



ELSEVIER

Available online at www.sciencedirect.com

How to obtain labeled proteins and what to do with them

Marlon J Hinner and Kai Johnsson

We review new and established methods for the chemical modification of proteins in living cells and highlight recent applications. The review focuses on tag-mediated protein labeling methods, such as the tetracysteine tag and SNAP-tag, and new developments in this field such as intracellular labeling with lipoic acid ligase. Recent promising advances in the incorporation of unnatural amino acids into proteins are also briefly discussed. We describe new tools using tag-mediated labeling methods including the super-resolution microscopy of tagged proteins, the study of the interactions of proteins and protein domains, the subcellular targeting of synthetic ion sensors, and the generation of new semisynthetic metabolite sensors. We conclude with a view on necessary future developments, with one example being the selective labeling of non-tagged, native proteins in complex protein mixtures.

Address

Institute of Chemical Sciences and Engineering, Laboratory of Protein Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Corresponding authors: Hinner, Marlon J (marlon.jakob@hinner.de) and Johnsson, Kai (kai.johnsson@epfl.ch)

Current Opinion in Biotechnology 2010, **21**:766–776

This review comes from a themed issue on
Chemical biotechnology
Edited by Phil Holliger and Karl Erich Jaeger

Available online 26th October 2010

0958-1669/\$ – see front matter
© 2010 Elsevier Ltd. All rights reserved.

DOI [10.1016/j.copbio.2010.09.011](https://doi.org/10.1016/j.copbio.2010.09.011)

Introduction

Chemists are becoming increasingly fascinated with derivatizing proteins by genetically non-encodable synthetic molecules. As discussed in this review, such molecules include fluorescent dyes, chemical crosslinkers, pharmacologically active compounds, and synthetic fluorescent probes for ions such as Ca^{2+} . The labeling of proteins with synthetic molecules provides exciting new tools for studying proteins and their function in the cell, and is promising to have a strong impact on drug development. In this review, we will provide a concise overview over established and new methods that provide proteins derivatized with a label, and give illustrative recent examples of their application. For further information, the reader is directed to recently published more exhaustive reviews [1–4]. The focus of our review lies on covalent tag-based labeling

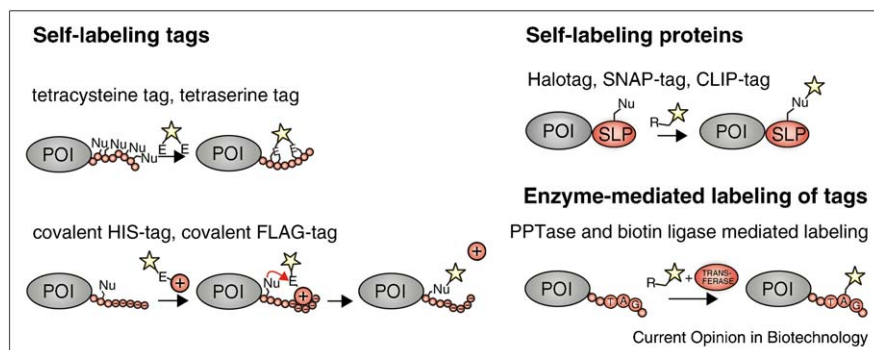
methods, as these allow an efficient and irreversible transfer of labels in mammalian cells. However, we will also comment on unnatural amino acid incorporation as a tool of increasing importance for mammalian cell studies. The discussion of methods and applications is preceded by general considerations on tag-mediated labeling. We will conclude with an outlook on future directions and highlight areas that would profit most from new developments.

Tag-mediated labeling

An ideal method for tag-based protein labeling should exhibit the following features: (i) the possibility to introduce any label of choice in one step, (ii) fast and quantitative labeling, (iii) no labeling of non-target proteins, (iv) a small tag to minimize its impact on protein function, (v) the formation of a stable, covalent bond between protein and label, and (vi) no side effects of the reagents used for labeling. Finally, an ideal tag should work *in vitro*, in complex protein samples, on the cell surface, within the cell and cellular compartments, and in living animals (*in vivo*), with this order representing an increasing level of difficulty. None of the existing labeling methods fulfils all these requirements. It is notable, however, that the existing methods of tag-mediated labeling can be grouped into three families (Figure 1) that have inherent strengths and weaknesses. *Self-labeling tags* and *self-labeling proteins* – such as the tetracysteine-tag and the SNAP-tag, respectively – are able to directly react with the labeling compound. Provided that the labeling molecule is cell membrane permeable, they can therefore be applied to labeling proteins in the cell interior. Compared to self-labeling tags, self-labeling proteins usually provide a higher labeling specificity, but the higher specificity inevitably comes at the price of a larger tag size. *Enzyme-mediated labeling of tags* (such as Biotin ligase mediated labeling) additionally requires an enzyme to covalently link the labeling compound to the tag. In many cases, this approach combines the small tag size provided by self-labeling tags and the specificity provided by enzymes. However, enzyme-mediated labeling of tags is restricted to labeling cell surface proteins unless the enzyme can be expressed within the cell and intracellular metabolites do not interfere with labeling. Some enzyme-mediated labeling methods are limited with regard to a free choice of the label. This applies to the cases where the label itself plays a role in substrate recognition, as discussed below. By contrast, the label usually has a minor impact on the labeling rate of self-labeling tags and self-labeling proteins, enabling the transfer of – in principle – arbitrary labels.

How important is the size of the tag? The plethora of reported successful fusions of autofluorescent proteins to

Figure 1



Families of tag-mediated protein labeling, prototypical family members, and generalized labeling scheme. Self-labeling tags, top: A peptide motif fused to the protein of interest (POI) reacts directly with the labeling compound. The specificity is provided by the matching spatial arrangement of reactive groups in the tag and the labeling compound. Self-labeling tags, bottom: A bifunctional compound containing a label and a recognition element first binds noncovalently to the tag sequence in the protein of interest. The noncovalent interaction is turned into a covalent one by reaction of a proximal amino acid side chain with the labeling compound with a concomitant release of the recognition element. Self-labeling proteins: The labeling compound contains both the label and a chemical moiety that mediates the covalent reaction with the self-labeling protein (SLP). Enzyme-mediated labeling of tags: A transferase that specifically recognizes a tag sequence appended to the protein of interest is used to label the protein. This requires a labeling compound containing the label and the chemical moiety recognized by the enzyme. Nu: Nucleophilic or basic residue; E: Electrophilic moiety; R: moiety required for recognition by the enzyme used for labeling.

various proteins demonstrates that a large tag is usually not problematic. However, while the impact of a tag on the function of the tagged protein is – unfortunately – rarely rigorously checked, in some cases small tags have been reported to perform significantly better than auto-fluorescent protein fusions. As may be expected, this involves proteins that are part of a tightly packed structure or involved in a translocation process. For example, an impact of tag size on protein function has been shown for virion surface proteins [5], beta-tubulin [6], beta-actin with regard to transport into the cell nucleus [31**], and effector proteins that are translocated by gram-negative bacteria into eukaryotic cells [7]. Further, large tags are usually restricted to N-terminal or C-terminal fusions, while small tags can in principle also be incorporated at internal sites of the protein of interest.

In the following, we introduce established methods and new developments for each of the tag-mediated labeling families. An overview over the self-labeling tags and proteins discussed in this review is shown in Table 1, providing data on the labeling rate and efficiency, and indicating whether the method works for intracellular labeling. An overview over methods used for enzyme-mediated labeling of tags is given in Table 2; as a rough guide to the labeling rate attainable with such methods, we provide typical published labeling protocols for the labeling of tagged cell surface proteins.

Self-labeling tags

The first tag developed for specific protein labeling is the tetracysteine tag [8], which exploits the large affinity of peptide sequences containing four cysteines – usually

Table 1

Overview of self-labeling tags and proteins. The second order rates given are rounded published literature values for typical substrates [12*,13,16*,76], except for the values marked with a star that were estimated from the published data [10,11,17,19]. The labeling efficiency is the maximum percentage of labeled proteins obtained by *in vitro* labeling as taken from published data. n.d.: not determined

	Labeling method	Rate [$M^{-1} s^{-1}$]	Intracellular labeling	Labeling Efficiency [%]
Self-labeling tags	Tetracysteine tag	10^4 – 10^5	Yes	100
	Tetraserine tag	n.d.	Not shown	n.d.
	Covalent His-tag labeling	5*	Not shown	80
	Covalent FLAG-tag labeling	1*	Not shown	30
	N-cyanobenzothiazole labeling	10	Not shown	100
Self-labeling proteins	SNAP-tag	3×10^4	Yes	100
	CLIP-tag	10^3	Yes	100
	HaloTag	3×10^6	Yes	100
	BL-tag	10^*	Not shown	n.d.
	Covalent TMP-tag	10–100	Yes	100

Table 2

Overview of enzymes used in enzyme-mediated labeling. Typical labeling protocols were taken from the literature [30,77–79]. ^aLipoic acid ligase needs to be engineered to accept different labels, but it seems likely that this is possible for various synthetic molecules

Enzyme used	Typical labeling protocol for cell surface labeling	Intracellular labeling	Free choice of labels
Sfp-labeling	2 μ M Enzyme, 1 μ M Label, 30 min	Not shown	Yes
AcpS-labeling	2 μ M Enzyme, 1 μ M Label, 20 min	Not shown	Yes
Biotin Ligase	0.3 μ M Enzyme, 10 μ M Label, 1-60 min	Yes	No
Lipoic acid ligase	10 μ M Enzyme, 500 μ M Label, 5 min	Yes	No ^a
Sortase	200 μ M Enzyme, 100 μ M Label, 10-30 min	Not shown	Yes

CCPGCC – towards synthetic molecules containing two arsenic atoms, also called biarsenicals. The detection of labeled proteins is simplified by the fact that the fluorescence of biarsenical dyes such as FIAsh and ReAsH is switched on by binding to the tag. FIAsh and ReAsH are cell permeable, making tetracysteine labeling the best established method for the fluorescent labeling of small tags within cells. The specificity of biarsenical labels for their tag is not perfect; labeling and washing procedures therefore have to be carried out in presence of a competing thiol reagent to reduce background labeling. This labeling procedure – which is also slightly cytotoxic [31^{••}] – may be problematic for tetracysteine tag applicability *in vivo*.

Recently, a similar self-labeling tag was designed with the aim of eliminating thiol reagents. In this method, dyes derivatized with boronic acids are designed to react with a tetraserine tag [9]. In a proof-of-principle experiment, the high-affinity binding of the label to the tag was demonstrated, but more development is needed to selectivity label tagged proteins within the cell. A series of self-labeling tags was also developed by Hamachi and co-workers [10,11]. They are based on bifunctional molecules containing a recognition element for the protein of interest and a reactive group. The initially noncovalent interaction between the protein of interest and the labeling compound is turned into a covalent one owing to reaction of the reactive group with a proximal amino acid side chain in the protein, as illustrated in Figure 1. The reaction leads to a release of the recognition element. The approach works selectively – but rather slowly – in complex mixtures using the interaction between positively charged Ni²⁺-containing reagents and the 6-His-tag or 10-His tag [10] or the negatively charged FLAG-tag [11]. A new development in self-labeling tags is based on the reaction of N-terminal cysteines with N-cyano-benzothiazole derivatives containing a label [12[•]]. The reaction is selective owing to the requirement of an amino group in β -position to the cysteine side chain, which is provided by the free N-terminal amine. N-terminal cysteines do not occur naturally in the cell, but can be obtained by enzymatic cleavage of an introduced recognition site for Tobacco Etch Virus (TEV) Protease (ENLYFQ↓C, cleavage site marked with an arrow) or from proteins fused to an intein. The authors demonstrate specific, but rather slow labeling in cell

lysates and on the surface of living cells; an intracellular labeling has not yet been shown.

Self-labeling proteins

Self-labeling proteins react covalently with a substrate that is linked to the label of interest and are mostly derived from enzymes. The most widely used self-labeling proteins are the HaloTag [13] and the SNAP-tag [14], the latter of which was developed in our laboratory. The advantages of these self-labeling proteins are their high speed and specificity (cf. Table 1), the large array of commercially available substrates, and the simple synthetic accessibility of arbitrary custom labels. Owing to the fact that labeling by HaloTag and SNAP-tag is irreversible and quantitative, these tags are well suited for the detection and quantitation of tagged proteins via in-gel fluorescence scanning of SDS-PAGE gels. HaloTag and SNAP-tag work well within cells and also within subcellular compartments when cell-permeable substrates are used. For example, SNAP-tag labeling has been described not only for proteins in the cytosol, but also in the nucleus, endoplasmic reticulum, mitochondria, and golgi apparatus (see, e.g. refs. [63,67]). Even non-permeable substrates that are microinjected do not lead to background staining, as excess dye usually does not remain within the cell [15].

Self-labeling proteins have recently been shown to be employable also for *in vivo* studies [43[•]], as discussed in the applications section below. Another recent development in this field is a further self-labeling protein, CLIP-tag [16[•]]. CLIP-tag allows a free choice of the label and owing to its orthogonality can be used in conjunction with SNAP-tag and HaloTag. An alternative method uses a mutant beta-lactamase that reacts covalently with a beta-lactam derivatized with a quencher and a fluorescent dye [17]. The method, termed BL-tag, is however slower than already existing methods (Table 1) and restricted to the cell surface. Recently, a covalent version of the noncovalent trimethoprim-tag [18] (TMP-tag) has been developed that is based on proximity labeling. It was shown to be applicable to intracellular labeling of proteins tagged with *E. coli* dihydrofolate reductase [19].

Enzyme-mediated labeling of tags

In this family of methods, an enzyme is used to attach a label to a usually small recognition sequence in the

protein of interest. A well established method is phosphotransferase (PPTase)-mediated labeling, using the PPTases Sfp [20] or AcpS [21]. The natural substrate of these enzymes is CoA, but CoA derivatives linked to any label can also be accepted as substrates. The method is restricted to labeling cell surface proteins because intracellular CoA would interfere with labeling. Owing to continuous engineering efforts, the tag size for PPTase-mediated labeling has been significantly reduced [22,23]. This makes PPTase-mediated labeling probably the best choice for labeling extracellular small tags with arbitrary substrates. *E. coli* Biotin Ligase (BirA) is another well established labeling enzyme that can be used for the specific transfer of biotin and of a biotin isostere with a ketone functionality [24]. Recent reports demonstrate that biotin ligases from other species can be used to label proteins with alkyne-functionalized and azide-functionalized biotin [25]. The functionalized biotins can serve as a handle to introduce arbitrary labels in a second step using selective reactive chemistry. BirA can be expressed within mammalian cells, and has been used for the labeling of intracellular proteins (cf., e.g. refs. [26,27]). The limitation of this broadly applicable method is the presence of other biotinylated proteins within cells and the fact that arbitrary labels such as fluorescent dyes cannot be transferred in one step.

A major breakthrough for the intracellular labeling of proteins with various substrates by enzyme-mediated labeling has recently been published, termed lipoic acid ligase mediated labeling. In initial reports, it has been shown that the enzyme efficiently transfers labels containing reactive bromides, azides, and photocrosslinkers to tagged proteins [28,29,30]. In a very recent publication, lipoic acid ligase was engineered to also accept the fluorescent dye coumarin as a substrate, and it has been shown that lipoic acid ligase can be expressed in mammalian cells, enabling intracellular labeling [31]. The utility of this method has been demonstrated for the specific labeling of intracellular tagged proteins in the cytosol and the nucleus. As of now, lipoic acid ligase does not allow a free choice of the label, because the introduced label has to fit into a cavity within the enzyme; new labels therefore require a reengineering of lipoic acid ligase. A further new method uses sortase to mediate specific protein labeling. Sortase is used to specifically join a peptide or protein containing a five-amino acid C-terminal recognition tag (usually LPXTG) with a second peptide or protein containing multiple glycines at its N-terminus. The method has been successfully used for the C-terminal [32,33] and N-terminal labeling of proteins [34,35]. The method works with high specificity on living cells, but is presumably restricted to the cell surface.

Applications of tag-mediated labeling

Study of protein function in living cells

In the past two years, tag-mediated labeling methods have increasingly been applied in live cells to study

protein localization, dynamics, and trafficking, as well as the stoichiometry of protein complexes [5,36,37–41,42]. An illustrating example [40,41] is given by two studies on a protein involved in Hedgehog signaling, which were carried out using SNAP-tag [40] and ACP labeling [41], respectively.

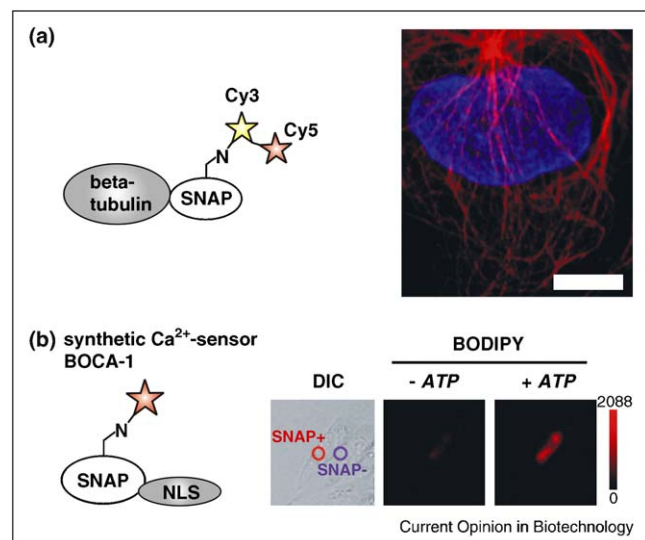
Self-labeling proteins have recently been demonstrated to work in living animals. This allows the introduction of fluorophores that absorb and emit light in the far-red spectral window, which improves imaging in deep tissue compared to autofluorescent proteins. In addition, the labeling time point can be freely chosen, allowing pulse-chase imaging *in vivo*. Kosaka *et al.* generated an ovarian cancer cell line stably expressing HaloTag at the cell surface [43]. This cell line was used to generate model tumors in the peritoneal cavity of living mice. The tumors were stained by direct injection of fluorescent HaloTag ligand into the peritoneal cavity and could be imaged noninvasively in the living animals.

Biophysical methods are increasingly being combined with tag-mediated labeling. The chromophore-assisted light inactivation of proteins (CALI) utilizing proteins labeled with a fluorophore via the tetracysteine-tag [44,45] is one such example. This previously established method has again successfully been applied to study the role of the Clathrin light chain in vesicle formation and its impact on neuronal signaling in the neuromuscular junction of drosophila flies [46]. Recently, it has been shown that CALI also works efficiently based on SNAP-tag labeling [47].

Tag-mediated labeling approaches are ideal platforms for the introduction of labels for super-resolution imaging in cells. This has now been demonstrated using HaloTag [48] or SNAP-tag fusion proteins [49] with the stimulated emission depletion (STED) approach, as well as using a SNAP-tag fusion protein with the stochastic optical reconstruction microscopy approach (STORM, Figure 2a) [50]. Super-resolution imaging based on tag-mediated labeling of the proteins of interest combines the advantages of conventional approaches based on antibodies or autofluorescent proteins, namely the possibility to apply super-resolution imaging to living cells and the free choice of the utilized fluorophore [49]. Compared to autofluorescent proteins, synthetic fluorophores offer the advantage of a wider array of available colors, and, depending on the chosen fluorophore, improved photostability. This is important because the attainable resolution depends on the photostability of the utilized dye [80].

As shown in a recent publication [51], tag-mediated labeling is also a good strategy for investigating protein mobility in living cells by fluorescence recovery after photobleaching or photoactivation.

Figure 2



New methods based on tag-mediated protein labeling. **(a)** Super-resolution imaging. SNAP-tagged beta-tubulin was labeled with the Cy5 and Cy3 fluorophores linked in tandem. The tandem chromophore is compatible with STORM imaging [50]. In the overlay of fluorescence images, beta-tubulin is colored red, while the nuclear counterstain is shown in blue. **(b)** Example for a localizable synthetic Ca^{2+} -sensor. SNAP-tag fused to a nuclear localization sequence (NLS) was expressed in CHO-K1 cells and labeled with the Ca^{2+} -sensitive fluorophore BOCA-1 [66]. Stimulation of Ca^{2+} -release by addition of ATP to the cells leads to a change in fluorescence that can be monitored locally.

New tools for biology

The possibility to generate hybrid proteins containing an expressed and a synthetic part is increasingly being exploited to make novel tools for biology. A particularly pressing problem in biology is the study of protein interactions. A previously introduced approach based on the tetracysteine tag (Figure 3a) uses a trifunctional compound containing a biarsenical dye to bind to the protein of interest, a crosslinker triggered by addition of sodium periodate to induce tethering to binding partners, and biotin for detection of the interaction partner after SDS-PAGE and western blotting [52]. This affinity labeling method has recently been applied to study the interaction between ubiquitin and the proteasome [53]. A different method based on *E. coli* Biotin Ligase works by fusing the two proteins of interest to biotin ligase and to the biotin ligase acceptor peptide, respectively (Figure 3b). The efficiency of biotinylation of the acceptor peptide by BirA is increased by induced proximity if the two tagged proteins interact. This method can be applied to study protein interactions with overexpressed proteins in the cell interior [54]. A number of examples have been published that utilize a split tetracysteine motif that can only be labeled if two proteins or protein domains interact [55–57]. This strategy has been applied to study protein association and protein folding. We have recently

further improved a method based on SNAP-tag developed in our laboratory [58]. In this improved approach, termed S-CROSS [59^{*}], the proteins of interest are fused to SNAP-tag and CLIP-tag, respectively (Figure 3c). The tool used for detecting protein interactions is a trifunctional molecule, containing a fluorescent dye and the SNAP-tag and CLIP-tag substrates. The efficiency of crosslinking of the tagged proteins of interest is strongly dependent on whether they are associated; the interaction can therefore be identified by detection of the fluorescent crosslinked protein pair by SDS-PAGE analysis and in-gel fluorescence scanning (Figure 3d).

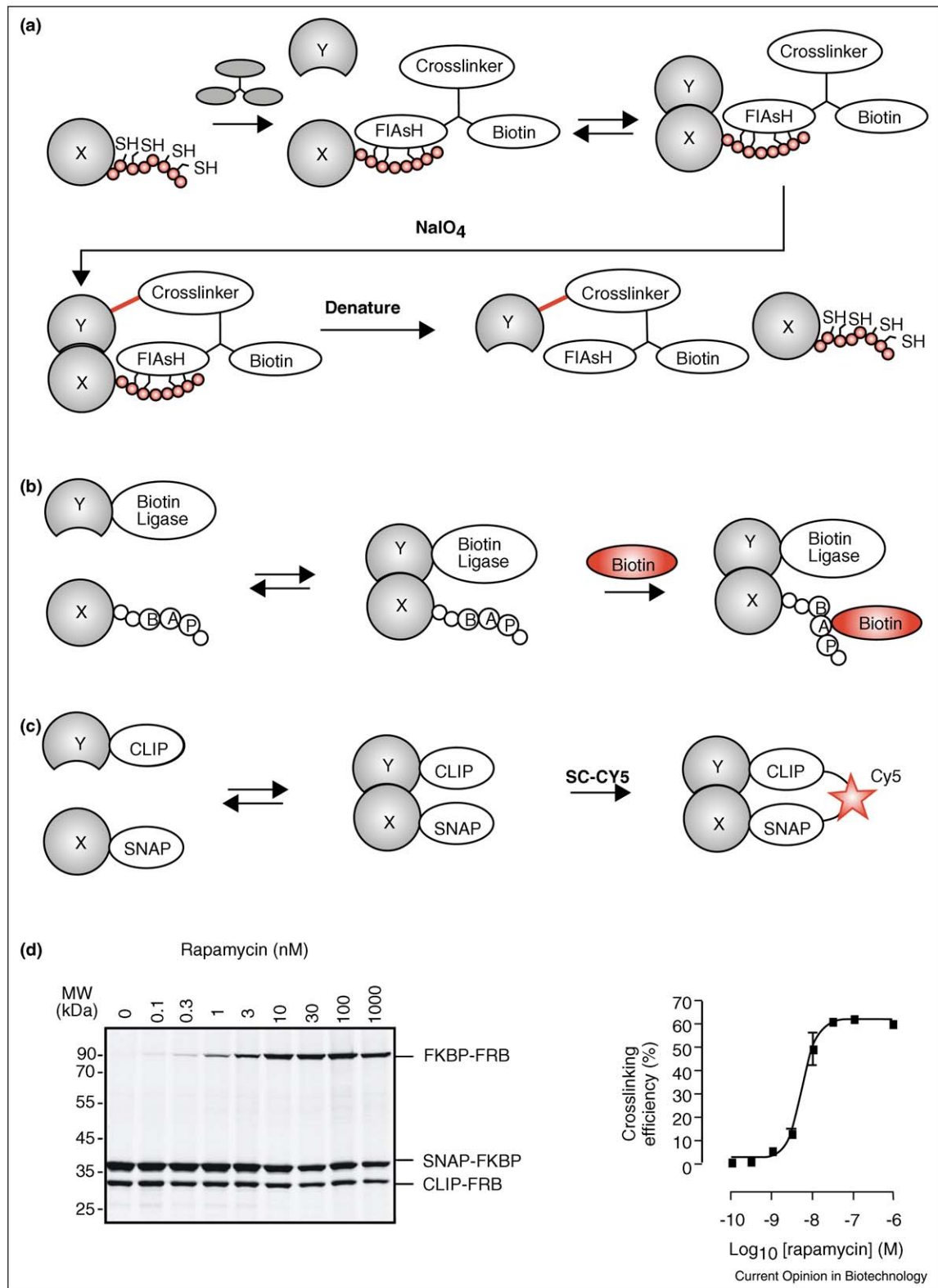
A further example how tag-mediated labeling can be exploited for the construction of novel tools is a small-molecule sensor principle as depicted in Figure 4 [60^{**}]. The method is similar to FRET-based biosensors that rely on a bacterial binding protein sandwiched between a FRET pair of autofluorescent proteins; the prototype of this established design was the maltose sensor developed by Fehr *et al.* [61], a recent prominent example are the glutamate sensors developed by the Tsien laboratory [62]. In a proof-of-principle experiment, the new sensor principle was tested with human carbonic anhydrase (HCA) as the binding protein, and used to build a sensor for HCA inhibitors. The sensor exhibited a favorable sensitivity compared to previous approaches. The decisive novelty in this design is the fact that sensors for metabolites can be generated based on metabolite binding proteins – such as HCA – that do not undergo a conformational change upon ligand binding.

With tag-mediated labeling, the targeting of synthetic sensors to specific subcellular locations and organelles is possible. This allows combining the subcellular targeting provided by genetic methods with the sensitive and fast response provided by synthetic sensors. The SNAP-tag technology has recently been applied to targeting sensors for Zn^{2+} [63], Ca^{2+} [64–66], and probes for hydrogen peroxide [67]. This strategy has also been applied to pH sensors that can be covalently linked to HaloTag [68]. A recent example for the nuclear localization of a highly sensitive BODIPY-based Ca^{2+} -sensitive dye is shown in Figure 2b [66]. Targeted ion sensors allow the detection of local ion concentration changes in the cell with a high spatial resolution; targeting the sensors to locations within the cell that experience a large analyte concentration change will inherently improve signal to noise ratios for such measurements. The application of targetable sensors to complex biological problems will help in revealing the importance of local concentration changes on cell function.

Incorporation of unnatural amino acids

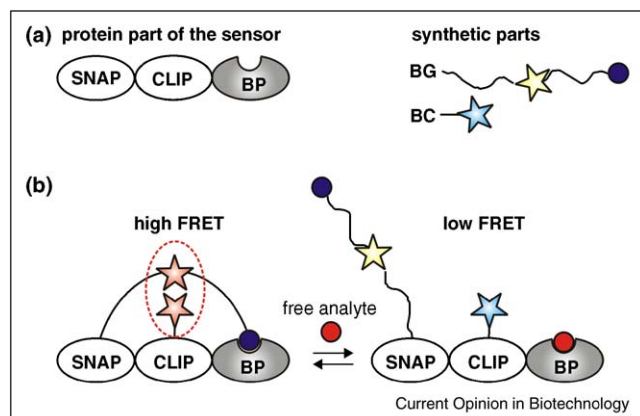
The most elegant way to obtain proteins containing non-natural functionalities and labels is to incorporate these already during protein synthesis using modified amino

Figure 3



New methods for the detection of protein interactions. **(a)** Label transfer chemistry based on tetracysteine tag labeling. A trifunctional molecule is used containing the biarsenical molecule FIASH that binds to the tetracysteine tag, biotin for the detection of interaction partners, and a crosslinker inducible by sodium periodate addition. **(b)** Proximity biotinylation. **(c)** Selective crosslinking of interacting proteins (S-CROSS). The trifunctional molecule used for crosslinking (SC-CY5) contains the fluorescent dye Cy5 and the substrates of SNAP-tag and CLIP-tag. **(d)** Application example for S-CROSS. The data shown correspond to the rapamycin-dependent interaction of FKBP and FRB [59]. The fluorescent SDS-PAGE gel image shows that crosslinking is dependent on the presence of rapamycin and therefore on the interaction of the two proteins.

Figure 4



Semisynthetic metabolite sensor. **(a)** Assembly. The sensor is semisynthetic: its proteinaceous part consists of a triple fusion protein of SNAP-tag, CLIP-tag, and a binding protein [60^{**}]. SNAP-tag is labeled with a synthetic part that consists of the SNAP-tag substrate benzylguanine (BG), a fluorophore, and a recognition element that interacts with the binding protein. CLIP-tag is labeled with a second fluorophore. **(b)** Mechanism. The efficiency of FRET between the two fluorophores is dependent on the presence of the analyte, yielding a ratiometric readout for metabolite concentration.

acids. Recent developments show the increasing utility of unnatural amino acid (UAA) incorporation for studies in mammalian cells. We therefore provide here a brief overview with selected examples, without going into the details of this large and complex field.

The roles of both tRNA and aminoacyl-tRNA synthetase in protein synthesis are central to UAA. Each tRNA is loaded with its respective amino acid by a specific aminoacyl-tRNA synthetase. Upon recognition of a codon in the coding mRNA by means of a specific anticodon, the tRNA transfers its amino acid to a nascent protein chain. There are three general ways to exploit these parts of the protein translation machinery to achieve UAA incorporation into proteins: (i) Aminoacyl-tRNA synthetases can also load their cognate tRNAs with a close analog of the respective canonical amino acid. When bacterial strains auxotrophic for the canonical amino acid are grown in the presence of the amino acid analog, the analog will replace the canonical amino acid in all proteins. A recent application of this method is the generation of intrinsically colored proteins using fluorescent tryptophan analogs [69^{*}]. (ii) *In vitro*, tRNA can be loaded with an unnatural amino acid that does not correspond to its original amino acid. If protein synthesis takes place in the presence of this mischarged tRNA, the unnatural amino acid is incorporated into the nascent protein upon translation of the codon corresponding to the utilized tRNA. To allow an orthogonal introduction of the unnatural amino acid, a tRNA corresponding to a codon not used for natural amino acid encoding has to be used, also referred

to as a free codon. This is accomplished by tRNA interacting with a stop codon or a four-base codon. This methodology is usually used in cell-free expression systems, but can also be applied to *Xenopus* oocytes by injection of the mischarged tRNA. The advantage of this approach is the possibility to use unnatural amino acids that are quite large, with a recent example being the incorporation of the fluorescent dye BODIPYFL into nicotinic acetylcholine receptors in *Xenopus* oocytes; the labeled receptors could subsequently be imaged as single molecules at the surface of intact oocytes [70^{*}]. (iii) The most elegant and currently most actively explored approach is the generation of pairs of tRNA and aminoacyl-tRNA synthetase that recognize an unnatural amino acid and that are orthogonal to the host cell's protein production machinery. With this approach, proteins incorporating an unnatural amino acid encoded by a free codon can simply be produced by growing cells equipped with the tRNA/aminoacyl-tRNA synthetase pair in the presence of the unnatural amino acid. Using this method, it is now possible to produce proteins containing various unnatural amino acids at defined positions in milligram quantities from bacterial expression systems [4]. Importantly, the approach has now been successfully implemented in mammalian cells [4], promising various previously impossible applications. Nevertheless, the use of UAA incorporation in mammalian cells could profit from further technology development to make it more accessible to non-expert users. Furthermore, the high concentrations of non-incorporated, free unnatural amino acid might be problematic, for example, if fluorescent amino acids are used for live cell imaging. At the moment, the most promising use of this method in cells therefore lies in applications where the free unnatural amino acid does not interfere with downstream applications, such as the photocaging of protein functions, as has recently been demonstrated by photocaging nuclear localization signals in proteins using a lysine bearing a photocleavable protecting group [71^{**}].

Conclusions and future developments

The utility of protein labeling methods to study protein and cell function has now been widely recognized in the scientific community and is increasingly being exploited. The specific introduction of diverse functional synthetic molecules into proteins by tag-mediated labeling expands the utility of methods such as super-resolution imaging and chromophore-assisted light inactivation (CALI) of proteins, and has led to new developments such as sub-cellular targeting of synthetic sensors, the generation of semisynthetic small metabolite sensors, and methods for the study of protein interactions like affinity labeling and S-CROSS. Regarding new labeling methods, the introduction of lipoic acid ligase is a major breakthrough for the intracellular labeling of proteins fused to small tags, as the approach promises to also be applicable *in vivo*. The further engineering of lipoic acid ligase should lead, in the

future, to a broader available palette of transferable labels. With the self-labeling proteins SNAP-tag and HaloTag and further developments of lipoic acid ligase mediated labeling, the toolbox for fluorescent labeling of tagged proteins from *in vitro* to *in vivo* will probably soon be essentially complete.

What is missing? Firstly, unnatural amino acid incorporation by genetic methods is, in principle, the least invasive method to obtain recombinant labeled proteins. The further development of this method in mammalian cells is eagerly awaited. Secondly, a more rigorous investigation of some methodological aspects of labeling seems a worthwhile endeavor. One area that merits a systematic investigation in this respect is the impact of tag size on protein function. In addition, getting a labeling compound of interest across the cell membrane is a recurring problem for intracellular labeling. New general methods that would allow the delivery of arbitrary compounds into the cell interior are therefore urgently required. Alternatively, an improved set of rules that allows the prediction of the membrane permeability of labeling compounds and therefore a better design of the compounds before their actual synthesis would be very valuable. Thirdly, one thing is still missing in the toolbox for protein labeling: general approaches that work on native, non-tagged proteins. While current methods are restricted to recombinant proteins, native protein labeling approaches would allow the study of proteins in native tissue, in their physiological context and at physiological protein concentrations. Thus, they would allow studies that can currently only be carried out with antibody techniques, but provide further possibilities by tagging with a reporter that is covalently bound and much smaller than an antibody.

We currently see three routes towards this aim. One route has become known as activity based protein profiling [72], and relies on the specific covalent interaction of small molecules that react with the active site of enzymes. However, this labeling procedure is invasive as it leads to inactivation of the enzyme of interest. A less invasive route has been described by Hamachi and co-workers and is based – similar to their self-labeling tag approach (see above) – on turning an initially noncovalent interaction into a covalent one. Here, the noncovalent interaction is provided by a molecule that specifically binds to the native protein of interest, for example, an enzyme inhibitor [73^{*}]. The reactive group used is the tosyl group, which is why this approach has been termed ligand-directed tosyl chemistry (LDT). The molecule mediating the interaction is released when the covalent coupling of the label to the protein takes place, which reduces its impact on protein function compared to previous approaches. This method has been shown to allow selective native protein labeling in the cell interior and even in blood cells of living animals [73^{*}]. The method is remarkably specific, but the general reactivity of the tosyl

group with nucleophiles may inherently limit the attainable selectivity towards proteins of low abundance. A third potential route may be provided by enzyme-mediated labeling using engineered enzymes that are specific for native, non-tagged proteins. With ubiquitinylation, Nature provides an illustrating general example for protein labeling that works specifically and with the same chemistry on hundreds of proteins within the cell. Unfortunately, the ubiquitinylation process is rather complex and involves a number of enzymes that work in conjunction. A different enzyme that may be easier to control is transglutaminase, which catalyzes the formation of a covalent bond between the carboxamide side chain of glutamine and primary alkylamines. The enzyme has been shown to have a relaxed substrate specificity and is capable of reacting with various dissimilar peptides [74,75]. We believe that transglutaminases hold great potential for native protein labeling if it should prove possible to increase their specificity towards target proteins of interest by protein engineering.

Acknowledgements

We thank Dr. Arnaud Gautier for interesting discussions and insights on unnatural amino acid incorporation, and Christopher Chidley for valuable comments on the manuscript. M.H. received a postdoctoral grant from the Deutsche Forschungsgemeinschaft under contract number HI-1363/1-1.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
 - of special interest
1. O'Hare HM, Johnsson K, Gautier A: **Chemical probes shed light on protein function.** *Curr Opin Struct Biol* 2007, **17**:488-494.
 2. Lin MZ, Wang L: **Selective labeling of proteins with chemical probes in living cells.** *Physiology* 2008, **23**:131-141.
 3. Sletten EM, Bertozzi CR: **Bioorthogonal chemistry: fishing for selectivity in a sea of functionality.** *Angew Chem Int Ed* 2009, **48**:6974-6998.
 4. Liu CC, Schultz PG: **Adding new chemistries to the genetic code.** *Annu Rev Biochem* 2010.
 5. Das SC, Panda D, Nayak D, Pattnaik AK: **Biarsenical labeling of vesicular stomatitis virus encoding tetracysteine-tagged M protein allows dynamic imaging of M protein and virus uncoating in infected cells.** *J Virol* 2009, **83**:2611-2622.
 6. Andresen M, Schmitz-Salue R, Jakobs S: **Short tetracysteine tags to beta-tubulin demonstrate the significance of small labels for live cell imaging.** *Mol Biol Cell* 2004, **15**:5616-5622.
 7. Enninga J, Mounier J, Sansonetti P, Tran Van Nhieu G: **Secretion of type III effectors into host cells in real time.** *Nat Methods* 2005, **2**:959-965.
 8. Griffin BA, Adams SR, Tsien RY: **Specific covalent labeling of recombinant protein molecules inside live cells.** *Science* 1998, **281**:269-272.
 9. Halo TL, Appelbaum J, Hobert EM, Balkin DM, Schepartz A: **Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid.** *J Am Chem Soc* 2009, **131**:438-439.
 10. Uchinomiya S, Nonaka H, Fujishima S, Tsukiji S, Ojida A, Hamachi I: **Site-specific covalent labeling of His-tag fused proteins with a reactive Ni(II)-NTA probe.** *Chem Commun* 2009:5880-5882.

11. Nonaka H, Fujishima SH, Uchinomiya SH, Ojida A, Hamachi I: **FLAG-tag selective covalent protein labeling via a binding-induced acyl-transfer reaction.** *Bioorg Med Chem Lett* 2009, **19**:6696-6699.
12. Ren HJ, Xiao F, Zhan K, Kim YP, Xie HX, Xia ZY, Rao J: **A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins.** 9658-L 9662. *Angew Chem Int Ed* 2009, **48**.
- The chemical reaction between N-cyanobenzothiazole and N-terminal cysteines is applied for protein labeling on the surface of living cells. This method is noteworthy because it requires just one cysteine as a tag for specific labeling.
13. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohane RF, Urh M *et al.*: **HaloTag: a novel protein labeling technology for cell imaging and protein analysis.** *ACS Chem Biol* 2008, **3**:373-382.
14. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K: **A general method for the covalent labeling of fusion proteins with small molecules in vivo.** *Nat Biotechnol* 2003, **21**:86-89.
15. Keppler A, Arrivoli C, Sironi L, Ellenberg J: **Fluorophores for live cell imaging of AGT fusion proteins across the visible spectrum.** *Biotechniques* 2006, **41**:167-175.
16. Gautier A, Juillerat A, Heinis C, Correa IR, Kindermann M, Beaufils F, Johnsson K: **An engineered protein tag for multiprotein labeling in living cells.** *Chem Biol* 2008, **15**:128-136.
- The self-labeling protein CLIP-tag is an engineered version of SNAP-tag that can also be used to introduce arbitrary labels into tagged proteins. Instead of benzylguanine derivatives that are the substrates of SNAP-tag, CLIP-tag reacts with benzylcytosine derivatives. The tag is orthogonal to SNAP-tag and can be used for simultaneous labeling of proteins fused to SNAP-tag and CLIP-tag in living cells.
17. Mizukami S, Watanabe S, Hori Y, Kikuchi K: **Covalent protein labeling based on noncatalytic beta-lactamase and a designed FRET substrate.** *J Am Chem Soc* 2009, **131**:5016-5017.
18. Miller LW, Cai Y, Sheetz MP, Cornish VW: **In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag.** *Nat Methods* 2005, **2**:255-257.
19. Gallagher SS, Sable JE, Sheetz MP, Cornish VW: **An in vivo covalent TMP-tag based on proximity-induced reactivity.** *ACS Chem Biol* 2009, **4**:547-556.
20. Yin J, Liu F, Li X, Walsh CT: **Labeling proteins with small molecules by site-specific posttranslational modification.** *J Am Chem Soc* 2004, **126**:7754-7755.
21. George N, Pick H, Vogel H, Johnsson N, Johnsson K: **Specific labeling of cell surface proteins with chemically diverse compounds.** *J Am Chem Soc* 2004, **126**:8896-8897.
22. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, Kelleher NL, Kolter R, Walsh CT: **Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase.** *Proc Natl Acad Sci USA* 2005, **102**:15815-15820.
23. Zhou Z, Koglin A, Wang Y, McMahon AP, Walsh CT: **An eight residue fragment of an acyl carrier protein suffices for post-translational introduction of fluorescent pantetheinyl arms in protein modification in vitro and in vivo.** *J Am Chem Soc* 2008, **130**:9925-9930.
- An octapeptide is generated by engineering that represents the so far smallest recognition sequence for AcpS-based labeling.
24. Chen I, Howarth M, Lin W, Ting AY: **Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase.** *Nat Methods* 2005, **2**:99-104.
25. Slavoff SA, Chen I, Choi YA, Ting AAY: **Expanding the substrate tolerance of biotin ligase through exploration of enzymes from diverse species.** *J Am Chem Soc* 2008, **130**:1160-1162.
26. de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, Heck A, Grosveld F, Strouboulis J: **Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice.** *Proc Natl Acad Sci USA* 2003, **100**:7480-7485.
27. Kulman JD, Satake M, Harris JE: **A versatile system for site-specific enzymatic biotinylation and regulated expression of proteins in cultured mammalian cells.** *Protein Expr Purif* 2007, **52**:320-328.
28. Fernandez-Suarez M, Baruah H, Martinez-Hernandez L, Xie KT, Baskin JM, Bertozzi CR, Ting AY: **Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes.** *Nat Biotechnol* 2007, **25**:1483-1487.
- Lipoic acid ligase is introduced as a tool for enzyme-mediated tag labeling. The enzyme is engineered to accept alkyl azide as a substrate that is transferred to an engineered acceptor peptide. The introduced azide serves as a handle to introduce various probes by click chemistry in a second step.
29. Baruah H, Puthenveetil S, Choi YA, Shah S, Ting AY: **An engineered aryl azide ligase for site-specific mapping of protein-protein interactions through photo-cross-linking.** *Angew Chem Int Ed* 2008, **47**:7018-7021.
30. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY: **Yeast display evolution of a kinetically efficient 13-amino acid substrate for lipoic acid ligase.** *J Am Chem Soc* 2009, **131**:16430-16438.
31. Uttamapinant C, White KA, Baruah H, Thompson S, Fernandez-Suarez M, Puthenveetil S, Ting AY: **A fluorophore ligase for site-specific protein labeling inside living cells.** *Proc Natl Acad Sci USA* 2010, **107**:10914-10919.
- This work represents the first example for an enzyme-mediated labeling method that can transfer fluorophores in one step in the interior of living mammalian cells. For this purpose, lipoic acid ligase is reengineered to accept the small hydrophobic fluorophore coumarin as a substrate, and it is shown that the enzyme can be expressed in mammalian cells, enabling intracellular labeling.
32. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL: **Sortagging: a versatile method for protein labeling.** *Nat Chem Biol* 2007, **3**:707-708.
33. Tanaka T, Yamamoto T, Tsukiji S, Nagamune T: **Site-specific protein modification on living cells catalyzed by Sortase.** *Chembiochem* 2008, **9**:802-807.
34. Antos JM, Chew GL, Guimaraes CP, Yoder NC, Grotenbreg GM, Popp MW, Ploegh HL: **Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity.** *J Am Chem Soc* 2009, **131**:10800-10801.
35. Yamamoto T, Nagamune T: **Expansion of the sortase-mediated labeling method for site-specific N-terminal labeling of cell surface proteins on living cells.** *Chem Commun* 2009: 1022-1024.
36. Farr GA, Hull M, Mellman I, Caplan MJ: **Membrane proteins follow multiple pathways to the basolateral cell surface in polarized epithelial cells.** *J Cell Biol* 2009, **186**:269-282.
- This work is an illustrating example how self-labeling protein tags can be applied to study protein trafficking in mammalian cells by pulse-chase methods. Notably, it shows how distinct protein pools can be synchronized without requiring protein mutants that display a temperature-sensitive trafficking behavior.
37. Neugart F, Zappe A, Buk DM, Ziegler I, Steinert S, Schumacher M, Schopf E, Bessey R, Wurster K, Tietz C *et al.*: **Detection of ligand-induced CNTF receptor dimers in living cells by fluorescence cross correlation spectroscopy.** *Biochim Biophys Acta* 2009, **1788**:1890-1900.
38. Copeland MF, Flickinger ST, Tuson HH, Weibel DB: **Studying the dynamics of flagella in multicellular communities of Escherichia coli by using biarsenical dyes.** *Appl Environ Microbiol* 2010, **76**:1241-1250.
39. Kropf M, Rey G, Glauser L, Kulangara K, Johnsson K, Hirling H: **Subunit-specific surface mobility of differentially labeled AMPA receptor subunits.** *Eur J Cell Biol* 2008, **87**: 763-778.
40. Milenkovic L, Scott MP, Rohatgi R: **Lateral transport of Smoothed from the plasma membrane to the membrane of the cilium.** *J Cell Biol* 2009, **187**:365-374.

41. Wang Y, Zhou Z, Walsh CT, McMahon AP: **Selective translocation of intracellular Smoothed to the primary cilium in response to Hedgehog pathway modulation.** *Proc Natl Acad Sci USA* 2009, **106**:2623-2628.
42. Maurel D, Comps-Agrar L, Brock C, Rives ML, Bourrier E, Ayoub MA, Bazin H, Tinel N, Durroux T, Prezeau L *et al.*: **Cell-surface protein-protein interaction analysis with time-resolved FRET and snap-tag technologies: application to GPCR oligomerization.** *Nat Methods* 2008, **5**:561-567.
- The self-labeling protein SNAP-tag is applied to study G protein-coupled receptor (GPCR) interactions on the surface of mammalian cells. GPCRs are expressed as SNAP-tag fusion proteins and labeled with fluorophores that are compatible with time-resolved fluorescence resonance energy transfer. This leads to the observation that different classes of GPCRs form various dimeric or oligomeric structures on the surface of mammalian cells.
43. Kosaka N, Ogawa M, Choyke PL, Karassina N, Corona C, McDougall M, Lynch DT, Hoyt CC, Levenson RM, Los GV *et al.*: **In vivo stable tumor-specific painting in various colors using dehalogenase-based protein-tag fluorescent ligands.** *Bioconjug Chem* 2009, **20**:1367-1374.
- A tumor expressing a HaloTag fusion protein at its cell surface is grafted into the peritoneal cavity of living mice. By direct injection of diverse fluorescent HaloTag substrates, the tumors can be fluorescently labeled and imaged noninvasively in living mice.
44. Marek KW, Davis GW: **Transgenically encoded protein photoinactivation (FIAsH-FALI): acute inactivation of synaptotagmin I.** *Neuron* 2002, **36**:805-813.
45. Tour O, Meijer RM, Zacharias DA, Adams SR, Tsien RY: **Genetically targeted chromophore-assisted light inactivation.** *Nat Biotechnol* 2003, **21**:1505-1508.
46. Heerssen H, Fetter RD, Davis GW: **Clathrin dependence of synaptic-vesicle formation at the *Drosophila* neuromuscular junction.** *Curr Biol* 2008, **18**:401-409.
47. Keppler A, Ellenberg J: **Chromophore-assisted laser inactivation of alpha- and gamma-tubulin SNAP-tag fusion proteins inside living cells.** *ACS Chem Biol* 2009, **4**:127-138.
- Alpha-tubulin and gamma-tubulin SNAP-tag fusions are labeled with fluorescein in living cells. After five minutes labeling and 1-15 s irradiation with high-intensity laser light, the proteins are inactivated owing to local production of singlet oxygen. The impact of protein inactivation on cell structures can be directly followed in living cells by time-lapse fluorescence microscopy.
48. Schroder J, Benink H, Dyba M, Los GV: **In vivo labeling method using a genetic construct for nanoscale resolution microscopy.** *Biophys J* 2009, **96**:L01-03.
49. Hein B, Willig KI, Wurm CA, Westphal V, Jakobs S, Hell SW: **Stimulated emission depletion microscopy of living cells using SNAP-tag fusion proteins.** *Biophys J* 2010, **98**:158-163.
- SNAP-tag fusion proteins of cytoskeletal and membrane proteins are overexpressed in mammalian cells and labeled with tetramethylrhodamine. Using the STED super-resolution imaging approach, the labeled proteins and the structures in which they were engaged can be imaged with a resolution down to 40 nm in the living cells.
50. Dellagiacoma C, Lukinavicius G, Bocchio N, Banala S, Geissbuhler S, Marki I, Johnsson K, Lasser T: **Targeted photoswitchable probe for nanoscopy of biological structures.** *ChemBiochem* 2010, **11**:1361-1363.
51. Maurel D, Banala S, Laroche T, Johnsson K: **Photoactivatable and photoconvertible fluorescent probes for protein labeling.** *ACS Chem Biol* 2010, **5**:507-516.
- A general strategy is presented for generating photoactivatable and photoconvertible fluorescent labels that can be selectively coupled to SNAP-tag fusion proteins. The labeling compounds can be applied to localized pulse-chase studies of tagged proteins in living cells. In this study, they are utilized to probe the lateral mobility of membrane proteins.
52. Liu B, Archer CT, Burdine L, Gillette TG, Kodadek T: **Label transfer chemistry for the characterization of protein-protein interactions.** *J Am Chem Soc* 2007, **129**:12348-12349.
53. Archer CT, Burdine L, Liu B, Ferdous A, Johnston SA, Kodadek T: **Physical and functional interactions of monoubiquitylated transactivators with the proteasome.** *J Biol Chem* 2008, **283**:21789-21798.
54. Fernandez-Suarez M, Chen TS, Ting AY: **Protein-protein interaction detection in vitro and in cells by proximity biotinylation.** *J Am Chem Soc* 2008, **130**:9251-9253.
55. Luedtke NW, Dexter RJ, Fried DB, Schepartz A: **Surveying polypeptide and protein domain conformation and association with FIAsH and ReAsH.** *Nat Chem Biol* 2007, **3**:779-784.
56. Krishnan B, Gierasch LM: **Cross-strand split tetra-cys motifs as structure sensors in a beta-sheet protein.** *Chem Biol* 2008, **59**:1104-1115.
57. Goodman JL, Fried DB, Schepartz A: **Bipartite tetracysteine display requires site flexibility for ReAsH coordination.** *ChemBiochem* 2009, **10**:1644-1647.
58. Lemercier G, Gendreizig S, Kindermann M, Johnsson K: **Inducing and sensing protein-protein interactions in living cells by selective cross-linking.** *Angew Chem Int Ed* 2007, **46**:4281-4284.
59. Gautier A, Nakata E, Lukinavicius G, Tan KT, Johnsson K: **Selective cross-linking of interacting proteins using self-labeling tags.** *J Am Chem Soc* 2009, **131**:17954-17962.
- A new method, termed S-CROSS, allows the investigation of protein interactions. For this purpose, the proteins of interest are fused to SNAP-tag and/or CLIP-tag, respectively, and incubated with a trifunctional molecule containing the substrates of SNAP-tag and CLIP-tag as well as a fluorescent dye. Crosslinking of the proteins of interest is only efficient when the two proteins interact, which can be detected by a gel-shift in SDS-PAGE after in-gel fluorescence scanning.
60. Brun MA, Tan KT, Nakata E, Hinner MJ, Johnsson K: **Semisynthetic fluorescent sensor proteins based on self-labeling protein tags.** *J Am Chem Soc* 2009, **131**:5873-5884.
- This publication presents a novel metabolite sensor design based on the protein human carbonic anhydrase and the self-labeling proteins SNAP-tag and CLIP-tag. The tags allow the introduction of synthetic fluorophores and an intramolecular inhibitor for human carbonic anhydrase, and ratiometric fluorescence changes of up to 100% are achieved using various fluorophore FRET pairs. The proof-of-principle with human carbonic anhydrase shows that the sensor design works without the requirement of a conformational change in the binding protein.
61. Fehr M, Frommer WB, Lalonde S: **Visualization of maltose uptake in living yeast cells by fluorescent nanosensors.** *Proc Natl Acad Sci USA* 2002, **99**:9846-9851.
62. Hires SA, Zhu Y, Tsien RY: **Optical measurement of synaptic glutamate spillover and reuptake by linker optimized glutamate-sensitive fluorescent reporters.** *Proc Natl Acad Sci USA* 2008, **105**:4411-4416.
63. Tomat E, Nolan EM, Jaworski J, Lippard SJ: **Organelle-specific zinc detection using zinpyr-labeled fusion proteins in live cells.** *J Am Chem Soc* 2008, **130**:15776-15777.
64. Bannwarth M, Correa IR, Sztretze M, Pouvreau S, Fellay C, Aebischer A, Royer L, Rios E, Johnsson K: **Indo-1 derivatives for local calcium sensing.** *ACS Chem Biol* 2009, **4**:179-190.
65. Ruggiu AA, Bannwarth M, Johnsson K: **Fura-2FF-based calcium indicator for protein labeling.** *Org Biomol Chem* 2010, **8**:3398-3401.
66. Kamiya M, Johnsson K: **A localizable and highly sensitive calcium indicator based on a BODIPY fluorophore.** *Anal Chem* 2010, **82**:6472-6479.
67. Srikun D, Albers AE, Nam CI, Iavaron AT, Chang CJ: **Organelle-targetable fluorescent probes for imaging hydrogen peroxide in living cells via SNAP-tag protein labeling.** *J Am Chem Soc* 2010, **132**:4455-4465.
68. Benink HA, McDougall MG, Klaubert DH, Los GV: **Direct pH measurements by using subcellular targeting of 5(and 6)-carboxysemaphthorhodafuran in mammalian cells.** *Biotechniques* 2009, **47**:769-774.
69. Lepthien S, Hoesl MG, Merkel L, Budisa N: **Azatriptophans endow proteins with intrinsic blue fluorescence.** *Proc Natl Acad Sci USA* 2008, **105**:16095-16100.
- (4-aza)indole is introduced as a fluorescent isostere of indole, which is metabolically converted within bacterial cells to (4-aza)tryptophan. In

bacteria that are auxotrophic for tryptophan, the 4-aza isostere is efficiently incorporated into the proteins produced by the cell. This leads to the production of intrinsically blue fluorescent proteins, as demonstrated with the human protein annexin A5.

70. Pantoja R, Rodriguez EA, Dibas MI, Dougherty DA, Lester HA:
 • **Single-molecule imaging of a fluorescent unnatural amino acid incorporated into nicotinic receptors.** *Biophys J* 2009, **96**:226-237.

A frameshift suppressor tRNA amino-acylated with the unnatural fluorescent amino acid Lys(BODIPYFL) is successfully incorporated into the muscle nicotinic acetylcholine receptor in oocytes. Single fluorescent channels that are still fully functional can subsequently be imaged in the oocyte plasma membrane using total internal reflection fluorescence microscopy.

71. Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW:
 •• **Genetically encoded photocontrol of protein localization in mammalian cells.** *J Am Chem Soc* 2010, **132**:4086-4088.

A pyrrolysyl-tRNA synthetase/tRNA pair is evolved to allow the incorporation of lysine bearing a photocleavable protecting group into the nuclear localization sequence of a protein. The nuclear import of the protein – at first localized to the cytoplasm – is triggered by light, which allows following the kinetics of nuclear import.

72. Barglow KT, Cravatt BF: **Activity-based protein profiling for the functional annotation of enzymes.** *Nat Methods* 2007, **4**:822-827.

73. Tsukiji S, Miyagawa M, Takaoka Y, Tamura T, Hamachi I:
 • **Ligand-directed tosyl chemistry for protein labeling in vivo.** *Nat Chem Biol* 2009, **5**:341-343.

A native protein labeling method is introduced that is based on trifunctional molecules containing a moiety that mediates binding to the protein of interest, the reactive tosyl group, and a label. The authors show a selective labeling of human carbonic anhydrase *in vitro* and in the blood

cells of living animals. Selective labeling is also demonstrated with the native proteins FKBP12 and congerin.

74. Keresztessy Z, Csoos E, Harsfalvi J, Csomos K, Gray J, Lightowlers RN, Lakey JH, Balajthy Z, Fesus L: **Phage display selection of efficient glutamine-donor substrate peptides for transglutaminase 2.** *Protein Sci* 2006, **15**:2466-2480.
75. Sugimura Y, Hosono M, Wada F, Yoshimura T, Maki M, Hitomi K: **Screening for the preferred substrate sequence of transglutaminase using a phage-displayed peptide library: identification of peptide substrates for TGASE 2 and Factor XIIIa.** *J Biol Chem* 2006, **281**:17699-17706.
76. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, Llopis J, Tsien RY: **New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications.** *J Am Chem Soc* 2002, **124**:6063-6076.
77. Zhou Z, Cironi P, Lin AJ, Xu YQ, Hrvatin S, Golan DE, Silver PA, Walsh CT, Yin J: **Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases.** *ACS Chem Biol* 2007, **2**:337-346.
78. Howarth M, Ting AY: **Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin.** *Nat Protoc* 2008, **3**:534-545.
79. Popp MW, Antos JM, Ploegh HL: **Site-specific protein labeling via sortase-mediated transpeptidation.** *Curr Protoc Protein Sci* 2009, **Chapter 15**:Unit 15.13.
80. Vogelsang J, Steinhauer C, Forthmann C, Stein IH, Person-Skegro B, Cordes T, Tinnefeld P: **Make them blink: probes for super-resolution microscopy.** *Chemphyschem* 2010, **11**:2475-2490.