sensors

Review

Selective Chemical Labeling of Proteins with Small Fluorescent Molecules Based on Metal-Chelation Methodology

Nobuaki Soh

Department of Applied Chemistry, Graduate School of Engineering, Kyushu University,
744 Moto-oka, Nishi-ku, Fukuoka 819-0395, Japan. E-mail: soh@cstf.kyushu-u.ac.jp;
Phone/Fax: +81-92-802-2891

Abstract: Site-specific chemical labeling utilizing small fluorescent molecules is a powerful and attractive technique for *in vivo* and *in vitro* analysis of cellular proteins, which can circumvent some problems in genetic encoding labeling by large fluorescent proteins. In particular, affinity labeling based on metal-chelation, advantageous due to the high selectivity/simplicity and the small tag-size, is promising, as well as enzymatic covalent labeling, thereby a variety of novel methods have been studied in recent years. This review describes the advances in chemical labeling of proteins, especially highlighting the metal-chelation methodology.

1. Introduction

Protein labeling with fluorescent molecules that allows the sensing and visualizing of protein dynamics, localization, and protein-protein interactions, is an invaluable technique to understand protein functions and networks in living cells. Genetic encoding of fluorescent proteins (FPs) such as green fluorescent protein (GFP) is the most popular technique for protein labeling because of some important advantages, including high labeling specificity and simplicity [1-3]. Unfortunately, however, the molecular sizes of FPs (i.e. 27 kDa for GFP) are large enough to potentially interfere with the structure or function of the protein to which they are fused [4, 5], resulting in recent efforts to develop new techniques concerning selective labeling of protein based on smaller chemical compounds.

Selective chemical labeling is generally achieved by site-specific binding between a tag fused with a protein of interest (POI) and a small fluorescent compound, although some different approaches, such

as incorporation of unnatural amino acids based on suppressor tRNA technology [6], have also been reported. In the general chemical labeling, a protein or peptide tag, whose size is small relative to the FPs, is utilized in order to reduce the potential interferences with the POI. From a binding mechanisms standpoint, most chemical labeling techniques can be classified into major two categories; enzymatic labeling and affinity labeling. The former includes the acyl carrier protein/phosphopantetheine transferase (ACP/PPTase) [7], Q-tag/transglutaminase (TGase) [8], biotin acceptor peptide/biotin ligase (AP/Bir A) [9], farnesylation motif/protein farnesyltransferase (PFTase) [10], aldehyde tag/formylglycine-generating enzyme [11], human O^6-alkylguanine transferase (hAGT) [12, 13], and mutated prokaryotic dehalogenase (HaloTag™) methods [14]. On the other hand, the latter includes noncovalent methods utilizing dihydrofolate reductase (DHFR) [15, 16] and Phe36Val mutant of FK506-binding protein 12 (FKBP12(F36V)) [17], as well as metal-chelation methods.

Protein labeling based on a specific chelation has some remarkable features, including (i) simplicity in labeling procedures, (ii) high selective and stable labeling based on metal-chelation, and (iii) applicability to various site-specific labeling (N-termini, C-termini, and internal sites). Due to these important advantages, metal-chelation labeling represents one of the most powerful and attractive methods in protein labeling, as well as enzymatic labeling [18-22]. In this review, the strategies and recent advances in fluorescent labeling of proteins were described, especially focusing on the metal-chelation methodology [Scheme 1(a)-(d)].

2. Tetracysteine-tag system

2.1. Pioneering of the tetracysteine/biarsenical system

The tetracysteine/biarsenical system was reported as the prototype for the specific fluorescent chemical labeling based on metal-chelation by Tsien's group in 1998 [23]. Biarsenical dyes such as 4',5'-bis(1,3,2-dithioarsolan-2-yl)fluorescein (FlAsH, also denoted as FlAsH-EDT$_2$, Figure 1(a)) bind specifically with a short peptide sequence called the tetracysteine motif (CCPGCC, some important improvements have also been accomplished, see section 2.3) with high affinity [23, 24]. Thus, POI that is expressed with the motif can be site-specifically labeled with the small biarsenical dyes [Scheme 1(a)]. An excess of 1,2-ethanedithiol (EDT) is simultaneously introduced with the fluorescent dye in order to minimize non-specific binding and toxicity. Notably, the biarsenical dyes are membrane-permeable and do not require complicated procedures such as microinjection. Furthermore, the biarsenical dyes are non-fluorescent until they bind to form covalent complexes with the corresponding motif, reducing a background noise in fluorescent measurements. After the development of green-fluorescent FlAsH ($\lambda_{ex} = 508$ nm, $\lambda_{em} = 528$ nm), several derivatives including red-fluorescent ReAsH ($\lambda_{ex} = 593$ nm, $\lambda_{em} = 608$ nm, Figure 1(b)) and blue-fluorescent CHoXAsH ($\lambda_{ex} = 380$ nm, $\lambda_{em} = 430$ nm, Figure 1(c)) were also reported [24]. The simultaneous use of such biarsenical dyes extends the availability of the functional molecules in biological applications [25].

Meanwhile, some limitations in the biarsenical system have also been pointed out. For example, a considerable affinity for monothiols requires troublesome wash processes, furthermore, the background

(a)

(b)

(c)

(d)
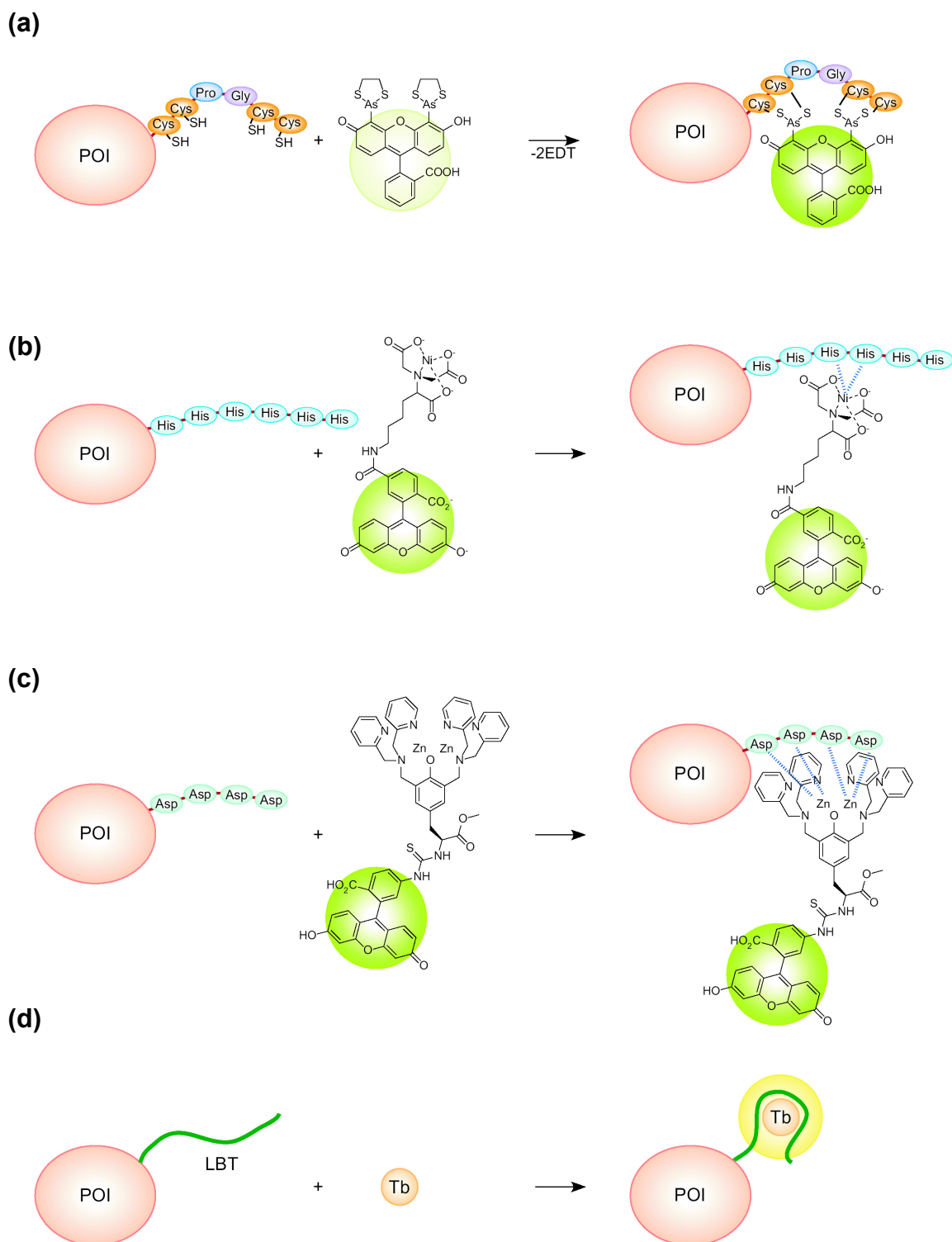
Scheme 1. Strategies for the selective chemical labeling of proteins with small fluorescent molecules based on metal-chelation: (a) tetracysteine/biarsenical system, (b) oligohistidine/nickel-complex system, (c) oligo-aspartate/zinc-complex system, (d) lanthanide-binding tag system
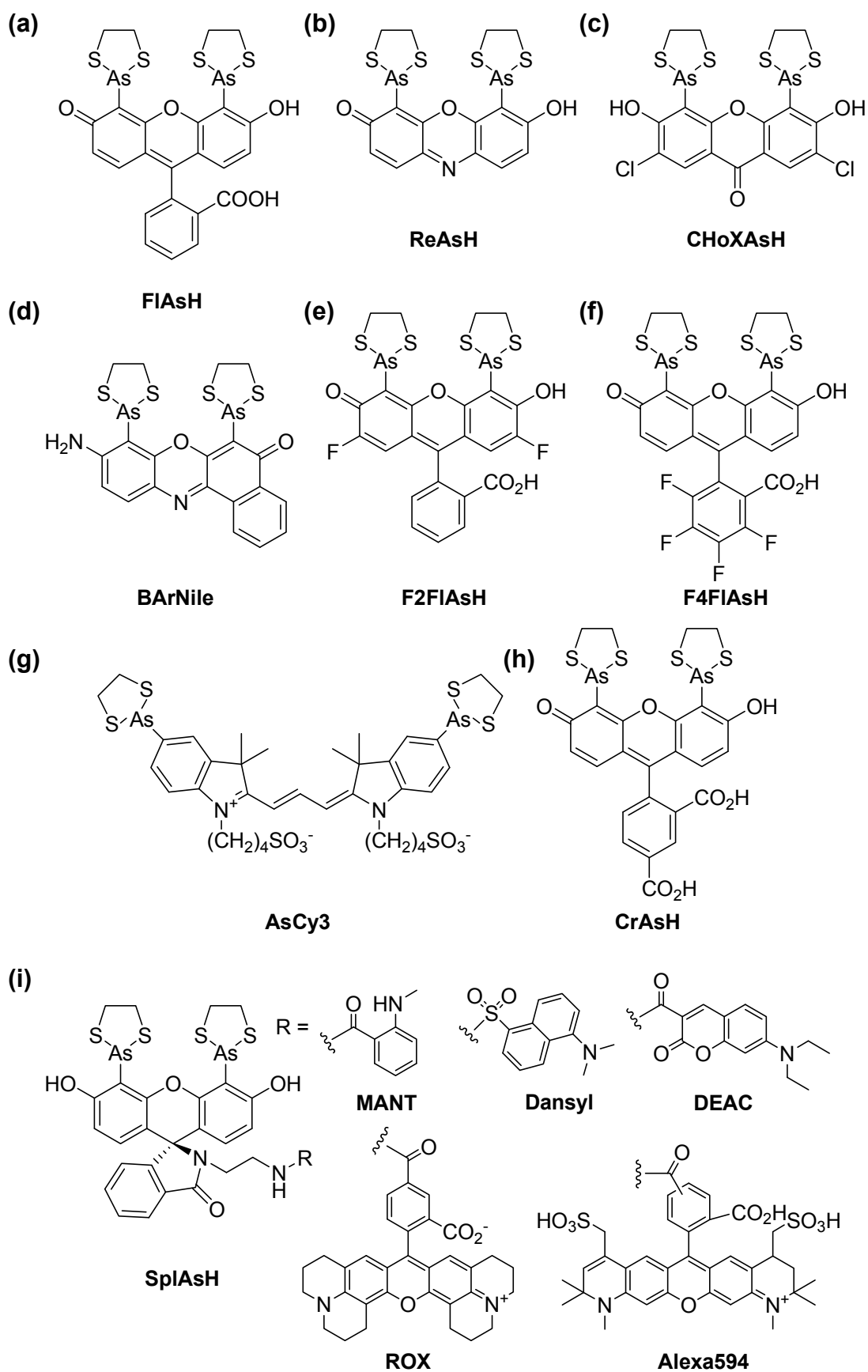
Figure 1. Biarsenical dyes: (a) FlAsH, (b) ReAsH, (c) CHoXAsH, (d) BArNile, (e) F2FlAsH, (f) F4FlAsH, (g) AsCy3, (h) CrAsH, (i) SplAsHs (SplAsH-MANT, SplAsH-Dansyl, SplAsH-DEAC, SplAsH-ROX, SplAsH-Alexa594)

fluorescence is not fully eliminated in spite of an excess addition of EDT and the extensive washing in some cases [26]. In oxidizing environments, a specific labeling would be difficult since the reduced form of tetracysteine motif seems to be easily converted into the oxidized form. Yet, the tetracysteine/biarsenical system is the most mature labeling technique with outstanding features, and therefore, has already contributed to studies concerning cellular biology [27-30].

2.2. Various biarsenical dyes

Since the first report on FlAsH dye by Tsien *et al.*, a variety of biarsenical dyes have been reported by other groups. Umezawa *et al.* synthesized a nile red-based biarsenical dye, 9-amino-6,8-bis(1,3,2-dithioarsolan-2-yl)-5H-benzo[*a*]phenoxazin-5-one, called BArNile-EDT$_2$ (*bisar*senical *nile* red analogue, bis-EDT adduct, Figure 1(d)) [31]. A 9-amino analogue of nile red was selected in order to avoid potential interference in tag recognition. Since the 9-amino analogue of nile red is environment-sensitive like nile red, the fluorescence spectra of BArNile would be highly influenced by changes in its molecular environment. Based on this expectation, the group applied the probe to tetracysteine-fused calmodulin (CaM), which is a Ca^{2+}-binding protein that exposes hydrophobic domains depending on the Ca^{2+} concentration, and successfully imaged the conformational change upon Ca^{2+} increase in living cells.

In order to improve the fluorescent character of FlAsH, Jares-Erijman *et al.* developed two FlAsH derivatives, difluorinated F2FlAsH [Figure 1(e)] and tetrafluorinated F4FlAsH [Figure 1(f)] [32]. In comparison to FlAsH, F2FlAsH shows higher absorbance, larger Stokes shift, higher quantum yield, higher photostability (50-fold), and reduced pH dependence. On the other hand, the emission of F4FlAsH exists in a region intermediate to that of FlAsH and ReAsH, allowing a new color and excellent luminosity. These two derivatives can work as an excellent pair for fluorescence resonance energy transfer (FRET) (F2FlAsH; λ_{abs} = 500 nm, λ_{em} = 522 nm, F4FlAsH; λ_{abs} = 528 nm, λ_{em} = 544 nm) with a large R_0 value (R_0; Förster critical distance for 50 % FRET efficiency).

In FlAsH and their derivatives, the conserved interatomic distance (~6 Å) between the two arsenic moieties complicates the selective labeling of multiple proteins with different reporters. To conquer the difficulties, Mayer *et al.* designed and synthesized a new Cy3-based biarsenical dye, AsCy3 [Figure 1(g)] [33]. The large interatomic distance between the two arsenics in AsCy3 (~14.5 Å) enables a high-affinity binding (K_d = 80 ± 10 nM, in the presence of 100 μM EDT) with Cy3TAG (CCKAEAACC). The absorbance spectrum is insensitive to the binding, while the fluorescence spectrum (λ_{em} = 568 nm) is red-shifted (λ_{em} = 576 nm) with a 6-fold increase in the fluorescence quantum yield (0.28). Compared to FlAsH and ReAsH, AsCy3 shows superior photostability and a minimal environmental sensitivity. AsCy3 provides a FRET partner to FlAsH because the fluorescence spectrum of FlAsH overlaps with the absorption spectrum of AsCy3. In the report, the utility of FlAsH/AsCy3 pair for monitoring protein-protein interactions was experimentally demonstrated utilizing CaM and CaM-binding peptide.

A non-specific binding to membranes and hydrophobic pockets of proteins, also increases the fluorescence of FlAsH. Although such background fluorescence can be decreased by prebinding with other dyes, an appropriate modification that renders FlAsH less hydrophobic would also reduce this undesirable binding. From this point of view, Mayer *et al.* selected carboxy-FlAsH (CrAsH, Figure

1(h)), which had formerly been synthesized by Tsien's group [24], and investigated the performance of the dye in terms of the specific binding [34]. As a result, it was confirmed that CrAsH exhibits lower non-specific binding to a hydrophobic protein than FlAsH although the affinity of CrAsH with tetracysteine sequence ($K_d = 407 \pm 11$ nM) is one order of magnitude weaker than FlAsH.

Color variations of biarsenical dyes are expected to promote the widespread use of the powerful labeling techinique based on tetracysteine/biarsenical system. Although several biarsenical dyes were reported, simultaneous structural requirements for both fluorescence and the rigid display of arsenic atoms strictly limit the range of fluorophores. Furthermore, the scope of compatible dyes is both narrow and difficult to predict, as found from the report that biarsenical rhodamines are nonfluorescent, even when bound to a tetracysteine motif [24]. Miller *et al.* developed a general strategy that overcomes the problems so as to show the recruitment of any fluorophore to a tetracysteine motif [35]. The strategy is based on a modular approach, wherein the biarsenical targeting moiety is separated from the fluorophore in order to remove any restriction on its structure. The fluorescein-based targeting moiety called SplAsH (*Spiro*lactam *Ar*senical *H*airpin binder, Figure 1(i)), which maintains the planarity of the xanthene ring and shows no inherent fluorescence, is conveniently attached with a variety of fluorophores. Based on this strategy, various SplAsH dyes, including SplAsH-MANT (MANT; *N*-methylanthranilate), SplAsH-Dansyl, SplAsH-DEAC (DEAC; diethylamino-coumarin), SplAsH-ROX (X-rhodamine), and SplAsH-Alexa594 were synthesized. Such SplAsH dyes are still fluorescent in the absence of tetracysteine motif, whereas the fluorescence of unbound FlAsH is strongly quenched. The advantage that any molecule such as fluorescent sensors and caged fluorophores can be attached to SplAsH platform by the same strategy is also attractive.

2.3. Improvement of tetracysteine tag

In the earliest design, Tsien *et al.* used α-helical tetracysteine tag, in which the four cysteines were located at the *i*, *i*+1, *i*+4, and *i*+5 positions of the helix, so that the four thiol groups would form a parallelogram on one side of the helix and thus geometrically fit the biarsenical dye (WEAAAREACCRECCARA was used as a model peptide for *in vitro* study) [23]. The group subsequently reported that CCPGCC peptide (called PG-tag, for simplicity here), in which helix-breaking proline and glycine between the dicysteine pairs were inserted, enhanced the affinity, suggesting that the preferred peptide conformation is a hairpin rather than α-helix [24]. To find more sophisticated peptide sequences, they utilized mammalian cell-based library afterward [36]. In the approach, residues surrounding the tetracysteine motif were randomized and fused to GFP, retrovirally transduced into cells and iteratively sorted using fluorescence-activated cell sorting (FACS) for high FRET from GFP to ReAsH with increasing dithiols. The selected sequences, HRWCCPGCCKTF and FLNCCPGCCMEP, exhibited higher fluorescence quantum yield and excellent dithiol resistance.

Basically, the application of biarsenical dyes is limited by the availability of a single PG-tag, inhibiting the differential labeling of coexpressed proteins. To conquer the drawback, Squier *et al.* used a peptide library so as to identify an alternate binding tag (CCKACC, KA-tag) [37]. The KA-tag shows a similar brightness to PG-tag upon complex formation, but exhibits different rates and affinities that allow the differential labeling by FlAsH and ReAsH. Utilizing these tags simultaneously, a specific labeling of two subunits of RNA polymerase in cellular lysates was accomplished.

In the strategies for searching short hairpin tetracysteine tags described so far, the overall design of the tags was fixed, namely, the separation between the vicinal cysteine pairs was constrained to two residues. Mayer *et al.* proposed a new strategy to identify optimal binding tags for biarsenical dyes based on a proteomic approach [38]. After finding fluorescent bands on SDS-PAGE gels of FlAsH-labeled *Eschrerichia coli* lysate, other than those arising from the tetracysteine-tagged protein, a prospecting approach was used to find improved protein tags. Based on this approach, two naturally occurring tags, (GCCGGSGNDAGGCCGG, S_peptide) and (GCCGGHGHDHGHEHGGEGCCGG, E_peptide), having better properties than those of existing designed tag sequences, were identified from *Shewanella oneidensis* MR-1 and *E.coli* K-12, respectively. Both peptides have substantially the same binding affinities for FlAsH as existing PG-tags, however, S_peptide was concluded to be more promising because it was essentially brighter in complex formation and did not form intermolecular cross-linking that limits the current biological application of PG-tag.

2.4. New generation of biarsenicals with unique functions

Recent advances in biarsenical dyes led to the appearance of a new type of biarsenical for further applications. Tsien *et al.* developed Calcium Green FlAsH (CaGF, Figure 2), which comprises a FlAsH moiety and a BAPTA-like chelator (BAPTA; 1,2-bis(*o*-aminophenoxy)ethane-*N,N,N',N'*-tetraacetic acid), for precisely localized fast detection of intracellular Ca^{2+} [39]. The selective binding of CaGF with Ca^{2+} results in 10-fold more fluorescence increase with a K_d of ~100 μM, < 1-ms kinetics, and good Mg^{2+} rejection. The Ca^{2+} waves through gap junctions were monitored in real time using CaGF in Hela cells expressing tetracysteine-tagged connexin 43. Total internal reflection microscopy (TIRF-M) of CaGF-labeled calcium channels displayed fast-rising depolarization-induced Ca^{2+} transient. While these transient decayed slowly with moderate Ca^{2+} buffering, the CaGF signals decayed as rapidly as the calcium currents with high Ca^{2+} buffering, indicating that CaGF can monitor highly localized and rapid Ca^{2+} dynamics.
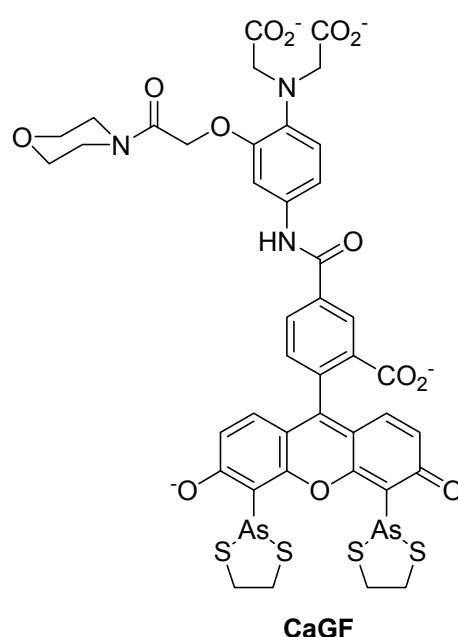

CaGF

Figure 2. Biarsenical Ca^{2+} indicator (CaGF)

3. His-tag system

3.1. Development and applications of oligohistidine/nickel-complex system

An oligohistidine sequence ((His)$_n$, $n \geq 6$, in general) called His-tag is known to interact robustly with transition-metal complexes, including nitrilotriacetic acid (NTA) complex of Ni^{2+}, thereby the sequence is widely used for purification of expressed proteins by affinity chromatography. The selective interaction between His-tag and the metal complexes is also applicable for site-specific fluorescent labeling of proteins. This strategy has important and unique advantages, including the compatibility to the large library of existing His-tagged proteins.

In 2001, Ebright *et al.* reported derivatives of Cy3 ($\lambda_{ex} \sim 550$ nm, $\lambda_{em} \sim 565$ nm) and Cy5 ($\lambda_{ex} \sim 650$ nm, $\lambda_{em} \sim 670$ nm), having one and two pendant of NTA-Ni^{2+} for specific labeling of His-tagged proteins [40]. Two NTA-Ni^{2+} pendant derivatives, (Ni^{2+}:NTA)$_2$-Cy3 and (Ni^{2+}:NTA)$_2$-Cy5 [Figure 3(a)], having higher affinity than one pendant derivatives, were separately applied to FRET-based distance measurements for fluorescein-labeled, His-tagged DNA-protein complex. At the same period, but individually, Katayama *et al.* developed NTA-FITC ($\lambda_{ex} \sim 495$ nm, $\lambda_{em} \sim 520$ nm, Figure 3(b)) and applied its Ni^{2+} complex to the detection of His-tagged protein on a nitrocellulose membrane [41].

His-tag/NTA system is also useful for *in vivo* labeling. Vogel *et al.* investigated the structure of cell-surface 5HT$_3$ serotonin receptor bearing His-tag, based on FRET measurements utilizing NTA-Ni^{2+}-chromophores [Figure 3(c)] [42]. Three chromophores, namely, QSY7 and QSY9 having large extinction coefficients, and fluorogenic tetraethyl rhodamine, were selected, and these chromophore-appended NTA-Ni^{2+} complexes (NTA-I, NTA-II, and NTA-Rho) were used as FRET acceptors. In the experiments, His$_{10}$-sequence bearing higher affinity than His$_6$-sequence was also utilized.

The use of paramagnetic Ni^{2+}, as a binding medium between fluorophore-NTA and His-tag, shows a decrease in the inherent fluorescence of the fluorophore, leading to the limitation of the fluorescent labeling based on Ni^{2+}-chelation technique for biological applications. Lippard *et al.* reported that the fluorescence of NTA-DCF (DCF; 2',7'-dichlorofluorescein, Figure 3(d)) is not significantly influenced by a chelation with Ni^{2+} (NTA-DCF; $\Phi = 0.78$, NTA-DCF-Ni^{2+} complex; $\Phi = 0.72$) [43]. Using the complex, extracellular proteins bearing His-tag was selectively labeled.

3.2. Multivalent chelator system

One of the most serious limitations in His-tag/NTA-Ni^{2+} system is the relatively low affinity (K_d values generally lie in the 1-20 μM range) between His-tag and NTA-Ni^{2+} complex, and this nature probably disturbs stable labeling and imaging. As expected from the results shown in Cy-NTA-Ni^{2+} complexes [40], the number of NTA-Ni^{2+} moiety has a significant effect on the entire stability between a fluorescent-NTA-Ni^{2+} complex and His-tag. Piehler *et al.* synthesized supramolecular entities containing 1-4 NTA moieties and a fluorescein (mono-NTA-Fluo, bis-NTA-Fluo, tris-NTA-Fluo, and tetrakis-NTA-Fluo, Figure 4(a)-(d)), and characterized their interaction with His-tag (His$_6$ and His$_{10}$-sequences) [44]. Analytical size exclusion chromatography showed that the binding stability increases substantially with increasing number of NTA moieties. An increase in stability of the multivalent chelator heads/His-tag complexes by four orders of magnitude compared to that of mono-NTA was
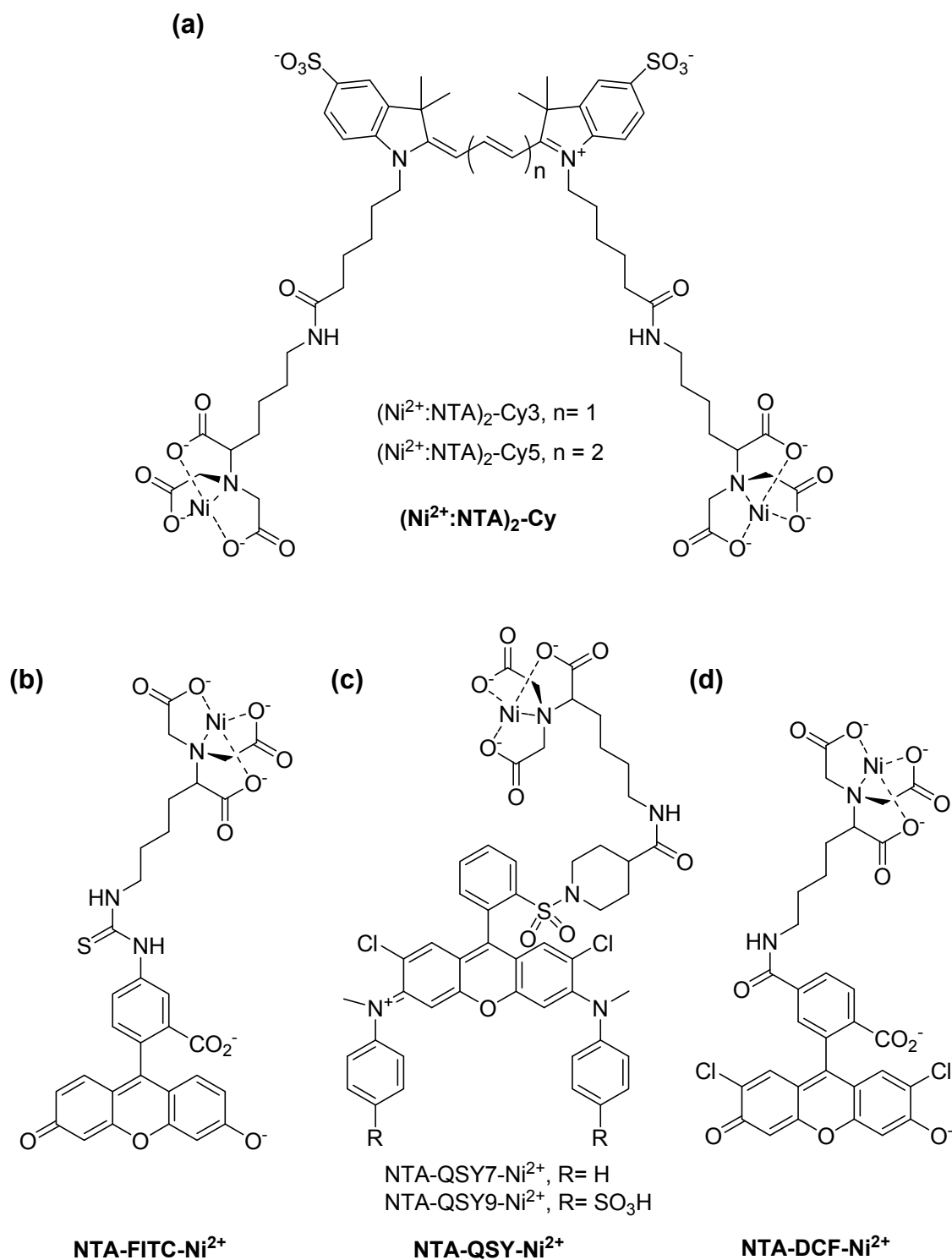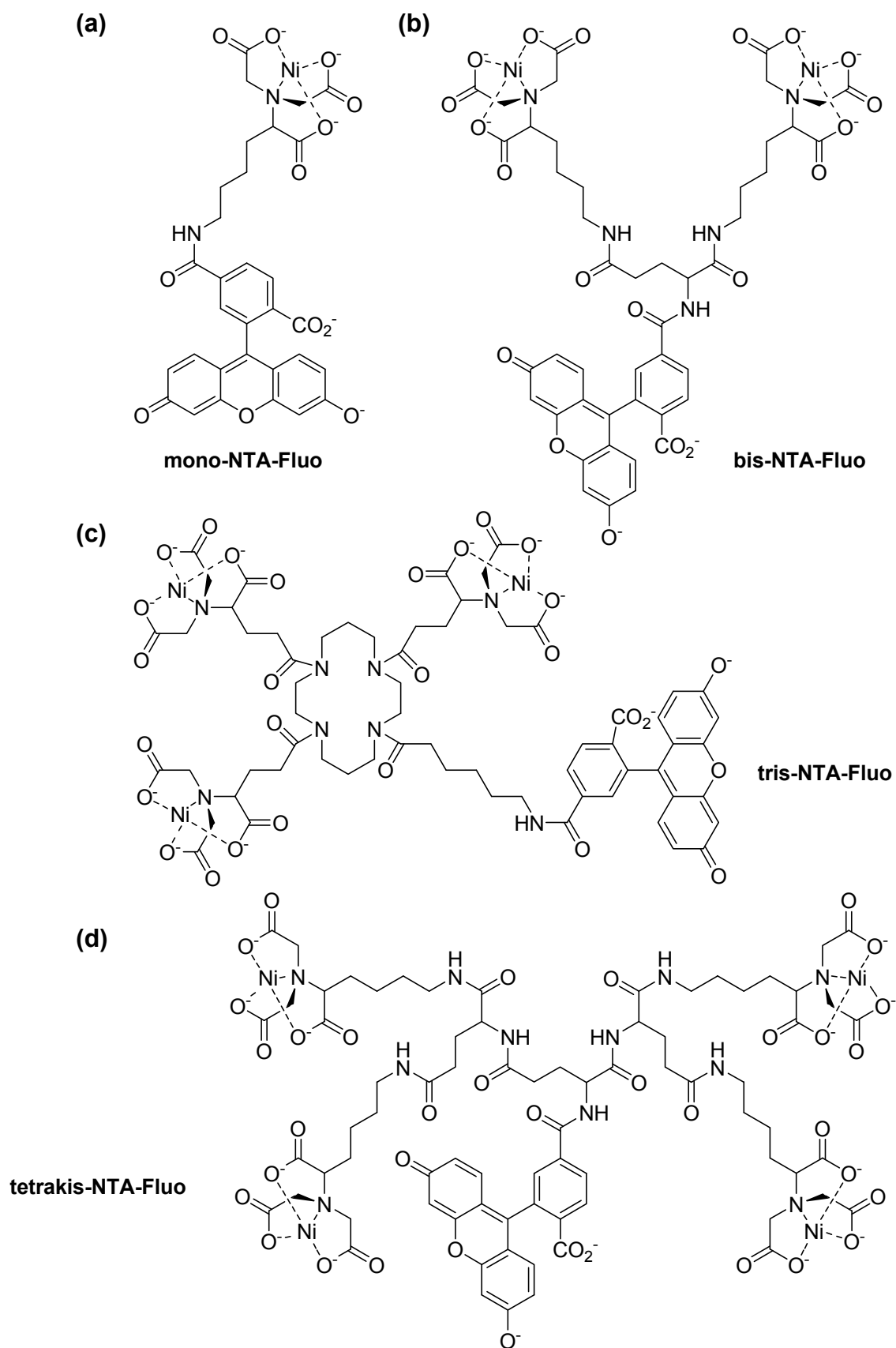
Figure 3. Fluorophore or chromophore-conjugated NTA-Ni^{2+} complexes: (a) (Ni^{2+}:NTA)$_2$-Cy, (b) NTA-FITC-Ni^{2+}, (c) NTA-QSY-Ni^{2+}, (d) NTA-DCF-Ni^{2+}

(a)

mono-NTA-Fluo

(b)

bis-NTA-Fluo

(c)

tris-NTA-Fluo
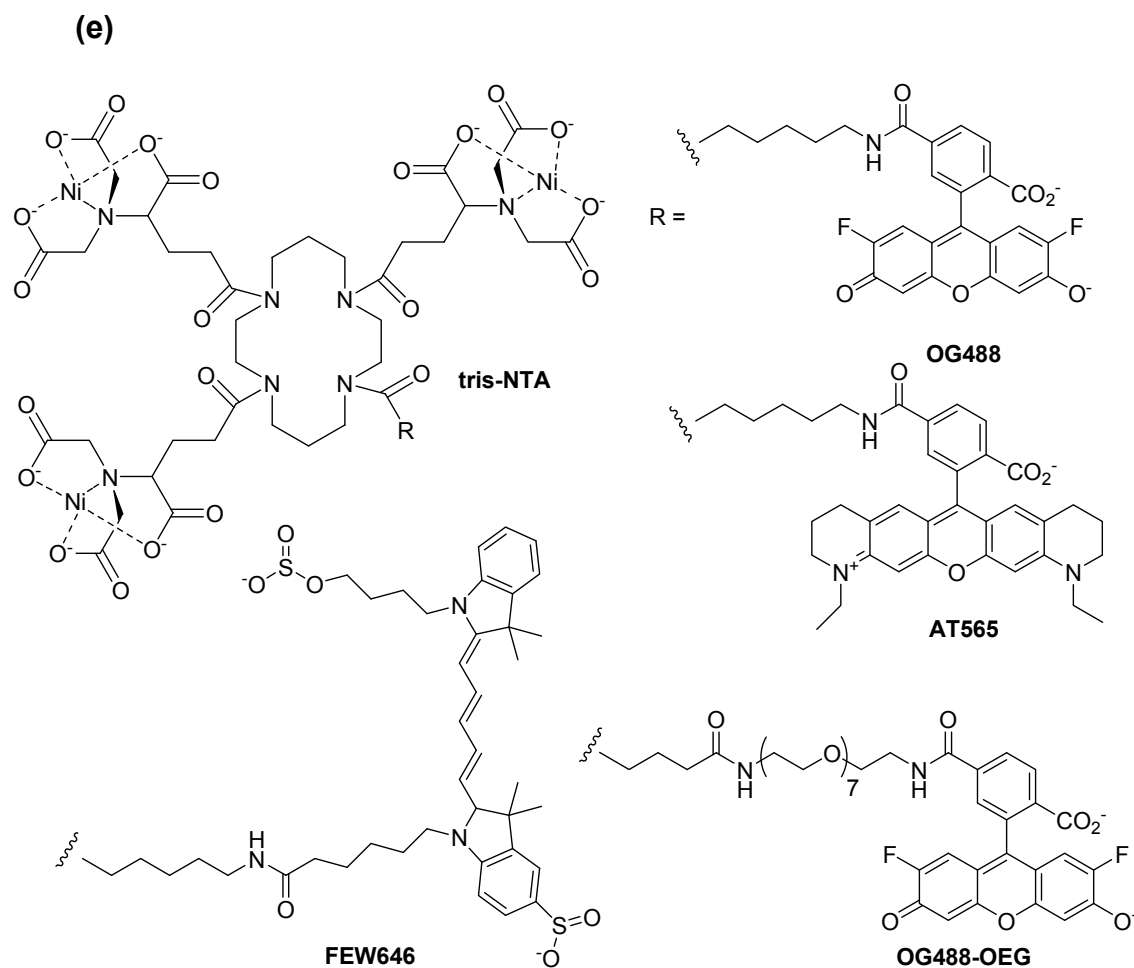
(d)

tetrakis-NTA-Fluo

(e)

Figure 4. Fluorophore-conjugated multivalent NTA-Ni^{2+} complexes: (a) mono-NTA-Fluo, (b) bis-NTA-Fluo, (c) tris-NTA-Fluo, (d) tetrakis-NTA-Fluo, (e) tris-NTAs for multicolor fluorescent detection (^{OG488}tris-NTA, ^{AT565}tris-NTA, ^{FEW646}tris-NTA, $^{OG488-OEG}$tris-NTA)

observed, and the subnanomolar affinity was reached for tris-NTA. Fast reversibility of interaction (switchability) under mild conditions (imidazole, EDTA) was preserved with the high affinity, is noteworthy. After the success, the same group reported color variations of fluorescent tris-NTA, utilizing Oregon Green 488 (^{OG488}tris-NTA), ATTO 565 (^{AT565}tris-NTA), Cy5-analogue FEW S0387 (^{FEW646}tris-NTA), and Oregon Green 488 with a long and flexible heptaethylene glycol spacer ($^{OG488-OEG}$tris-NTA, Figure 4(e)) [45]. Fluorescence labeling by tris-NTA-Ni^{2+} was applied for a protein complex in solution and on surfaces. Because of the remarkable features, the multivalent strategy is also useful for constructing high-performance sensors [46-49]. As other examples, Gierasch *et al.* reported NTA2-BM (BM; dibromobimane) containing two NTA-Ni^{2+} moieties, which was synthesized based on a simple approach using thiol chemistry [50]. Meredith and Allbritton *et al.* achieved stable binding between NTA-Ni^{2+} complex and His-tag, not by multivalent technique, but by photo-crosslinking technique [51].

3.3. New chelate compounds utilizing expanded tag or other metal ion

Recently, new chelate compounds that use a different tag or a metal ion were also reported. Soh *et al.* proposed a novel strategy for ratiometric monitoring of proteins, utilizing dansyl-NTA-Ni^{2+} complex [Figure 5(a)], a NTA-Ni^{2+} complex conjugated with a field-sensitive dansyl fluorophore, and His-Trp-tag, a peptide tag composed of a normal His-tag and a few adjacent tryptophans [52]. When the dansyl-NTA-Ni^{2+} complex is used for a target protein fused with His-Trp-tag, the complex coordinates with the His-Trp-tag due to the affinity between the NTA-Ni^{2+} moiety and the His-tag moiety. Simultaneously, the aromatic dansyl group of the complex would associate with hydrophobic tryptophans, and as a result, both an increase in fluorescence intensitiy and a blue shift in the fluorescence spectrum are observed (I_{bound}/I_{free} = 5.2, $\lambda_{em,free}$ = 491 nm, $\lambda_{em,bound}$ = 538 nm, for His$_6$Trp$_3$-tag). Using the complex, His-Trp-tagged RNase S' and His-tagged RNase S' were clearly identified under light irradiation. The technique would be useful not only for protein monitoring but also one-pot protein analysis without complicated B/F separation.

Although Ni^{2+} is widely used as a mediator for combining a fluorescent ligand with His-tag selectively, the heavy metal is harmful and tends to quench any nearby fluorohpores due to the paramagnetism. Tsien *et al.* developed a novel chelating molecule, 2',7'-bis(pyridyl-2-sulfonamido)-4',5'-dimethylfluorescein (histidine-zinc fluorescent *in vivo* tag, called HisZiFiT, Figure 5(b)) that utilizes Zn^{2+}, which is ubiquitous in biological systems, nutritionally and physiologically essential, diamagnetic, and redox-inert [53]. Since Zn^{2+} is not a fluorescence quencher, the fluorophore of the probe molecule participates directly in metal chelation, resulting in a more compact probe structure, as well as excitation and emission wavelength shifts on Zn^{2+} binding ($\lambda_{ex,free}$ = 515 nm, $\lambda_{em,free}$ = 545 nm, $\lambda_{ex,bound}$ = 534 nm, $\lambda_{em,bound}$ = 560 nm). The K_d value of HisZiFiT-Zn^{2+} to His-tag (His$_6$ sequence) was calculated to nearly 40 nM, which is comparable to bis-NTA ligands, and much lower than mono-NTA ligands. It was successfully demonstrated that the labeling technique using HisZiFiT is helpful to overcome a current controversy related to externalization of the stromal interaction molecule 1 (STIM1) upon depletion of Ca^{2+} from the endoplasmic reticulum.
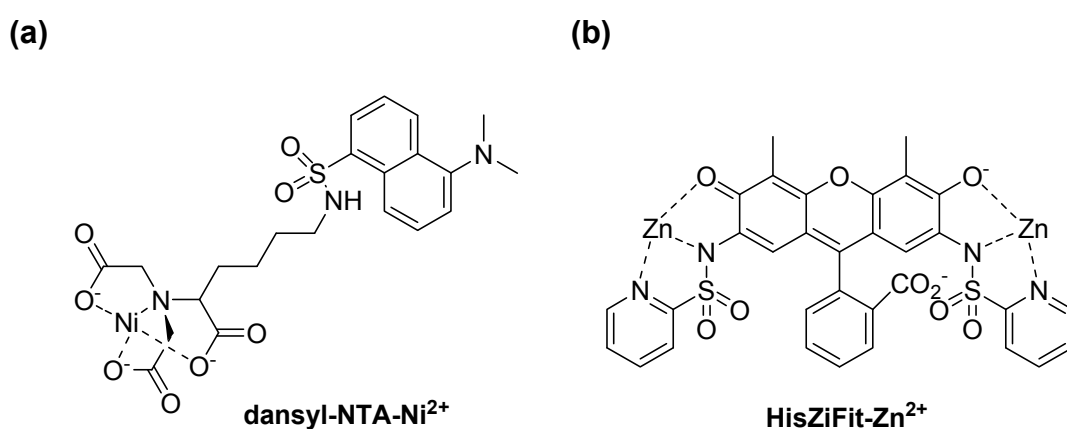

(a)   **(b)**

dansyl-NTA-Ni^{2+}   **HisZiFit-Zn^{2+}**

Figure 5. Chelate compounds utilizing expanded tag or other metal ion: (a) dansyl-NTA-Ni^{2+} (for His-Trp-tag), (b) HisZiFit-Zn^{2+} (comprising Zn^{2+} mediator)

4. D4-tag system

*4.1. Selective protein labeling based on oligo-aspartate/zinc-complex system*

Another peptide tag/artificial probe system, different from tetracysteine/biarsencal or His-tag/NTA systems, was also reported. Hamachi *et al.* developed a protein labeling system composed of a oligo-aspartate tag (D4 tag, $(D4)_n$, *n* = 1-3) and multinuclear Zn^{2+} complexes (Zn(II)-DpaTyrs); FITC-conjugated binuclear complex (Figure 6(a)) and FITC or Cy5-conjugated tetranuclear complexes (Figure 6(b),(c)) [54]. Higher affinity was observed with increasing the number of Asp in D4 tag (binding constants: D2; 7.1×10^3 M^{-1}, D3; 6.3×10^4 M^{-1}, D4; 6.9×10^5 M^{-1}, D5; 8.6×10^5 M^{-1} (comparable to D4) for FITC-conjugated binuclear complex). In addition, the moderate binding affinity between FITC-conjugated binuclear complex and D4-RNase (1.2×10^4 M^{-1}) was significantly enhanced when FITC-conjugated tetranuclear complex and RNase fused with a longer tag (D4-G-D4) were used alternatively (1.8×10^7 M^{-1}). Fluorescent labeling of muscarinic acetylcholine receptor (mlAChR) by the tetranuclear complexes showed that the complexes did not interfere with the original activity of the receptor or the evoked intracellular signaling.

4.2. Oligo-aspartate/zinc-complex system for ratiometric measurements

The oligo-aspartate/zinc-complex system is able to be expanded to ratiometric fluorescent detection of proteins. Hamachi *et al.* synthesized a binuclear complex appended with pH-sensitive seminaphtho-rhodafluor (SNARF, Figure 6(d)) [55]. The ratiometric emission change ($\lambda_{em,free}$ = 586 nm, $\lambda_{em,bound}$ = 628 nm) was induced upon the binding with D4-tag due to the pK_a shift of the SNARF unit, as a result of the electrostatic neutralization of the cationic binuclear complex. The detection system was also applicable to a dual-emission sensing of D4-tagged RNase. The same group also reported a binuclear complex bearing a pyrene [Figure 6(e)], which is known to display an excimer emission at around 480 nm that is more than 50 nm longer than a wavelength of the monomer emission [56]. Using D4-tag containing two D4 sequences, two probe complexes bind to the D4-tag simultaneously so as to decrease the emission at 378.5 nm and 407 nm derived from the monomer, and increase the emission at 472 nm derived from the excimer. The first and second binding of the binuclear complexes with the D4-tag were calculated to be 6.1×10^6 M^{-1} and 4.9×10^6 M^{-1}, respectively, indicating that the second binding was not considerably interfered by the first binding. The application to protein analysis demonstrated the high selectivity of the complex toward the D4-tagged protein than other proteins.

4.3. Non-enzymatic covalent labeling utilizing affinity interaction and a reactive group

Irreversible covalent labeling is advantageous in terms of the allowance of clear analyses of proteins, as well demonstrated by various enzyme-catalyzed labeling methods. Non-enzymatic covalent labeling is attractive because of both non-requirement of a large enzyme or protein domain, and a wide suitability of labeling reagents and reaction conditions, although such labeling methods with sufficient selectivity have been less successful. Hamachi *et al.* recently reported a method for non-enzymatic selective covalent labeling of tag-fused proteins with small molecules including fluorophores [57]. Binuclear Zn(II)-DpaTry bearing both a thiol-reactive *N*-α-chloroacetyl moiety and another functional
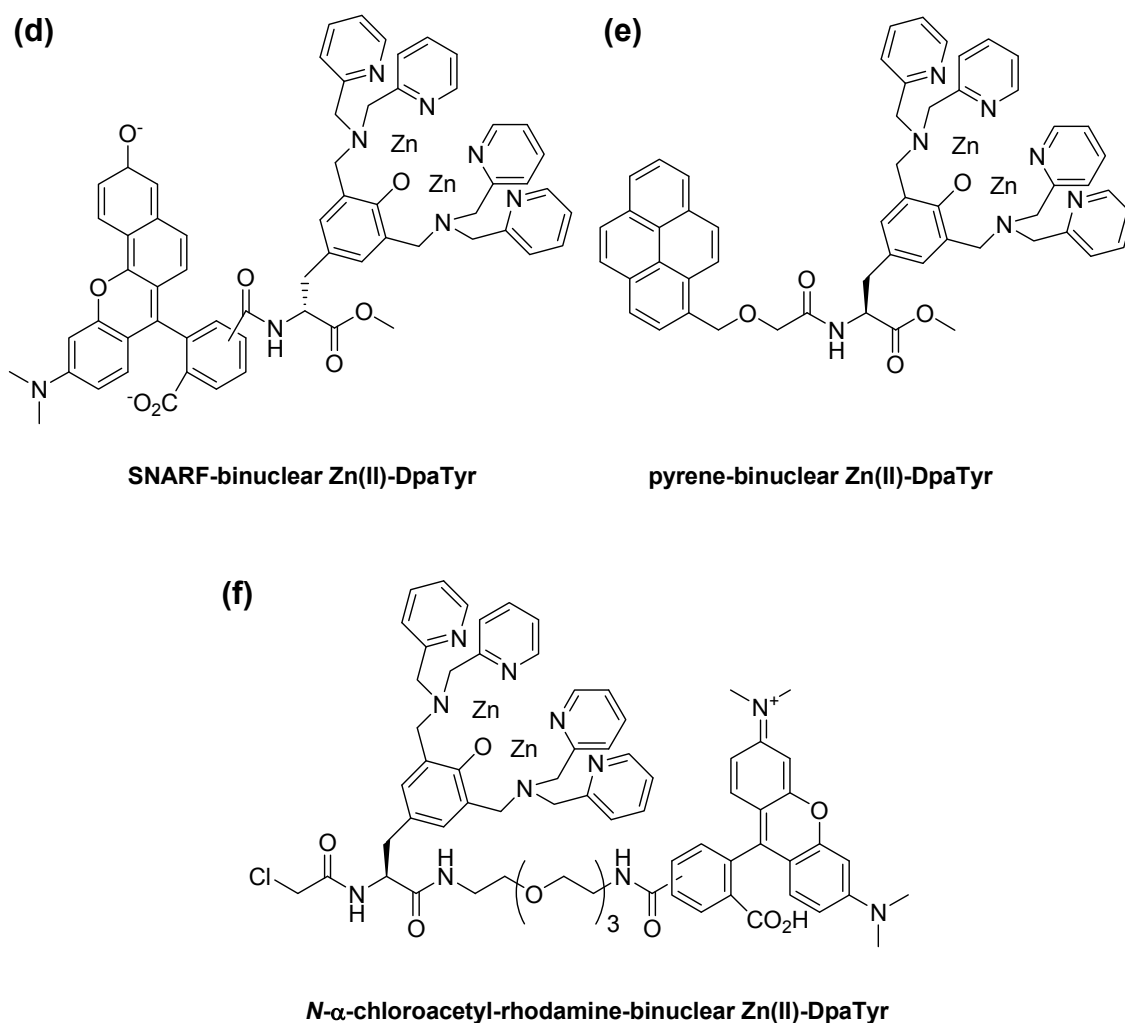
(a) FITC-binuclear Zn(II)-DpaTyr

(b) FITC-tetranuclear Zn(II)-DpaTyr

(c) Cy5-tetranuclear Zn(II)-DpaTyr
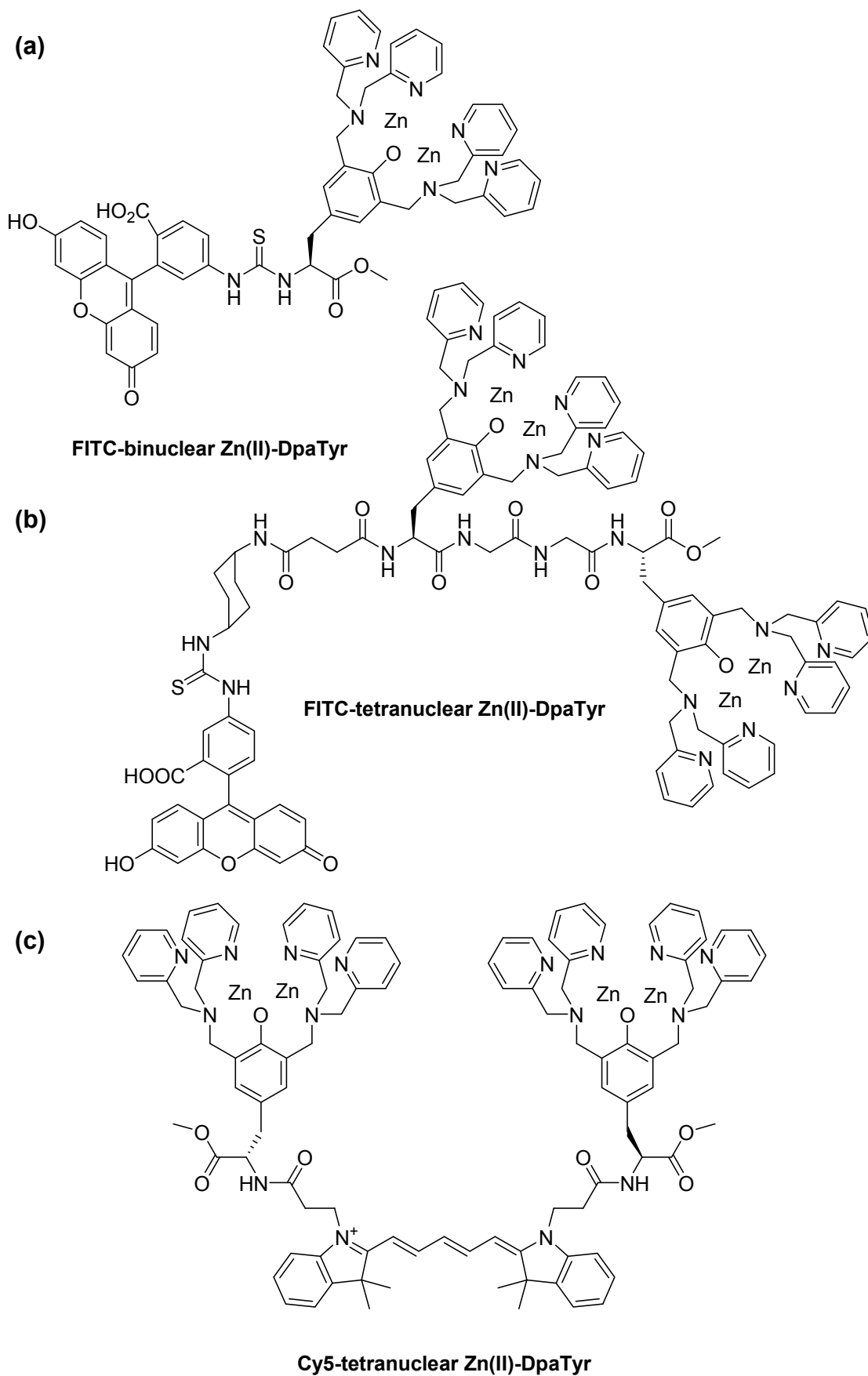
Figure 6. Fluorophore-conjugated multinuclear Zn^{2+} complexes (Zn(II)-DpaTyrs): (a) FITC-binuclear Zn(II)-DpaTyr, (b) FITC-tetranuclear Zn(II)-DpaTyr, (c) Cy5-tetranuclear Zn(II)-DpaTyr, (d) SNARF-binuclear Zn(II)-DpaTyr, (e) pyrene-binuclear Zn(II)-DpaTyr, (f) *N*-α-chloroacetyl-rhodamine-binuclear Zn(II)-DpaTyr

group (fluorescent rhodamine unit, Figure 6(f), or azide-reactive propargyl unit) is selectively associated and bound covalently with D4-tag peptide possessing one Cys group (Cys-Ala$_n$-Asp$_4$), which could be regarded as the expanded version of D4-tag, as like expanded His-tag described in 3.3. Selective fluorescent labeling using the binuclear complex appending propargyl unit was achieved by Huisgen reaction with coumarin azide. SDS-PAGE analysis showed that the selective and covalent labeling of protein with Cys-appended D4 tag was accomplished in protein mixture and a crude lysate. The system was further applied to the protein labeling inside *E. coli* cells.

5. Lanthanide-binding tag system

5.1. Development and applications of lanthanide-binding tag system

Lanthanides display several appealing properties such as long-lived luminescence emissions, which are ideal for time-resolved fluorescent measurements, under appropriate conditions in the presence of a sensitizing chromophore. Imperiali *et al.* proposed lanthanide-binding tag (LBT) system for protein labeling. Based on 14-mer peptide reported by Szabo *et al.* [58], the group searched alternative peptide sequences utilizing a split-and-pool library, and subsequently optimized the position of Cys residues, resulting to find a peptide sequence (ACADYNKDGWYEELECAA), which has a 40-fold higher affinity for Tb^{3+} (K_d = 220 nM) and exhibits significantly brighter luminescence intensity compared to an existing reference peptide (GDYNADGWIEFEEL, K_d = 9.3 μM) [59]. Tb^{3+}-loaded LBT fusion proteins was selectively detected in complex solution mixtures, as well as in SDS-PAGE gels. Using the strategy of solid-phase split-and-pool combinatorial peptide synthesis, the same group subsequently found cysteine-free peptides that bind Tb^{3+} with higher affinity (YIDTNNDGWYEGDELLA (K_d = 57 nM) [60], FIDTNNDGWIEGDELLLEEG (K_d = 19 nM) [61]). The origin of the enhanced luminescence of the LBT and its selectivity and high affinity for Tb^{3+} was investigated from the crystal structure and time-resolved luminescence measurements [62].

The availability of LBT has also been demonstrated. Imperiali *et al.* investigated binding interactions between the SH2 domains of Src and Crk kinase genetically encoded with an LBT, and phosphorylated and non-phosphorylated peptides chemically labeled with BODIPY fluorophores, based on luminescence resonance energy transfer (LRET) [63]. The LRET between the protein-bound Tb^{3+} (LRET donor) and the BODIPY fluorohore (LRET acceptor) showed the specific recognition of the SH2 domain and the peptide binding partner. LBT also accomplished significant alignment of a fusion protein in a magnetic field in the presence of stoichiometric lanthanide ion, and the alignment provided residual dipolar couplings of a magnitude that can be accurately measured by simple methods [64].

5.2. Double-lanthanide-binding tag

In LBT techniques, the increase in the number of bound lanthanides is expected to improve the capabilities for some applications. Imperiali *et al.* developed a double-LBT (dLBT), which concatenates two lanthanide-binding motifs, utilizing a well-characterized single-LBT sequence [65]. The ubiquitin construct fused with dLBT (GPGYIDTNNDGWIEGDELYIDTNNDGWIEGDELLA) showed 3-fold greater luminescence intensity upon lanthanide binding. NMR studies on the dLBT-tagged ubiquitin demonstrated that not only the improved photophysical properties, but also the remarkably increased efficiency of the tag in mediating alignment between the lanthanide ions and protein. In the accompanying paper [66], dLBT was applied for the crystallographic structure determination of protein by the single-wavelength anomalous diffraction method.

6. Conclusions

In this review, advances in techniques for fluorescent labeling of proteins were described, highlighting on the metal-chelation methodology, a kind of affinity labeling. The metal-chelation technique is applicable for a simple and selective labeling based on a small peptide tag, thereby it is regarded as one of the most useful techniques for protein labeling, as well as enzymatic labeling technique. Since the first report on FlAsH by Tsien's group, an increasing number of researchers have energetically developed new selective chemical labeling of proteins with small fluorescent molecules. After this, not only fluorescent labeling methods, having higher affinity and simplicity based on smaller tags, but also fluorescent probes, with more unique or sophisticated functions like CaGF, would be enthusiastically developed. For example, we recently reported a protein probe that can switch the fluorescence states by irradiation of light [67]. A variety of methods, developed so far and from now on, for chemical labeling of proteins with small functional molecules, would be highly expected to contribute to the studies on of protein biochemistry in living cells.

References

1. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509-544.
2. Shaner, N.C.; Steinbach, P.A.; Tsien, R.Y. A guide to choosing fluorescent proteins. *Nat. Meth.* **2005**, *2*, 905-909.
3. Lippincott-Schwartz, J; Patterson, G.H. Development and use of fluorescent protein markers in living cells. *Science* **2003**, *300*, 87-91.
4. Lisenbee, C.S.; Karnik, S.K.; Trelease, R.N. Overexpression and mislocalization of a tail-anchored GFP redefines the identity of peroxisomal ER. *Traffic* **2003**, *4*, 491-501.
5. Andresen, M.; Schmitz-Salue, R.; Jakobs, S. Short tetracysteine tags to β-tubulin demonstrate the significance of small labels for live cell imaging. *Mol. Biol. Cell* **2004**, *15*, 5616-5622.
6. Wang, L.; Schultz, P.G. Expanding the genetic code. *Chem. Commun.* **2002**, 1-11.
7. 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.
8. Lin, C.W.; Ting, A.Y. Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins *in vitro* and on the surface of living cells. *J. Am. Chem. Soc.* **2006**, *128*, 4542-4543.
9. Chen, I.; Howarth, M.; Lin, W.Y.; Ting, A.Y. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat. Meth.* **2005**, *2*, 99-104.

10. Duckworth, B.P.; Zhang, Z.; Hosokawa, A.; Distefano, M.D. Selective labeling of proteins by using protein farnesyltransferase. *ChemBioChem* **2007**, *8*, 98-105.

11. Carrico, I.S.; Carlson, B.L.; Bertozzi C.R. Introducing genetically encoded aldehydes into proteins. *Nat. Chem. Biol.* **2007**, *3*, 321-322.

12. 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.

13. Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9955-9959.

14. Los, G.; Learish, R.; Karassina, N.; Zimprich, C.; McDougall, M.G.; Encell, L.P.; Friedman-Ohana, R.; Wood, M.; Vidugiris, G.; Zimmerman, K.; Otto, P.; Berstock, S.; Klaubert, D.; Wood, K.V. Halotag^TM technology: cell imaging and protein analysis. *Cell Notes* **2006**, *14*, 10-14.

15. Miller, L.W.; Sable, J.; Goelet, P.; Sheetz, M.P.; Cornish V.W. Methotrexate conjugates: a molecular *in vivo* protein tag. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 1672-1675.

16. Miller, L.W.; Cai, Y.; Sheetz, M.P.; Cornish, V.W. In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. *Nat. Meth.* **2005**, *2*, 255-257.

17. Marks, K.M.; Braun, P.D.; Nolan, G.P.; A general approach for chemical labeling and rapid, spatially controlled protein inactivation. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9982-9987.

18. Chen, I.; Ting, A.Y. Site-specific labeling of proteins with small molecules in live cells. *Curr. Opin. Biotechnol.* **2005**, *16*, 35-40.

19. Marks, K.M.; Nolan, G.P. Chemical labeling strategies for cell biology. *Nat. Meth.* **2006**, *3*, 591-596.

20. Miller, L.W.; Cornish, V.W. Selective chemical labeling of proteins in living cells. *Curr. Opin. Chem. Biol.* **2005**, *9*, 56-61.

21. O'Hare, H.M.; Johnsson, K.; Gautier, A. Chemical probes shed light on protein function. *Curr. Opin. Sruct. Biol.* **2007**, *17*, 488-494.

22. Dragulescu-Andrasi, A.; Rao, J. Chemical labeling of protein in living cells. *ChemBioChem.* **2007**, *8*, 1099-1101.

23. Griffin, B.A.; Adams, S.R.; Tsien, R.Y. Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **1998**, *281*, 269-272.

24. Adams, S.R.; Campbell, R.E.; Gross, L.A.; Martin, B.R.; Walkup, G.K.; Yao, Y.; Llopis, J.; Tsien, R.Y. 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.

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

26. Stroffekava, K.; Proenza, C.; Beam, K.G. The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflugers Archiv. Eur. J. Physiol.* **2001**, *442*, 859-866.

27. Poskanzer, K.E.; Marek, K.W.; Sweeney, S.T.; Davis, G.W. Synaptotagmin I is necessary for

compensatory synaptic vesicle endocytosis *in vivo*. *Nature* **2003**, *426*, 559-563.

28. Ju, W.; Morishita, W.; Tsui, J., Gaietta,G.; Deerinck, T.J.; Adams, S.R.; Garner, C.C.; Tsien ,R.Y.; Ellisman, M.H.; Malenka, R.C. Activity-dependent regulation of dendritic synthesis and trafficking of AMPA receptors. *Nat. Neurosci.* **2004**, *7*, 244-253.

29. Hoffmann, C.; Gaietta, G.; Bunemann, M., Adams, S.R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J.P.; Tsien, R.Y.; Eisman, M.H.; Lohse, M.J. A FlAsH-based FRET approach to determine G protein-coupled receptor activation in living cells. *Nat. Meth.* **2005**, *2*, 171-176.

30. Gaietta, G.M.; Giepmans, B.N.G.; Deerinck, T.J.; Smith, W.B.; Ngan, L.; Llopis, J.; Adams, S.R.; Tsien, R.Y.; Ellisman, M.H. Golgi twins in late mitosis revealed by genetically encoded tags for live cell imaging and correlated electron microscopy. *Proc. Natl. Acad. Sci.* **2006**, *103*, 17777-17782.

31. Nakanishi, J.; Nakajima, T.; Sato, M.; Ozawa, T.; Tohda, K.; Umezawa, Y. Imaging of conformational changes of proteins with a new environment-sensitive fluorescent probe designed for site-specific labeling of recombinant proteins in live cells. *Anal. Chem.* **2001**, *73*, 2920-2928.

32. Spagnuolo, C.C.; Vermeij, R.J.; Jares-Erijman, E.A. Improved photostable FRET-competent biarsenical-tetracysteine probes based on fluorinated fluoresceins. *J. Am. Chem. Soc.* **2006**, *128*, 12040-12041.

33. Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T.C.; Mayer, M.U. A red Cy3-based biarsenical fluorescent probe targeted to a complementary binding peptide. *J. Am. Chem. Soc.* **2007**, *129*, 8672-8673.

34. Cao, H.; Chen, B.; Squier, T.C.; Mayer, M.U. CrAsH: a biarsenical multi-use affinity probe with low non-specific fluorescence. *Chem. Commun.* **2006**, 2601-2603.

35. Bhunia, A.K.; Miller, S.C. Labeling tetracysteine-tagged proteins with a SplAsH of color: a modular approach to bis-arsenical fluorophores. *ChemBioChem* **2007**, *8*, 1642-1645.

36. Martin, B.R.; Giepmans, B.N.G.; Adams, S.R.; Tsien, R.Y. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol.* **2005**, *23*, 1308-1314.

37. Chen, B.; Cao, H.; Yan, P.; Mayer, M.U.; Squier, T.C. Identification of an orthogonal peptide binding motif for biarsenical multiuse affinity probes. *Bioconjugate Chem.* **2007**, *18*, 1259-1265.

38. Wang, T.; Yan, P.; Squier, T.C.; Mayer, M.U. Prospecting the proteome: identification of naturally occurring binding motifs for biarsenical probes. *ChemBioChem* **2007**, *8*, 1937-1940.

39. Tour, O.; Adams, S.R.; Kerr, R.A.; Meijer, R.M.; Sejnowski, T.J.; Tsien, R. W.; Tsien, R.Y. *Nat. Chem. Biol.* **2007**, *3*, 423-431.

40. Kapanidis, A.N.; Ebright, Y.W.; Ebright, R.H. Site-specific incorporation of fluorescent probes into protein: hexahistidine-tag-mediated fluorescent labeling with $(Ni^{2+}:nitrilotriacetic\ acid)_n$-fluorochrome conjugates. *J. Am. Chem. Soc.* **2001**, *123*, 12123-1215.

41. Amano, H.; Ohuchi, Y.; Katayama, Y.; Maeda, M. A new fluorescent reagent for the detection of proteins having histidine-tag (his-tag). *Anal. Sci.* **2001**, *17(suppl.)*, i1469-i1471.

42. Guignet, E.G.; Hovius, R.; Vogel, H. Reversible site-selective labeling of membrane proteins in living cells. *Nat. Biotechnol.* **2004**, *22*, 440-444.

43. Goldsmith, C.R.; Jaworski, J.; Sheng, M.; Lippard, S.J. Selective labeling of extracellular proteins

containing polyhistidine sequences by a fluorescein-nitrilotriacetic acid conjugate. *J. Am. Chem. Soc.* **2006**, *128*, 418-419.

44. Lata, S.; Reichel, A.; Brock, R.; Tamplé, R.; Piehler, J. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J. Am. Chem. Soc.* **2005**, *127*, 10205-10215.

45. Lata, S.; Gavutis, M.; Tampé, R.; Piehler, J. Specific and stable fluorescence labeling of histidine-tagged proteins for dissecting multi-protein complex formation. *J. Am. Chem. Soc.* **2006**, *128*, 2365-2372.

46. Tinazli, A.; Tang, J.; Valiokas, R.; Picuric, S.; Lata, S.; Piehler, J.; Liedberg, B.; Tampé, R. High-affinity chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. *Chem. Eur. J.* **2005**, *11*, 5249-5259.

47. Lata, S.; Gavutis, M.; Piehler, J. Monitoring the dynamics of ligand-receptor complexes on model membranes. *J. Am. Chem. Soc.* **2006**, *128*, 6-7.

48. Klenkar, G.; Valiokas, R.; Lundström, I.; Tinazli, A.; Tampé, R.; Piehler, J.; Liedberg, B. Piezo dispensed microarray of multivalent chelating thiols for dissecting complex protein-protein interactions. *Anal. Chem.* **2006**, *78*, 3643-3650.

49. Valiokas, R.; Klenkar, G.; Tinazli, A.; Tampé, R.; Liedberg, B.; Piehler, J. Differential protein assembly on micropatterned surfaces with tailored molecular and surface multivalency. *ChemBioChem* **2006**, *7*, 1325-1329.

50. Krishnan, B.; Szymanska, A.; Gierasch, L.M. Site-specific fluorescent labeling of poly-histidine sequences using a metal-chelating cysteine. *Chem. Biol. Drug. Des.* **2007**, *69*, 31-40.

51. Meredith, G.D.; Wu, H.Y.; Allbritton, N.L. Targeted protein functionalization using His-tags. *Bioconjugate Chem.* **2004**, *15*, 969-982.

52. Soh, N.; Seto, D.; Nakano, K.; Imato, T. Methodology of reversible protein labeling for ratiometric fluorescent measurement. *Mol. BioSyst.* **2006**, *2*, 128-131.

53. Hauser, C.T.; Tsien, R.Y. A hexahistidine-Zn^{2+}-dye label reveals STIM1 surface exposure. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3693-3697.

54. Ojida, A.; Honda, K.; Shinmi, D. Kiyonaka, S.; Mori, Y.; Hamachi, I. Oligo-Asp Tag/Zn(II) complex probe as a new pair for labeling and fluorescence imaging of proteins. *J. Am. Chem. Soc.* **2006**, *128*, 10452-10459.

55. Honda, K.; Nakata, E.; Ojida, A.; Hamachi, I. Ratiometric fluorescence detection of a tag fused protein using the dual-emission artificial molecular probe. *Chem. Commun.* **2006**, 4024-4026.

56. Honda, K.; Fujishima, S.-H.; Ojida, A.; Hamachi, I. Pyrene excimer-based dual-emission detection of a oligoaspartate tag-fused protein by using a Zn^{II}-DpaTyr probe. *ChemBioChem* **2007**, *8*, 1370-1372.

57. Nonaka, H.; Tsukiji, S.; Ojida, A.; Hamachi, I. Non-enzymatic covalent protein labeling using a reactive tag. *J. Am. Chem. Soc.* **2007**, *129*, 15777-15779.

58. MacManus, J.P.; Hogue, C.W.; Marsden, B.J.; Sikorska, M.; Szabo, A.G. Terbium luminescence in synthetic peptide loops from calcium-binding proteins with different energy donors. *J. Biol. Chem.* **1990**, *265*, 10358-10366.

59. Franz, K.J.; Nitz, M.; Imperiali, B. Lanthanide-binding tags as versatile protein coexpression probes. *ChemBioChem* **2003**, *4*, 265-271.

60. Nitz, M.; Franz, K.J.; Maglathlin, R.L.; Imperiali, B. Powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides. *ChemBioChem* **2003**, *4*, 272-276.

61. Martin, L.J.; Sculimbrene, B.R.; Nitz, M.; Imperiali, B. Rapid combinatorial screening of peptide libraries for the selection of lanthanide-binding tags (LBTs). *QSAR Comb. Sci.* **2005**, *24*, 1149-1157.

62. Nitz, M.; Sherawat, M.; Franz, K.J.; Peisach, E; Allen, K.N.; Imperiali, B. Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide. *Angew. Chem. Int. Ed.* **2004**, *43*, 3682-3685.

63. Sculimbrene, B.R.; Imperiali, B. Lanthanide-binding tags as luminescent probes for studying protein interactions. *J. Am. Chem. Soc.* **2006**, *128*, 7346-7352.

64. Wöhnert, J.; Franz, K.J.; Nitz, M.; Imperiali, B.; Schwalbe, H. Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *J. Am. Chem. Soc.* **2003**, *125*, 13338-13339.

65. Martin, L.J.; Hähnke, M.J.; Nitz, M.; Wöhnert, J.; Silvaggi, N.R.; Allen, K.N.; Schwalbe, H.; Imperiali, B. Double-lanthanide-binding tags: design, photophysical properties, and NMR applications. *J. Am. Chem. Soc.* **2007**, *129*, 7106-7113.

66. Silvaggi, N.R.; Martin, L.J.; Schwalbe, H.; Imperiali, B.; Allen, K.N. Double-lanthanide-binding tags for macromolecular crystallographic structure determination. *J. Am. Chem. Soc.* **2007**, *129*, 7114-7120.

67. Soh, N.; Yoshida, K.; Nakajima, H.; Nakano, K.; Imato, T.; Fukaminato, T.; Irie, M. *Chem. Commun.* **2007**, 5206-5208.