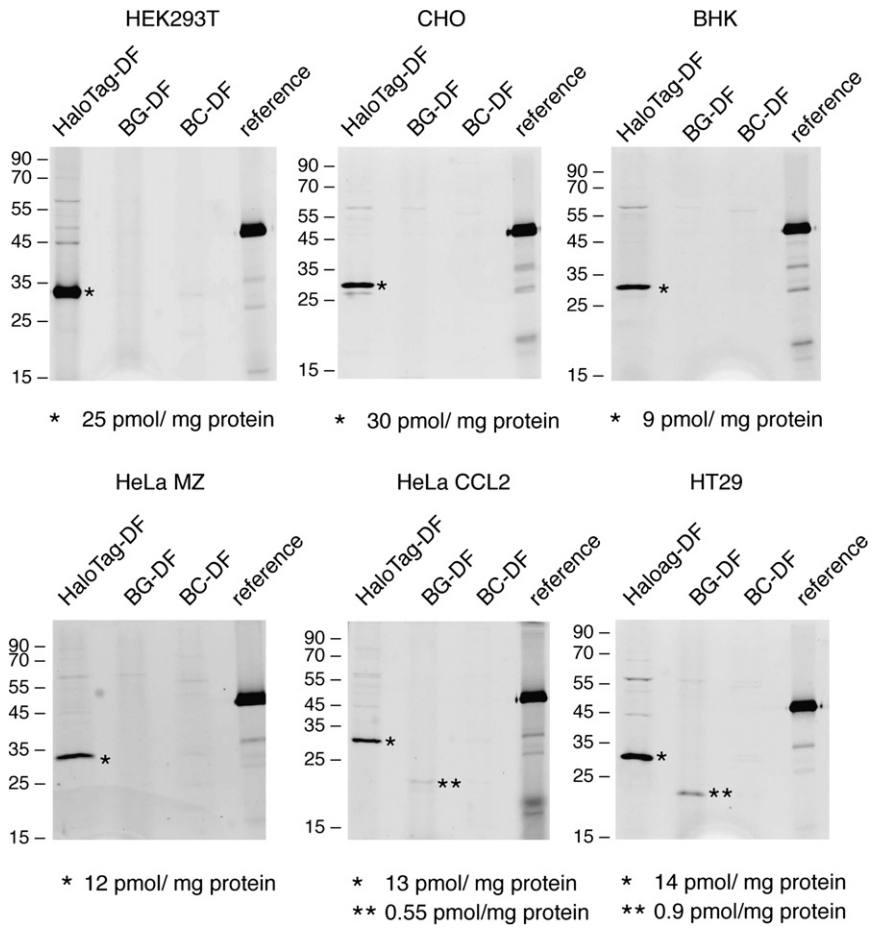

Figure 2. Reactivity of BC-DF, BG-DF, and HaloTag-DF with the Mammalian Proteome

Cells were incubated with 10 μM substrate for 1 hr at 37°C. After cell lysis, equal amounts of protein from crude extracts were analyzed by SDS-PAGE and in-gel fluorescence scanning. The 28 kDa endogenous protein labeled with fluorescein by HaloTag-DF (*) and the endogenous AGT labeled with BG-DF (**) were quantified (in pmol/mg of soluble extract) by comparison with the fluorescence intensity of a known amount of fluorescein-labeled GST-SNAP (reference).

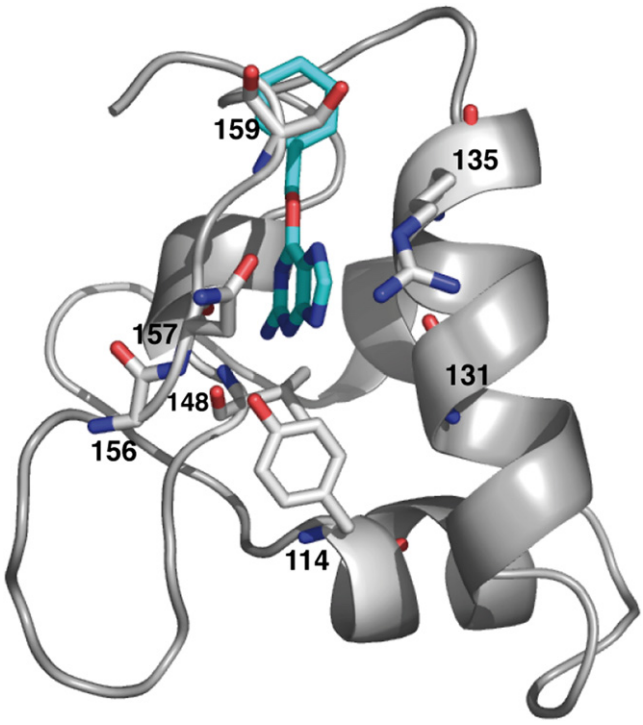

Figure 3. Structure of the Active Site of Wild-Type Human AGT
BG is shown docked into the active site of wild-type human AGT (Juillerat et al., 2003; Wibley et al., 2000). Highlighted residues 114, 131, 135, 148, 156, 157, and 159 were randomized in SNAP-tag for the directed evolution experiment.

with the Aga2p mating agglutinin protein (Colby et al., 2004). Yeast display combined with fluorescence-activated cell sorting (FACS) was chosen because two different fluorophores can be detected, allowing direct selection for specificity using two different competing fluorescent substrates (vide infra) and because the endoplasmic reticulum acts as a quality control that should favor the selection of mutants that possess the stability necessary for applications in protein labeling (Boder and Wittrup, 2000; Colby et al., 2004). The final library contained 2×10^7 individual clones, which represents approximately 2% of the theoretical size of the library on the protein level (1.3×10^9).

The library was subjected to three rounds of screening by FACS. For the first round, the pool of yeast cells was incubated for 30 min with 5 μM BC-FL and cells labeled with fluorescein were retrieved by FACS (Figure S3). For the next two rounds, yeast cells were incubated with both 5 μM BC-FL and 5 μM BG-Cy5 to select for preference of BC over BG. Cells labeled with fluorescein but not with Cy5 were selected by FACS (Figure S3). Thirteen clones retrieved from the third round of sorting were FACS analyzed for their ability to react with BC-FL and BG-FL, and six clones labeled by BC-FL but not BG-FL were analyzed by DNA sequencing. All clones possessed the same sequence (Mut1; Table 1); five of the seven randomized residues were mutated and Gly156 and Val148 were conserved. Kinetic studies with purified protein demonstrated that Mut1 possesses a 10^5-fold switch in substrate specificity (Table 1). However, the protein reacts only 4-fold faster with BC-FL (90 $M^{-1}s^{-1}$) than with SNAP-tag, and this rather low activity toward BC would limit future applications in protein labeling: for example, 26 min is necessary to achieve 50% labeling of Mut1 when incubated with 5 μM BC-FL, whereas 5 s is sufficient to reach 50% labeling when SNAP-tag is incubated with 5 μM BG-FL. Although the activity was lower than desired, Mut1 possessed stability comparable to SNAP-tag (Table 1; Figure S4) and was therefore a reasonable starting point for further improvements in activity. In order to increase the activity of Mut1 toward BC derivatives, we used error-prone PCR to introduce further mutations into Mut1 and phage display to select improved mutants. Phage display was chosen as the selection method, as we have previously shown that it is well suited for increasing the reactivity of AGT (Juillerat et al., 2003). The Mut1 gene was mutated and then inserted in pAK100 phagemid to generate a pool of M13 filamentous phages displaying Mut1 mutants fused to the M13 phage capsid protein pIII (Krebber et al., 1997). The resulting library contained 8×10^7 individual clones with an average of five to six base mutations per gene. In order to isolate mutants with increased activity and the desired specificity, phages were incubated with 0.5 μM BC-FL for 20 min in the presence of 5 μM BG. After four rounds of panning using magnetic beads covered with an anti-fluorescein antibody, 38 clones from the selected pool were analyzed by DNA sequencing. Two thirds of the sequences had a leucine in position 159, which by itself was shown to increase the reactivity of Mut1 by a factor of 6 (see mutant Mut2 in Table S1). The predominance of the mutation F159L confirmed the decision to randomize this position in the first library even though this mutant was not isolated. The failure to select clones containing Leu159 using yeast display is probably because of the fact that only 2% of the theoretical sequence space was screened. The fastest mutant identified in these selections, dubbed CLIP-tag, combined F159L with three additional mutations and was shown to have a 13-fold greater reactivity toward BC-FL (1130 $M^{-1}s^{-1}$) than Mut1. Consequently, the time to achieve 50% labeling of this mutant when incubated with 5 μM BC-FL was reduced to 2 min. Furthermore, CLIP-tag retained the 10^5-fold switch in substrate specificity, exhibiting a 100-fold greater preference for BC over BG (Table 1). In addition, the stability of CLIP-tag was shown to be comparable to SNAP-tag (Table 1; Figure S4).

To verify that CLIP-tag could be efficiently labeled inside mammalian cells, we fused three consecutive simian virus 40 nuclear localization sequences at its C terminus (CLIP-NLS_3) and transiently expressed the corresponding fusion in CHO cells. After incubation for 20 min with 5 μM either BC-DF or BC-OG, fluorescence imaging revealed the nuclear localization of CLIP-NLS_3, demonstrating that CLIP-tag possesses the required activity and stability for covalent labeling in mammalian cells (Figure 4).

Simultaneous and Specific Labeling of Two Fusion Proteins

To demonstrate that SNAP and CLIP fusion proteins can be labeled simultaneously and specifically with different molecular probes, a mixture of CLIP-tag fused to glutathione S-transferase (GST-CLIP) and hexahistidine-tagged SNAP-tag (His-SNAP) was incubated with an equimolar mixture of either BC-Cy3 and BG-Cy5 or BC-Cy5 and BG-Cy3. Reactions were analyzed by SDS-PAGE and in-gel fluorescence scanning, revealing that both proteins were labeled with their preferred substrates with more than 99% selectivity (Figure 5A).

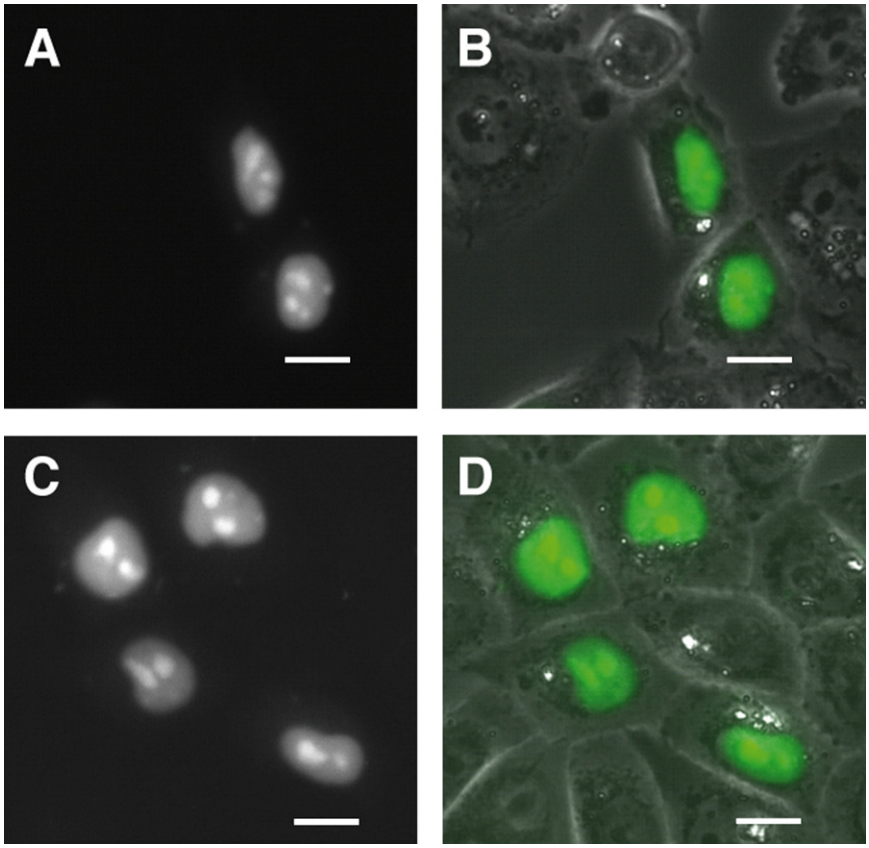

Figure 4. Labeling of CLIP-Tagged Proteins with BC Derivatives

Wide-field micrographs of CHO cells transiently expressing CLIP-NLS_3 and labeled with BC-DF (A and B) and BC-OG (C and D).

(A and C) Fluorescence channel.

(B and D) Merge of transmission and fluorescence micrographs.

Cells were labeled by addition of 5 μM BC derivative for 20 min at 24°C and imaged directly after washing three times with HBSS. The scale bars represent 10 μm.

To show that specific double labeling can also be achieved in living cells, we used HEK293T cells transiently coexpressing either (1) FK506 binding protein (FKBP) fused at the C terminus of CLIP-tag (CLIP-FKBP) plus the FKBP-rapamycin binding domain (FRB) fused at the C terminus of SNAP-tag (SNAP-FRB), or (2) SNAP-FKBP plus CLIP-FRB. FKBP and FRB were arbitrarily chosen because they are cytoplasmic proteins that can be distinguished in SDS-PAGE by their different sizes, thereby enabling easy evaluation of the specificity of labeling by in-gel fluorescence scanning. Incubation of cells with BG-DF and BC-TMR led to the specific labeling of SNAP and CLIP fusions by their native substrates: in the experiment with CLIP-FKBP and SNAP-FRB (Figure 5B, lane 1), fluorescein-labeled SNAP-FRB makes up 95% of the fluorescein-labeled proteins and TMR-labeled CLIP-FKBP makes up more than 99% of TMR-labeled proteins. The difference in the labeling specificities measured in vitro (Figure 5A) and in cells can be explained, at least partially, by the 3.3-fold higher expression level of the FKBP fusion. The specificity of the labeling in these experiments will also be influenced by the relative membrane permeability of the substrates. Similar labeling specificities were obtained in the experiment with CLIP-FRB and SNAP-FKBP (Figure 5B, lane 2). These experiments thus demonstrate that SNAP and CLIP fusions can be simultaneously and specifically labeled by their native substrates in living cells.

One application of simultaneous labeling of two different fusion proteins is the investigation of protein localization and dynamics by fluorescence microscopy. To test the utility of SNAP-tag and CLIP-tag for this application, we transiently expressed CLIP and SNAP fusions with different localizations. For localization at the cytoplasmic side of the plasma membrane, CLIP-tag was expressed with a C-terminal farnesylation motif (CLIP-CaaX); for cytosolic expression, CLIP-tag was fused to the N terminus of β-galactosidase (CLIP-β-Gal); and for nuclear localization, the SNAP-tag was attached to three consecutive simian virus 40 nuclear localization sequences (SNAP-NLS_3). CHO cells transiently coexpressing SNAP-NLS_3 and CLIP-CaaX were incubated with BC-DF and TMR-star, a SNAP-tag substrate leading to labeling with TMR. Analysis of the cells by fluorescence microscopy demonstrated that double labeling can be used for parallel determination of protein localization, in this case the nuclear localization of SNAP-NLS_3 and the predominant insertion of CLIP-CaaX into the plasma membrane (Figure 5C), and also confirmed the specificity of the labeling. Similarly, double labeling also allowed the simultaneous observation of the nuclear localization of SNAP-NLS_3 and the diffuse cytosolic localization of CLIP-β-Gal (Figure 5C).

The ability to distinguish old proteins and newly synthesized proteins by labeling at different time points with different fluorophores is a powerful approach to study dynamic cellular processes. The ability to investigate two dynamic processes simultaneously in one cell through double pulse-chase labeling experiments of SNAP and CLIP fusion proteins would be an attractive extension of this approach. To demonstrate the feasibility of such double pulse-chase experiments, we followed the incorporation of Aga2p into the cell wall of the budding yeast *Saccharomyces cerevisiae* (Figure 6A). Two yeast strains expressing either Aga2p-CLIP or Aga2p-SNAP were generated. A mixture of the two yeast strains was incubated with 5 μM BC-Cy3 and 2 μM BG-Cy5 for 15 min. Imaging by confocal fluorescence microscopy revealed that individual yeast cells were labeled with either Cy3 or Cy5 but not with both fluorophores (Figure 6B), demonstrating the specific labeling of CLIP and SNAP fusion proteins by their native substrates. The yeast cells were then incubated in growth medium for 60 min to allow new protein synthesis and subsequently labeled with BC-FL and BG-Cy3. Analysis by confocal fluorescence microscopy demonstrated that cell wall growth and incorporation of new Aga2p was

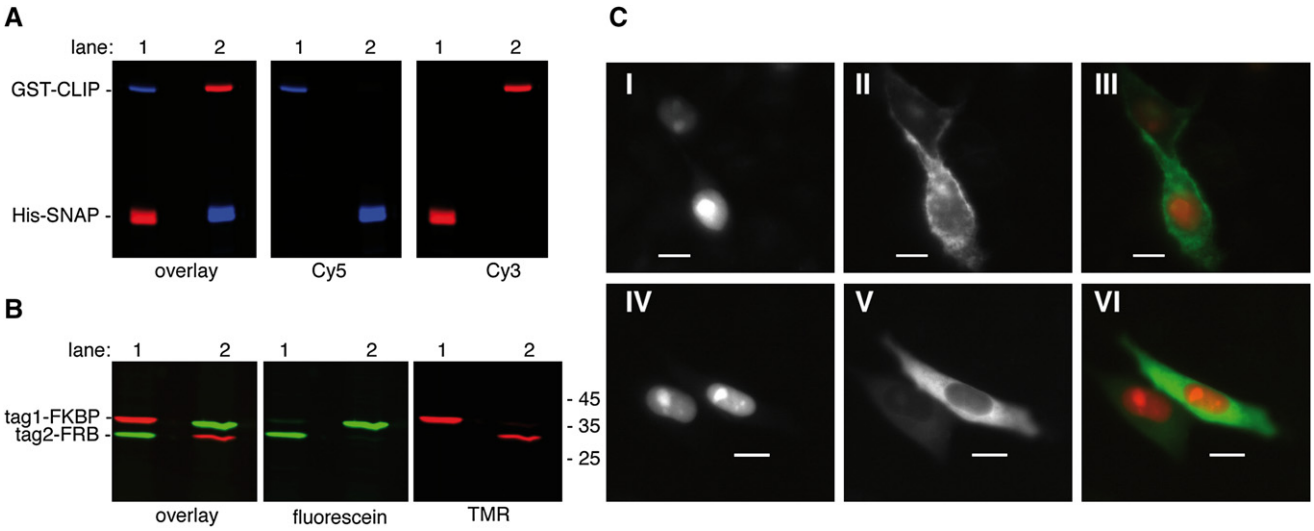

Figure 5. Simultaneous and Specific Labeling of SNAP and CLIP Fusion Proteins In Vitro and in Living Cells

(A) Mixtures of purified His-SNAP and GST-CLIP (0.5 μM each) were labeled by addition of either 5 μM BC-Cy5 and 5 μM BG-Cy3 (lane 1) or 5 μM BC-Cy3 and 5 μM BG-Cy5 (lane 2) for 30 min at 24°C. Samples were analyzed by SDS-PAGE and in-gel fluorescence scanning (red, Cy3; blue, Cy5).

(B) HEK293T cells transiently coexpressing either CLIP-FKBP and SNAP-FRB (lane 1) or CLIP-FRB and SNAP-FKBP (lane 2) were labeled by addition of 5 μM BG-DF and 5 μM BC-TMR for 1 hr at 37°C. After cell lysis, equal amounts of proteins from crude extracts were analyzed by SDS-PAGE and in-gel fluorescence scanning (green, fluorescein; red, TMR). Numerical values of the fluorescence intensities of all bands are shown in Table S2.

(C) (I–VI) Wide-field fluorescence micrographs of CHO cells transiently expressing SNAP-NLS_3 and CLIP-CaaX (I–III), and SNAP-NLS_3 and CLIP-β-Gal (IV–VI) labeled with BC-DF and TMR-star. (I and IV) SNAP-tagged proteins labeled with TMR-star; (II and V) CLIP-tagged proteins labeled with BC-DF; (III and VI) overlay of the fluorescein (green) and TMR (red) channels. Cells were labeled by simultaneous addition of 2 μM TMR-star and 5 μM BC-DF for 20 min at 24°C and imaged directly after washing three times with HBSS. The scale bars represent 10 μm.

directed toward the bud in both yeast strains (Figure 6C): yeast expressing Aga2p-CLIP can be identified by their Cy3-labeled mother cell and a fluorescein-labeled bud, whereas yeast expressing Aga2p-SNAP can be identified by their Cy5-labeled mother cell and a Cy3-labeled bud (Figure 6C). This proof-of-principle experiment shows that SNAP-tag and CLIP-tag can be used in combination for pulse-chase experiments to study dynamic processes such as biological structure formation.

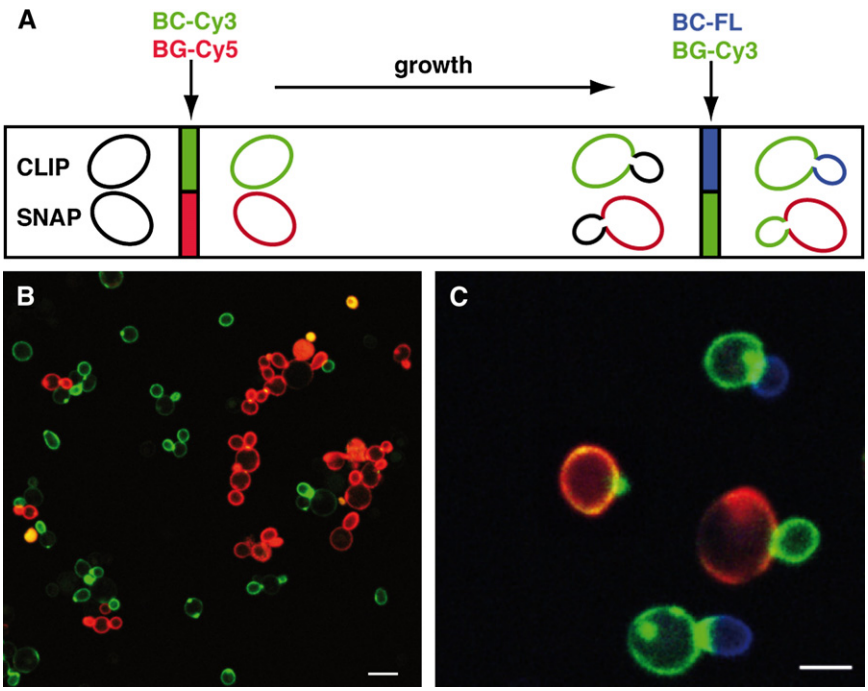

Figure 6. Double Pulse-Chase Experiments of SNAP and CLIP Fusion Proteins

(A) Schematic representation of the double pulse-chase experiment. The two yeast strains expressing either SNAP-Aga2p or CLIP-Aga2p are abbreviated as SNAP and CLIP.

(B) Confocal fluorescence micrograph of the two yeast strains after mixing and labeling with 2 μM BG-Cy5 (red) and 5 μM BC-Cy3 (green) for 15 min.

(C) Confocal fluorescence micrograph of the same yeast cells as in (B) after an additional growth period of 60 min and labeling with 2 μM BG-Cy3 (green) and 5 μM BC-FL (blue) for 15 min.

The scale bars represent 10 μm (B) and 5 μm (C).

DISCUSSION

Mutagenesis of eight amino acids in SNAP-tag led to the generation of CLIP-tag, a self-labeling protein that reacts with high speed and selectivity with O^2-benzylcytosine (BC) derivatives. BC derivatives are accessible by a simple synthetic route that allows the preparation of a large variety of different substrates for protein labeling. In contrast to other substrates used for covalent protein labeling, BC derivatives are chemically relatively inert and show no significant reactivity toward the mammalian proteome. CLIP-tag can be expressed and labeled in different cellular compartments, including the cell surface. Considering the kinship between CLIP-tag and SNAP-tag, we expect that CLIP-tag can be used in all applications where SNAP-tag has been used. This versatility and the current availability of membrane-permeable and -impermeable molecular probes should make CLIP-tag labeling an important addition to existing methods for the covalent labeling of fusion proteins. Concerning applications of CLIP-tag, we have focused in this work on its use in conjunction with SNAP-tag for the simultaneous and specific labeling of two different fusion proteins. Because SNAP-tag shows high selectivity for O^6-benzylguanine (BG) derivatives over BC derivatives, SNAP and CLIP fusion proteins can be used simultaneously for specific labeling with different molecular probes in vitro and in living cells. In the context of multiprotein studies in living cells, SNAP-tag and CLIP-tag have a number of advantages, including (1) the low intrinsic reactivity of their substrates toward other proteins compared to other tags such as tetracysteine tag and HaloTag, (2) the high specificity toward their native substrates, (3) the ability to label these proteins in any cellular compartment, and (4) the similar properties of the two proteins that would aid the comparison of the properties of one fusion protein to another.

Applications for the specific labeling of two proteins inside living cells demonstrated in this work include the simultaneous determination of the cellular localization of two different proteins. Although autofluorescent proteins already provide a straightforward solution to this problem (Giepmans et al., 2006), chemical labeling methods are attractive complements because they allow, for instance, the visualization of proteins in organisms that are not suitable for the expression of autofluorescent proteins (Regoes and Hehl, 2005). Furthermore, chemical labeling is well suited when molecular imaging is followed by other biochemical characterizations such as PAGE or pull-down assays. One of the most attractive applications of chemical labeling is the ability to distinguish young and old copies of a protein by labeling at different time points with different fluorophores. This approach is an elegant alternative to the use of photo-activable or -switchable autofluorescent proteins to track protein over time (Chapman et al., 2005), and gives greater flexibility with respect to colors used. Prominent examples of such pulse-chase experiments include the study of gap junction plaque formation through pulse-chase labeling of connexin-43 using the tetracysteine tag (Gaietta et al., 2002) and the determination of the time point of insertion of CENP-A in centromeres during the cell cycle using SNAP-tag (Jansen et al., 2007). We have shown in this work that it is possible to discriminate different generations of two different proteins, demonstrating that double pulse-chase labeling experiments of SNAP and CLIP fusion proteins could enable the simultaneous investigation of two different dynamic processes. The similar properties of SNAP-tag and CLIP-tag other than their substrate specificity will aid comparison in such experiments.

Finally, future applications of SNAP-tag and CLIP-tag could also include the labeling of two interacting proteins with fluorophores well suited for fluorescence resonance energy transfer (FRET) experiments, including fluorophores for time-resolved FRET (Bazin et al., 2002). Together, these applications should make SNAP and CLIP fusion proteins powerful tools for cell biology.

SIGNIFICANCE

The labeling of proteins with synthetic probes in living cells is a powerful approach to study and manipulate protein function. We have introduced a new approach for the specific labeling of fusion proteins that is based on the irreversible reaction of O^2-benzylcytosine (BC) derivatives with an engineered O^6-alkylguanine-DNA alkyltransferase named CLIP-tag. The reaction between CLIP-tag and BC is fast, and BC derivatives do not possess any significant activity toward the mammalian proteome, thereby giving the system unique specificity among the covalent labeling systems already described. CLIP-tag is also functional in different cellular compartments with no particular restriction. One important application of CLIP-tag will be its use in conjunction with other labeling technologies for the specific labeling of two (or more) different proteins in one cell. We demonstrate here how CLIP and SNAP fusion proteins can be simultaneously and specifically labeled with different synthetic probes in one cell. As a result, simultaneous pulse-chase experiments can be carried out to differentiate different generations of two different proteins in living cells. This represents a significant innovation in the available methodology for studying protein dynamics and the formation of cellular structures. In summary, the labeling of CLIP-tag fusion proteins by BC derivatives is highly specific and orthogonal to other existing labeling approaches, making the method a highly valuable tool for chemical biology.

EXPERIMENTAL PROCEDURES

General

Detailed protocols for chemical syntheses of the BC derivatives used in this work and recombinant DNA work (library construction and cloning) are available in the Supplemental Data. BG derivatives (BG-Cy3, BG-Cy5, BG-FL, BG-DF, and TMR-star) for the labeling of SNAP-tagged proteins were provided by Covalys Biosciences. The HaloTag substrate was obtained from Promega. Cell lines used in this work were CHO-9-neo-C5, HEK293T, HeLa CCL2, HeLa MZ, BHK, and HT29. HeLa CCL2 was obtained from Prof. Pierre Gönczy, Swiss Institute for Experimental Cancer Research (ISREC), and HeLa MZ, BHK, and HT29 cell lines were obtained from Prof. Gisou van der Goot, École Polytechnique Fédérale de Lausanne (EPFL).

SNAP-Tag Mutant Used in This Work

The SNAP-tag mutant used in this work is a recently described engineered AGT with improved expression properties and low DNA binding and high activity toward BG (Gronemeyer et al., 2006). It is a 182 amino acid mutant of the wild-type human AGT in which the last 25 amino acids were deleted and the following mutations were introduced: K32I, L33F, C62A, Q115S, Q116H, K125A, A127T, R128A, G131K, G132T, M134L, R135S, C150Q, S151G, S152D, G153L, A154D, N157G, and S159E.

Reactivity of Substrates against the Mammalian Proteome

Cells resuspended in Hank's balanced salt solution (HBSS; Lonza) were incubated with 10 μM BC-DF, BG-DF, or Halotag-DF at 37°C for 1 hr. After labeling, cells were harvested, washed once with HBSS, and then resuspended in lysis buffer (150 mM KH_2PO_4 [pH 7.0], 100 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100) and subjected to three freeze-thaw cycles (liquid nitrogen/37°C). Cell extracts were analyzed by SDS-PAGE and in-gel fluorescence scanning. Loading on the same gel of a solution of fluorescein-labeled GST-SNAP (0.25 or 5 μM) allowed estimating the molar concentration of the labeled proteins in the extract. The quantity of protein (in pmol/mg of soluble protein) was then determined by dividing the molar concentration (in nM) by the protein concentration of the cell extract (in mg/ml) as determined by Bradford assay.

Selection by Yeast Display

Yeast cells were grown in SD-CAA medium (25 μg/ml kanamycin) at 30°C until the optical density OD_{600} reached 1. Cells were harvested by centrifugation, resuspended in galactose-containing SG-CAA (25 μg/ml kanamycin), and grown overnight at 24°C. After collecting 1 ml of culture with an OD_{600} = 5, cells were harvested by centrifugation and resuspended in 1 ml ice-cold PBS/BSA (1 mg/ml bovine serum albumin). For selection, BC-FL (and BG-Cy5 for the last two rounds) was added to a final concentration of 5 μM, and cells were incubated for 30 min at room temperature. Cells were centrifuged, washed twice with ice-cold PBS/BSA, resuspended in 1 ml PBS/BSA, and filtered through a 100 μm net filter. Cells strongly labeled with fluorescein were sorted with a FACSVantage FACSDiVa (BD Biosciences) equipped with a Coherent Enterprise II laser producing a multiline UV and a 488 nm laser line, and a Coherent Innova Spectrum laser producing a 647 nm laser line. The collected cells were grown at 30°C in SD-CAA (25 μg/ml kanamycin) before storage at −80°C or use for the next round of selection. The first two sortings were performed in enrichment mode and the third one in purification mode. After the third round, individual clones were tested for their ability to react with BC-FL and BG-FL using a FACS CyAn ADP Lx9 (Dako) equipped with 405, 488, and 633 nm laser sources. For FACS analysis, cells were prepared as described above and 100 μl cell solutions in PBS/BSA were incubated with either 0.5 μM BC-FL or 5 μM BG-FL for 30 min.

Selection by Phage Display

Phages were prepared as previously described (Juillerat et al., 2003). Selections were done with a combination of BC-FL (0.5 μM) and BG (5 μM) for 20 min at room temperature. After the labeling reaction, the 1 ml phage preparation was quenched by addition of excess BC and BG. Phage solutions were subjected to two cycles of PEG precipitation using 250 μl of 20% w/v polyethylene glycol 8000, 15% w/v NaCl solution and incubating on ice for 20 min. After centrifugation at 13,000 rpm, 4°C for 15 min, phages were resuspended in 1 ml PBS/milk (4% skimmed milk powder) and gently rotated for 60 min. Two hundred microliters of magnetic beads coated with anti-fluorescein antibody (QIAGEN) were added to the phage preparation and rotated for 30 min. Beads were washed three times with PBS/milk, five times with PBS/Tween (0.05% Tween 20), and twice with PBS. Phages were eluted by incubating the beads with 100 μl of 0.1 M glycine (pH 2.5) for 5 min. The supernatant was neutralized with 50 μl of 1 M Tris-HCl (pH 8) and used to infect *Escherichia coli* JM101 for 30 min at 37°C. Cells were then plated on 2YT supplemented with 1% glucose and 25 μg/ml chloramphenicol and incubated overnight at 37°C.

Characterization of AGT Mutants

Rate constants of the labeling reactions with BC-FL and BG-FL were determined by incubation of purified (GST-) AGT mutants (0.2–0.4 μM) with fluorescent substrate (2–20 μM) in reaction buffer (50 mM HEPES [pH 7.2], 1 mM DTT) at 24°C. Aliquots were taken at different times, boiled at 95°C in SDS buffer for 5 min, and analyzed by SDS-PAGE and in-gel fluorescence scanning using a Pharox FX molecular imager. The data were fitted to a pseudo-first-order reaction model using the Prism software package (GraphPad Software). Second-order rate constants were then obtained by dividing the pseudo-first-order constant by the concentration of substrate. Values given are an average of at least three independent measurements.

To determine the concentration of urea leading to 50% inactivation of protein, purified GST-AGT mutants (0.5 μM) were incubated in reaction buffer (50 mM HEPES [pH 7.2], 1 mM DTT) supplemented with varying concentrations of urea ranging from 0 to 8 M for 30 min. The solutions were then adjusted to 10–20 μM fluorescent substrate and incubated for 1–2 hr. Samples were boiled for 5 min at 95°C in SDS buffer and analyzed by SDS-PAGE and in-gel fluorescence scanning. The fluorescence data set was fitted with a sigmoidal dose-response plot (variable slope) using the Prism software package (GraphPad Software).

In Vitro Double Labeling

A mixture of purified GST-CLIP and His-SNAP (0.5 μM) was incubated in reaction buffer (50 mM HEPES [pH 7.2], 1 mM DTT) at 24°C with 5 μM each BG and BC substrates for 30 min. Labeling reactions were quenched by addition of SDS buffer and incubation for 5 min at 95°C. Samples were analyzed by SDS-PAGE and subsequent in-gel fluorescence scanning.

Double Labeling in Living Cells

HEK293T cells were grown in suspension culture in ExCell-293 medium (JRH Biosciences). For cotransfection of CLIP-FKBP/SNAP-FRB or CLIP-FRB/SNAP-FKBP, 7.5 μg PEI (polyethylenimine in water [pH 7.1]) and 2.5 μg DNA (1:1) were mixed and diluted to a final volume of 100 μl in 150 mM NaCl and incubated at room temperature for 10 min. This transfection cocktail was added to 1 ml of cell suspension (2×10^6 cells/ml in RPMI 1640 medium; Cambrex). After 4 hr of incubation with agitation at 37°C, the transfection mixture was diluted with 1 ml of Pro293s-CD (Cambrex), and the incubation was continued for 24 hr. Transfected cells were harvested and resuspended in HBSS. Cells were labeled by addition of 5 μM BG-DF and 5 μM BC-TMR for 60 min at 37°C. After the labeling step, cells were harvested, washed once with HBSS, and then resuspended in lysis buffer (150 mM KH_2PO_4 [pH 7.0], 100 mM NaCl, 0.5 mM EDTA, 0.1% Triton X-100) and subjected to three freeze-thaw cycles (liquid nitrogen/37°C). Cell extracts were analyzed by SDS-PAGE and in-gel fluorescence scanning. The relative expression level of FKBP and FRB fusions was estimated by comparing the TMR fluorescence of CLIP-FRB and CLIP-FKBP and the fluorescein fluorescence of SNAP-FRB and SNAP-FKBP in the two experiments.

Fluorescence Imaging of Mammalian Cells

Chinese hamster ovary (CHO) 9-neo-C5 cells were grown in DMEM/F12 (Cambrex) supplemented with 5% fetal bovine serum (Cambrex) in a humidified atmosphere under 5% CO_2. Twenty-four hours before transfection, cells were seeded on a μ-Dish (Ibidi) to a density of 75,000 cells per dish. Transient cotransfections were performed using FuGENE-6 transfection reagent (Roche) following the supplier's instructions. For the labeling experiments, BG and BC substrates were added to a final concentration of 2 μM and 5 μM, respectively, in HBSS for 20 min, before washing three times with HBSS. Cells were imaged in HBSS using a Zeiss Axiovert 200 inverted microscope, equipped with an objective LD Plan Neofluar 63×/0.75 corr Ph2 and an AxioCam MR digital camera (Zeiss). Zeiss filter sets 10 (excitation 450–490 nm; emission 515–565 nm) and 43 (ex. 545–625 nm; em. 605–670 nm) were used for fluorescence microscopy.

Fluorescence Imaging of Yeast Cells

Yeast cells were grown in SD-CAA medium at 30°C until the optical density OD_{600} reached 1. Cells were harvested, resuspended in galactose-containing SG-CAA, and grown overnight at 24°C. One and a half milliliters of yeast cells was centrifuged and resuspended in 0.1 ml PBS for labeling experiments. Cells were incubated with 2 μM BG substrate and 5 μM BC substrate for 15 min at 24°C, washed three times with PBS, and resuspended in 0.1 ml PBS for imaging. For the pulse-chase experiment, yeast cells were grown for 60 min in yeast rich medium containing galactose at 30°C before the second labeling step. Cells were imaged with a confocal microscope (Leica TCS-SP2 AOBS) equipped with a glycerol immersion objective HCX PL APO 63×/1.30, with a 488 nm argon laser and a 561 and 633 nm HeNe laser. Fluorescence emission was measured between 505 and 550 nm for fluorescein, 570 and 600 nm for Cy3, and 645 and 750 nm for Cy5.

SUPPLEMENTAL DATA

Supplemental Data include four figures and two tables and can be found with this article online at http://www.chembiol.com/cgi/content/full/15/2/128/DC1/.

ACKNOWLEDGMENTS

The authors thank Didier Surdez for technical expertise during FACS sorting, Gisou van der Goot and Pierre Gönczy for providing cell lines, and Helen M. O'Hare, Marlon Hinner, and Simone Schmitt for critical reading of the manuscript. This work was supported by the European Research Training Network "ProSA," the Swiss National Science Foundation, and the Human Frontiers Science Program. The authors declare financial interests: M.K., F.B., and K.J. are cofounders (M.K., K.J.) or employees (F.B.) of Covalys Biosciences, which aims to exploit CLIP-tag for applications in protein labeling.

Received: December 3, 2007
Revised: January 12, 2008
Accepted: January 17, 2008
Published: February 22, 2008

REFERENCES

Bazin, H., Trinquet, E., and Mathis, G. (2002). Time resolved amplification of cryptate emission: a versatile technology to trace biomolecular interactions. J. Biotechnol. *82*, 233–250.

Boder, E.T., and Wittrup, K.D. (2000). Yeast surface display for directed evolution of protein expression, affinity, and stability. Methods Enzymol. *328*, 430–444.

Chapman, S., Oparka, K.J., and Roberts, A.G. (2005). New tools for in vivo fluorescence tagging. Curr. Opin. Plant Biol. *8*, 565–573.

Colby, D.W., Kellog, B.A., Graff, C.P., Yeung, Y.A., Swers, J.S., and Wittrup, K.D. (2004). Engineering antibody affinity by yeast surface display. Methods Enzymol. *388*, 348–358.

Daniels, D.S., Woo, T.T., Luu, K.K., Noll, D.M., Clarke, N.D., Pegg, A.E., and Trainer, J.A. (2004). DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat. Struct. Mol. Biol. *11*, 714–720.

Fasman, G.D. (1975). Handbook of Biochemistry and Molecular Biology, Third Edition (Cleveland, OH: CRC Press).

Foote, R.S., and Mitra, S. (1984). Lack of induction of O^6-methylguanine-DNA methyltransferase in mammalian cells treated with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine. Carcinogenesis *5*, 277–281.

Gaietta, G., Deerinck, T.J., Adams, S.R., Bouwer, J., Tour, O., Laird, D.W., Sosinsky, G.E., Tsien, R.Y., and Ellisman, M.H. (2002). Multicolor and electron microscopic imaging of connexin trafficking. Science *296*, 503–507.

Giepmans, B.N.G., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. (2006). The fluorescent toolbox for assessing protein location and function. Science *312*, 217–224.

Griffin, B.A., Adams, S.R., and Tsien, R.Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. Science *281*, 269–272.

Gronemeyer, T., Chidley, C., Juillerat, A., Heinis, C., and Johnsson, K. (2006). Directed evolution of O^6-alkylguanine-DNA alkyltransferase for applications in protein labeling. Protein Eng. Des. Sel. *19*, 309–316.

Heinis, C., Schmitt, S., Kindermann, M., Godin, G., and Johnsson, K. (2006). Evolving the substrate specificity of O^6-alkylguanine-DNA alkyltransferase through loop insertion for applications in molecular imaging. ACS Chem. Biol. *1*, 575–589.

Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. J. Cell Biol. *176*, 795–805.

Johnsson, N., and Johnsson, K. (2007). Chemical tools for biomolecular imaging. ACS Chem. Biol. *2*, 31–38.

Juillerat, A., Gronemeyer, T., Keppler, A., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2003). Directed evolution of O^6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. Chem. Biol. *10*, 313–317.

Kaina, B., Christmann, M., Naumann, S., and Roos, W.P. (2007). MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. DNA Repair (Amst.) *6*, 1079–1099.

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat. Biotechnol. *21*, 86–89.

Keppler, A., Kindermann, M., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2004a). Labeling of fusion proteins of O^6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. Methods *32*, 437–444.

Keppler, A., Pick, H., Arrivoli, C., Vogel, H., and Johnsson, K. (2004b). Labeling of fusion proteins with synthetic fluorophores in live cells. Proc. Natl. Acad. Sci. USA *101*, 9955–9959.

Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Boshard, H.R., and Pluckthun, A. (1997). Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. J. Immunol. Methods *201*, 35–55.

Lemercier, G., Gendreizig, S., Kindermann, M., and Johnsson, K. (2007). Inducing and sensing protein-protein interactions in living cells via selective crosslinking. Angew. Chem. Int. Ed. Engl. *46*, 4281–4284.

Los, G.V., Darzins, A., Karassina, N., Zimprich, C., Learish, R., McDougall, M.G., Encell, L.P., Friedman-Ohana, R., Wood, M., Vidurgiris, G., et al. (2005). HaloTag™ interchangeable labeling technology for cell imaging, protein capture and immobilization. Promega Cell Notes *11*, 2–6.

O'Hare, H.M., Johnsson, K., and Gautier, A. (2007). Chemical probes shed light on protein function. Curr. Opin. Struct. Biol. *17*, 488–494.

Regoes, A., and Hehl, A.B. (2005). SNAP-tag mediated live cell labeling as an alternative to GFP in anaerobic organisms. Biotechniques *39*, 809–812.

Schultz, C., Schleifenbaum, A., Goedhart, J., and Gadella, T.W.J., Jr. (2005). Multiparameter imaging for the analysis of intracellular signaling. ChemBioChem *6*, 1323–1330.

Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide for choosing fluorescent proteins. Nat. Methods *2*, 905–909.

Tour, O., Adams, S.R., Kerr, R.A., Meijer, R.M., Sejnowski, T.J., Tsien, R.W., and Tsien, R.Y. (2007). Calcium Green FlAsH as a genetically targeted small-molecule calcium indicator. Nat. Chem. Biol. *3*, 423–431.

Wibley, J.E.A., Pegg, A.E., and Moody, P.C.E. (2000). Crystal structure of the human O^6-alkylguanine-DNA alkyltransferase. Nucleic Acids Res. *28*, 393–401.

