[Frontiers in Bioscience 13, 1716-1732, January 1, 2008]

Site-specifically modified fusion proteins for molecular imaging

Hui Wang[1], Xiaoyuan Chen[1]

[1]The Molecular Imaging Program at Stanford (MIPS), Department of Radiology and Bio-X Program, Stanford University School of Medicine, Stanford, California

TABLE OF CONTENTS

1. ABSTRACT

To visualize and quantify specific molecular pathways *in vivo*, it is essential to develop molecular imaging probes with high target affinity and specificity. The introduction of radioisotope, fluorophores and other detectable labels onto a protein ligand in a site-specific manner without compromising its function and binding affinity has been a valuable tool for such molecular imaging. Site-specific labeling of proteins poses an enormous challenge, because the correct functional groups of one protein amidst all of the other expressed proteins containing the same range of amino acids must be accurately targeted. Chemists have developed many diverse strategies for accomplishing sit-specific modification; unnatural amino acids have been incorporated into targeted protein for site-specific modification. Enzymatic post-translational modification of proteins is an alternative approach and beneficial complement for site-specific labeling. A small peptide tag is a substrate for the enzyme that can be appended to N-, C-terminus or internal of target protein for site-specific labeling. In this review, we will discuss several valuable tags for addressing this problem, and comment on the potential usage of tagged fusion protein for molecular imaging.

2. INTRODUCTION TO MOLECULAR IMAGING PROBES

In the past decade, numerous novel treatments have aimed to prolong survival, induce remission, and provide better quality of life for cancer patients. Among the successful examples are the molecularly targeted cancer treatment drugs such as the anti-HER-2/neu antibody trastuzumab for Erb-2 expressing breast cancers (1), and the kinase inhibitor imatinib (Glivec®) for chronic myelogenous leukemia and gastrointestinal stromal tumors (2). Despite their efficacy in certain settings, molecularly targeted drugs are resource-intensive to develop, being expensive and requiring protracted time, and numerous patients during development, which result in the inevitably high cost of the approved therapy (3). Moreover, molecular targets have been identified and characterized in relatively few oncology patients. Improving the phenotypic characterization of the underlying molecular lesions would therefore expand the overall effect of targeted agents in oncology.

The development of targeted therapy strategy highlights the need for a faster, more efficient and cost-effective development to better define patients likely to

benefit from treatment. As addressed by the recent Food and Drug Administration (FDA) Critical Path Initiative, collaborative interactions among such scientific knowledge areas as bioinformatics, genomics, materials science, and imaging technologies are needed to design and implement better drug development tools (3). Important among these tools are molecular imaging probes that image specific molecular pathways *in vivo*, enabling visualization of phenotypic expression of key targets in cancer disease processes.

Probes to image molecular targets have included small molecules, peptides, proteins, and antibodies labeled with radionuclides for positron emission tomography (PET) and single photon emission computed tomography (SPECT) imaging, near-infrared (NIR) fluorescence imaging, molecular magnetic resonance imaging (mMRI) that employs paramagnetic and superparamagnetic-based moieties, and microbubbles for contrast enhanced ultrasound (CEU).

Despite their promise for characterizing targets and the functional consequences of drug-target interactions, the development of molecular imaging probes has faced several significant hurdles. For example, the targets are expected to have low (nanomolar to micromolar) concentrations; thus, adequate intracellular delivery and acceptable signal-to-noise amplification are necessary. Furthermore, the molecular imaging probe should have only a trace level of mass in tissues of the target(s) so as not to exert mass or pharmacologic effects. In addition, the molecular imaging probe should have very high target affinity and specificity so that the image contrast will truly reflect the receptor density or enzyme activity found in target organ or tissue. This review article focuses on protein based molecular imaging probes and the importance of site-specific labeling to maintain the receptor binding affinity and specificity after appropriate modification, bioconjugation, or labeling.

3. SITE-SPECIFIC CHEMICAL MODIFICATION OF PROTEIN PROBES

Fluorophores, isotopes, and other detectable labels are most commonly introduced into a protein through *in vitro* modifications with suitable amine or thiol reactive reagents. Proteins typically contain multiple amines. Even with the reduced pK_a of the amino-terminus of the protein compared to the lysine side chain, protein labeling is generally nonspecific. Cysteine labeling is typically more specific than amine labeling as many natural cytosolic proteins lack cysteine residues and a single cysteine can be added by site-directed mutagenesis without affecting the function of the protein. However, cysteine-based methods alone cannot easily perform multiple modifications to a protein (such as introducing two fluorophores for fluorescence resonance energy transfer (FRET) analysis) except when the reactivities of the two cysteines are quite different, and not all protein targets allow for the introduction of cysteine without impairing function. Recent work by Francis *et al.* has identified reactions that can specifically target tryptophan(4) and tyrosine(5) residues. Using rhodium carbenoids, the selective modification of the indole functionality in tryptophan (at N-1 or C-2) can be achieved. When applied to myoglobin or subtilisin Carlsberg (with 2 and 1 surface tryptophan, respectively), modification occurs exclusively at the tryptophan residues in about 50-60% yield. The results do suggest that transition metal complexes can be used for specific bioconjugation of aromatic residues in proteins. However, this reaction requires low pH (1.5-3.5) and an organic co-solvent to solubilize the organic reagents, both of which may denature the target protein. The same group also developed conditions to perform a Mannich-type of reaction to selectively target tyrosine residues using aldehyde and aniline at pH 6.5(5). By attaching a fluorophore to the aniline, conjugation of rhodamine to chymotrypsinogen A is achieved without altering the protein's activity. Another strategy for selective tyrosine modification using diazonium salts (6)has been applied to modify the exterior of the rodlike tobacco mosaic virus. The addition of a variety of groups, including biotin, fluorophores, and crown ethers modulates the properties of this protein scaffold.

Site-specific chemical modification is strictly defined as a process which yields a stoichiometrically altered protein with the quantitative covalent derivatization of a single, unique amino acid residue without either modifying the other amino acid residue or changing the conformation of the protein. In practice, this objective is rarely obtained because of several major problems. First, only very few reagents are specific for the modification of a single functional group. Second, even with a reagent that appears to be functional group specific, the modification of only one of the several residues within a functional group class is unlikely unless the specific residue is uniquely reactive. Finally, it is also unlikely that the site-specific modification of a protein can be achieved without any conformational change.

4. SITE-SPECIFIC MODIFICATION VIA INCORPORATED UNNATURAL AMINO ACID IN TARGET PROTEIN

Biological methods such as site-directed mutagenesis can be used to systematically alter proteins at the genetic level. Through unnatural amino acid mutagenesis, chosen modifications can be added to proteins during their cellular synthesis. As far as we know, all known organisms specify the same 20 amino acids (with the rare exceptions of selenocysteine (7) and pyrrolysine (8)). Although a 20-amino-acid code might be sufficient for life, it might not be optimal. The methodology to incorporate amino acid with defined steric and electronic properties at unique sites in proteins will provide powerful new tools for exploring protein structures and functions *in vitro* and *in vivo*. To-date, over 30 unnatural amino acids (Figure 1)-including those containing spectroscopic probes, post-translational modifications, metal chelators, photoaffinity labels and other chemical moieties- have been selectively incorporated into proteins with high fidelity and efficiency in response to unique three and four base codons (9). For example, Ohno *et al.* (10) utilized the yeast amber suppressor $tRNA^{Tyr}$/mutated tyrosyl synthetase pair as a

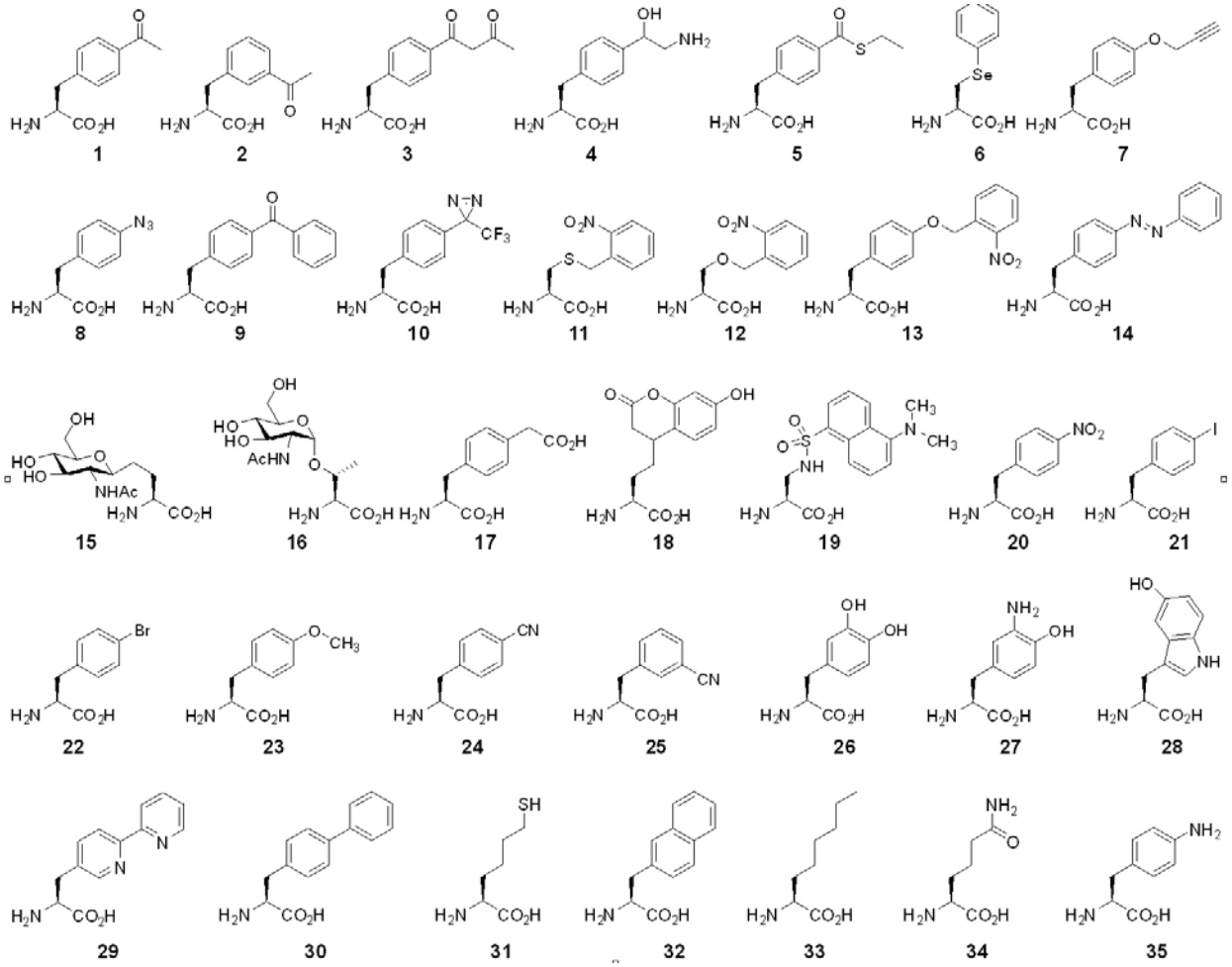


Figure 1. Unnatural amino acids that have been added to the genetic codes of prokaryotes and eukaryotes. Reproduced with permission from (9).

carrier for 3-azidotyrosine in an *E. coli* cell-free translation system, and triarylphosphine derivatives for site-specific modification of the azido group. Using rat calmodulin (CaM) as a model protein, several unnatural CaM molecules were prepared, each carrying an azidotyrosine at predetermined positions 72, 78, 80, or 100, respectively. Post-translational modification of these proteins with a conjugate compound of triarylphosphine and biotin produced site-specifically biotinylated CaM molecules. The CBP-binding activity of CaM was not affected by this biotinylation. This site-specific modification method is easy and fast and can be expanded to introduce other desirable compounds into selected site of proteins as long as appropriate triarylphosphine derivatives can be chemically synthesized. The limitation of this strategy is that the overall modification efficiency is typically only in the range of about 30-50%. One possible explanation for this incomplete modification might be the reduction of 3-azidotyrosine to 3-amino-tyrosine during the course of reaction. Establishing the optimal reaction conditions to minimize such reduction reaction would be valuble to improve the overall modification efficiency.

5. Fusion proteins for site-specific modification

Recombinant DNA construction of protein fusions in which the protein segment has a dedicated site for labeling or could be recognized by a coenzyme ligase fused to a protein of interest, may offer distinct advantages over conventional protein fusions and chemical or mutagenesis-based site-specific modification of target protein. Some obvious benefits of the fusion protein approach over chemical modification and site-specific mutagenesis include the following: (1) the fusion proteins can be specifically labeled by growth of cells cultures in the presence of labeled coenzymes, or be site-specifically labeled under certain mild conditions; (2) if the coenzyme possesses specific binding toward an immobilized ligand, the fusion protein can be readily purified, perhaps in a native form; (3) the protein segment that is fused to the protein of interest would be much smaller than those commonly used, with less alteration of the chemical and biological properties of the protein of interest; and (4) the tag is usually fused to the N- or C-terminus of the targeted protein, with minimum effect on the protein structure and function.

The addition of unnatural moieties, such as fluorophores, affinity labels, spin-label probes, isotope, and bioorthogonal functional groups to proteins has also proven useful for a variety of processes and applications both *in vitro* and *in vivo*, including molecular imaging. In this review, we will focus on the strategies for constructing fusion proteins containing small peptide tag for site-specific modification as a beneficial complementary of the abovementioned methods and the potential application of such fusion protein in molecular imaging.

5.1. Cys-tag

As previously mentioned, thiol-reactive reagents such as the maleimide-based reagents and the haloacetic-based reagents are somewhat more selective because of the lower amounts of cysteine residues in peptides and proteins. In cases where all thiol groups are involved in disulfide bond formation, one can introduce a single surface-accessible thiol into the target protein by site-directed mutagenesis(11).

Introducing N-terminus or C-terminus cysteine residue for site-specific labeling using thiol-specific reagents provides an alternative approach. Olafsen *et al.* (12) have previously produced an anti-carcinoembryonic antigen (anti-CEA) diabody assembling V_L-eight amino acid linker-V_H. Following radiolabeling at random sites on the protein with radioiodine or radiometals, this fragment exhibits rapid tumor targeting in a nude mouse/LS174T human colon carcinoma xenograft model in biodistribution and imaging studies (13). In order to allow site-specific radiolabeling using thiol-specific reagents, they (12) described four mutant anti-CEA diabodies engineered by substitution or addition of unique cysteine residues. Four variants of anti-CEA were constructed, expressed and purified, two of which (with the C-terminal sequence –LGGC or –SGGC) were found to exist as a stable disulfide-linked dimer. The purified –LGGC Cys-diabody was reduced and conjugated with a bifunctional chelating agent comprised of the macrocyclic chelator DOTA (1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetraacetic acid), a tetrapeptide linker and a hexanevinyl sulfone group for chemical attachment to thiol groups. This compound, DOTA-glyclyleucylglycyl(ε-aminobis-1,6-hexanevinyl sulfone)lysine, was abbreviated as DOTA-GLGK-HVS. DOTA-GLGK-HVS-conjugated Cys-diabody was then labeled with copper-64 for microPET imaging using CEA-positive LS174T and CEA-negative C6 rat glioma xenograft. Although the initial biodistribution studies using radioiodinated Cys-diabody showed equivalent targeting ability to parental anti-CEA diabody, the performance of ^{64}Cu-DOTA-GLGK-HVS-conjugated Cys-diabody was not optimal *in vivo*, resulting in high kidney and liver activities. Nevertheless, the novel covalently-linked Cys-diabody described here provides a platform for ready conjugation of a wide variety of effector moieties to this rapid targeting anti-tumor molecule. These include alternative radiolabeling approaches such as developing an F-18 tag with thiol-specific chemistry for generating additional PET tracers. Other tags such as optical or fluorescent probes can be site-specifically attached. In addition, the C-terminal cysteine residues can be used for conjugation of other small molecules like chemotherapeutic drugs. Using a similar strategy, Natarajan *et al.* (14) incorporated a single cysteine to C-terminus of four different scFv fragments for site-specific PEGylation. Their results demonstrated that scFv-c can be PEGylated in a site-specific manner using PEG-maleimide.

To load vascular enthothelial growth factor (VEGF) with ^{99m}Tc without destructive random modification of the protein with chelating agents, Blankenberg *et al.* (15) expressed $VEGF_{121}$ with a cysteine-containing peptide tag, named C-tag, which was engineered by R→C substitution in Hu-tag (sequence: KETAAAKPERQHMDS, viz. S-tag, (16)). C-tagged VEGF (C-VEGF) was reduced, conjugated with HYNIC (6-hydrazinopyridine-3-carboxylic acid) and then radiolabeled with ^{99m}Tc. The labeled ^{99m}Tc-HYNIC-C-VEGF fully retained VEGF bioactivity and receptor binding ability *in vitro*. Imaging of orthotopic mouse mammary 4T1 carcinomas revealed that ^{99m}Tc-HYNIC-C-VEGF is preferentially accumulated at the tumor rim, where the most extensive angiogenesis takes place. They also reported a similarly enhanced circumferential uptake with near-infrared fluorescence and positron emission tomography (PET) imaging of subcutaneous 4T1 tumors (16). Using the same 15-amino acid Cys-tag, Backer *et al.* (17) expressed an fusion protein containing N-terminal Cys-tag and single chain VEGF (Cys-tag scVEGF); this fusion protein was conjugated to fibronectin (FN) through a common thiol-directed bi-functional cross-linking agent. FN-scVEGF conjugates, when coated on cell culture plate, retained soluble VEGF activity on stimulating VEGFR2 expressing 293/KDR cell growth. These results demonstrated that site-specific immobilization via Cys-tag can provide a facile and reliable method for permanent deposition of functionally active growth factors on synthetic or protein scaffolds that has applications for advanced tissue engineering. Furthermore, to image the VEGFRs expression in angiogeneic vasculature with the single-chain VEGF, this cysteine-tagged single-chain VEGF was site-specifically labeled with contrast agents for near-infrared fluorescence imaging, single-photon emission computed tomography or photon emission tomography (18).These probes retained VEGF activities *in vitro* and underwent selective and highly specific focal uptake into the vasculature of tumors and surrounding host tissue *in vivo*. The fluorescence contrast agent showed long-term persistence and co-localized with endothelial cell markers, indicating that internalization was mediated by the receptors. The multimodal imaging of VEGF receptors with these probes will be useful for clinical diagnosis and therapeutic monitoring, and will help accelerate the development of new angiogenesis-directed drugs and treatments.

Besides its usefulness in site-specific radiolabeling, fluorophore coupling, and PEGylation via Cys-tag incorporated into target protein, another interesting application of Cys-tag is to introduce multiple, covalent modifications involving several different functionalities in a site-specific manner (Figure 2). Cysteines that are oxidized in a disulfide bridge, or coordinated to a metal can be protected from covalent modification by thiol-reactive

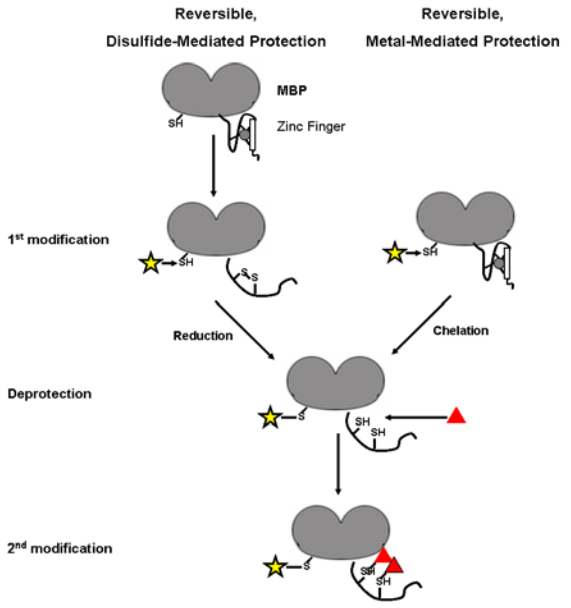


Figure 2. Schemes for producing multiple, site-specific modifications in zinc finger fusion proteins using either reversible metal coordination or disulfide-mediated protection strategies. Two distinct thiol-reactive modifications are represented as ☆ and ▲. Reproduced with permission from (21).

reagents. Protection is readily reversed by reduction or metal chelation. It is therefore possible to engineer proteins with multiple, independently addressable, site-specific covalent attachment points by constructing several cysteines that (1) are unprotected, (2) react to form a disulfide, and (3) participate in metal binding. One such example is made by fusing a given protein with a consensus zinc-finger domain, ZifQNK (19). This 32-residue domain has a Cys_2His_2 primary coordination sphere that binds Zn^{2+} reversibly with 10^{-9}–10^{-11} M affinity (20). In the absence of Zn^{2+}, the two cysteines can form a disulfide bond under oxidizing conditions. ZifQNK can therefore be used in either metal-dependent or redox-dependent RTP strategies (MRTP, RRTP). Another example is a truncated, 18-residue version of ZifQNK (βZIF) in which the single α-helix bearing the two histidines has been deleted, leaving a two-stranded β-sheet containing the two cysteines that readily oxidize to form a disulfide but do not bind Zn^{2+} in the reduced form. Thus, βZIF can be used in a RRTP strategy. Smith *et al.* (21) constructed a variety of N- and C-terminal fusions of these domains with maltose-binding protein, which were then labeled with two or three different fluorophores. MBP protein has a single cysteine engineered at position 141 (C141). For double fluorophore labeling, ZifQNK or βZIF was fused to either N-, or C-terminal of MBP, Cy5-maleimide mono-reactive dye and tetramethylrhodamine-5-maleimide (TMR) were used as the fluorescence labels. Cy5 was first attached to the unprotected C141, followed by deprotection (chelation or reduction) and attachment of TMR labels to ZifQNK or βZIF. For triple labeling, they constructed a MBP141 with βZIF fused to the N-terminus, and ZifQNK fused to the C terminus; the Cy5, TMR and 5-iodoacetamide fluorescence (IAF) were used as fluorescent labels. The modification and deprotection were carried out in the following order: (1) the unreacted thiol was modified; (2) βZIF was deprotected by reduction, and modified; (3) ZifQNK was deprotected by chelation, and modified. Their results demonstrated the feasibility of orthogonal protein modification without apparent mislabeling and the successful illustration of intramolecular FRET (double labeling) and FRET relay (triple labeling). The same results were also obtained using glucose-binding protein (GBP), in which a cysteine was engineered at position 149 instead of MBP, as a model protein. This strategy provides a rapid, straightforward method for producing multiple, site-specific modifications to almost any proteins.

Another strategy is to introduce two functional labels into the target protein by combining chemical and enzymatic modifications. Transglutaminase (TGase) catalyzes the acyl transfer reaction between the γ-carboxyamide group of an acceptor glutamine (Gln) residue and a primary amine donor (22). This TGase-base site-specific modification of protein will be discussed in the following part of this review. Jager *et al.* (23) combined traditional chemical thiol labeling with enzymatic modification of an engineered, reactive Gln by TGase to obtain highly pure, stoichiometrically D/A-labeled protein. In this study, A 64-residue, single domain serine protease inhibitor, chymotrypsin inhibitor 2 (CI2), served as a model system. A unique Cys residue (Cys400 was introduce at a solvent-exposed position into CI1 for site-specific labeling with Alexa488-maleimide (A488 hereafter). A substrate sequence-recognition tag (Pro-Lys-Pro-**Gln**-Gln-Phe) for tissue TGase was appended at the N-terminus of CI2. TGase then catalyzed the acyl transfer reaction between the primary amine of Alexa746 cadaverine (A647-$(CH_2)_6$-NH_2, A647 hereafter) and the γ-carboxyamine moiety of the reactive Gln (Gln4) in the TGase tag. The three internal Gln were imbedded into a rigid tertiary structure context not recognized as substrate for tissue TGase. These results showed that chemical labeling of cysteine, followed by enzymatic modification of reactive Gln in this fusion protein made stoichiometrically D/A labeled substrate suitable for single-molecule FRET experiments. Thermodynamic data indicated that neither the presence of TGase tag nor D/A/ labeling perturbed protein stability. This two-step chemical/enzymatic labeling may thus present a simple, low-cost, and widely available strategy for the D/A labeling of protein in FRET-based single-molecule protein folding studies.

The strategies to introduce two or more functional group into target protein in a site-specific manner may also provide a platform to site-specifically double label a target protein with amine-reactive reagents and sulfhydryl-reactive agents simultaneously for dual-modality molecular imaging. For example, amine reactive chelator DOTA can be conjugated to one site and subsequently labeled with

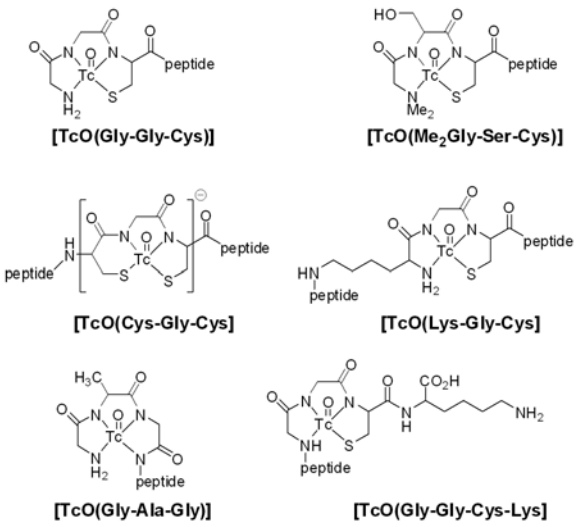


Figure 3. Examples of short amino acid sequences that form stable $[^{99m}TcO]^{3+}$ complexes. Reproduced with permission from (27).

Cu-64 for PET imaging, whereas Cy5-maleimide can be conjugated to a thiol reactive site for optical imaging. This will be a complement the current multimodality imaging that uses a triple/dual fusion reporter gene strategy. However, introducing unpaired engineered cysteines is problematic for routine production and storage, because it can interfere with the disulfides in the target protein and cause misfolding, leading to poor production yields and covalent aggregations. Lastly, trans-chelation to other sulfur-containing ligands has been shown to be an important route of radiopharmaceutical instability (24).

5.2. Chelator tag

^{99m}Tc is a widely used isotope for SPECT and scintigraphic imaging. ^{99m}Tc is characterized by optimal physical properties such as a half-life of 6 h, a nearly monoenergenic γ-emission of 140 keV, and its instant availability through a generator system. The creation of ^{99m}Tc binding site by the conjugation of a chelator to the targeting molecule is associated with instability of the labeled proteins. Conjugates of a cleavable, ester-linked bifunctional chelator RP-1 with scFv has been used to prepare ^{99m}Tc-labeled scFv species (25, 26) for *in vivo* biodistribution and scintigraphic imaging studies. However, such methods require specialized expertise in coupling chemistry and typically result in heterogeneous products with possible damage to the labeled protein (27). Intact IgG can be labeled directly with ^{99m}Tc through the hinge-region cysteine residue, but unmodified scFv cannot be directly labeled, as the conserved cysteine pairs in each variable domain form essential disulfide bonds.

Several short amino acid sequences (Figure 3) (27) that provide coordinating nitrogen and sulfur atoms have been incorporated into the N- or C-terminus of bioactive molecules before radiolabeling with ^{99m}Tc. This approach has several attractive features. First, the ^{99m}Tc-binding amino acid sequence can be added to or incorporated within proteins by DNA recombinant techniques. Second, these amino acid sequences form stable $[TcO]^{3+}$ complex when ^{99m}Tc pertechnetate is reduced by $SnCl_2$ in the presence of a peptide containing one of the chelating amino acid sequences. Third, the charge and polarity of the complex can be modified by changing the chelating peptide sequence. For example, George *et al.* (28) introduced a Gly_4Cys as a peptide chelate into the C-terminus of 741F8 scFv, which targeted the tumor-associated antigen c-erbB-2. This Gly_4Cys peptide is typically coordinated with ^{99m}Tc as oxotechnetium through its amide nitrogens and the terminal cysteine, forming a square planar complex. The glycines allow maximal flexibility of the peptide to facilitate chelation of ^{99m}Tc without disrupting the structure of adjacent framework regions. These results showed that the complex formed between the peptide and ^{99m}Tc was stable both *in vitro* and *in vivo*, and could be used to image tumor very efficiently. In addition, a (Gly)3-Cys-Ala sequence was genetically inserted at the C-terminus of L19, which is a single-chain antibody fragment that specifically targets the ED-B sequence of fibronectin (29). The resulting fusion protein AP39 could be directly labeled with ^{99m}Tc. In this molecule, technetium forms a Tc (V) oxo metal complex with three amide nitrogen atoms and one thiol sulfur atom in the Gly-Gly-Cys sequence. Bogdanov *et al.* (30) showed specific Tc chelation to a construct of green fluorescent protein with the added C-terminal sequence LGGGGCGGGCGI. Tait *et al.* (31) expressed ACGGGHM-Annexin V, AGGCGHM-Annexin V and ACGCGHM-Annexin V fusion proteins. Two derivatives form N_3S and one forms N_2S_2 chelation site for subsequent ^{99m}Tc radiolabeling for imaging purpose. All three proteins could be labeled with ^{99m}Tc to specific activities of at least 50-100 μCi/μg. The proteins retained membrane binding activity after the radiolabeling procedure, and quantitative analysis indicated a dissociation constant (K_d) of 7 nmol/L. The radiolabeled proteins were stable when incubated with phosphate-buffered saline or serum *in vitro*. Annexin V site specifically labeled with ^{99m}Tc at the N-terminus showed twice as much apoptosis-specific liver uptake as did Annexin V derivatized randomly via amino groups. Another widely used specific peptide sequences, mercaptoacetyltriglycine (MAG3) (32-35), have also been introduced into recombinant proteins in order to create endogenous chelation sites for technetium labeling (34-36).

All the Tc complexes mentioned above that have been used for Tc labeling of proteins are "classical Werner type" coordination compounds without direct metal-carbon bonds. Technetium is usually in the oxidation state +5 as $[M=O]^{3+}$ core, wrapped by tetradentate N- or S-ligands, which provide high thermodynamic stability because of their chelate effect. Waibel *et al.* (37) reported a new approach using Tc(I)-carbonyl compound to form an extremely stable complex to the His-tag, which is one of the most convenient attachment sites and has been used widely as a purification tag for immobilized metal ion affinity chromatography.

Besides technetium, a tetracysteine (TC) tag: Cys-Cys-Xaa-Xaa-Cys-Cys, where Xaa is an noncysteine amino

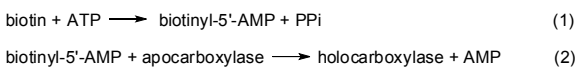
acid (38-40) is reported able to specifically chelate biarsenical compounds such as FIAsH, a fluorescein derivative, and ReAsH, a resoeufin derivative, for site-specific protein modification. The main advantage of TC tag over autofluorescent proteins is its relatively small size, which can be a mere six amino acids (CCPGCC in single-letter amino acid code).

5.3. BirA-based site specific biotinylation

Protein biotinylation is a powerful technique for many molecular biology and biomedical applications due to the high affinity and the specificity of the biotin-avidin interaction. The biotin-avidin interaction has been extensively exploited for labeling biochemical reagents such as antibodies due to the availability of methods to covalently attach biotins to reagents *in vitro*. Likewise, *in vitro* biotin labeling has been applied to a variety of drug-targeting and viral gene therapy vector-targeting strategies. While the existing biotinylation technology is certainly robust, it is limited by the fact that the *in vivo* labeling procedure is not site-specific, requires a prior purification of the labeled substrate, and can lead to inactivation of the biotinylated protein.

Given the difficulties with chemical biotinylation, more recent efforts have been devoted to developing systems in which protein biotinylation can be achieved directly in living cells without chemical reagents. Heterologous expression systems emulating such a "metabolic" approach offer distinct advantages over chemical biotinylation of target proteins. For example, recombinant biotinyl domain fusion proteins can be biotinylated in living cells without the use of specialized reagents and without non-specific and heterogeneous biotinylation at sites other than the acceptor lysine of the biotinyl domain.

Biotin-dependent carboxylases (41) are a class of enzymes that undergo post-translational modification in which the biotin moiety is covalently linked to a single lysine residue via an amide bond. This two-step reaction, as summarized in the following equations, is catalyzed by a class of enzymes termed the biotin holoenzyme synthetases (BHS)



In the first step, the enzyme catalyzes formation of an activated intermediate, biotinyl-5'-adenylate (bio-5'-AMP) from biotin and ATP. Subsequent nucleophilic attack at the activated carboxylate of biotin results in formation of the amide between the biotin moiety and the target lysine residue. The biotinylation reaction is highly specific, with only the biotin-dependent carboxylases serving as substrates *in vivo*. Here, we reviewed the application of the *Escherichia coli* enzyme biotin ligase (BirA) (42) and different tags recognized by BirA for site-specific biotinylation. The tags range from 123-amino acid (PSTCD) to a 15-amino acid peptide called the acceptor peptide (AP) (43-46). These tags have been fused to target proteins and site-specifically biotinylated *in vitro* and *in vivo* successfully.

The PSTCD is a domain of the 1.3S *Propionibacterium shermanii* transcarboxylase that is naturally biotinylated at lysine 89 of the domain. This PSTCD, when expressed in both *E.coli* and *Saccharomyces cerevisiae* as a fusion protein, is metabolically biotinylated by each organism's cellular biotin ligase enzyme. Chapman-Smith *et al.* (47) have fused the DNA sequence encoding the PSTCD to the carboxyl terminal of β-galactosidase, neomycin phosphotransferase, Tn5 and Tn903, and chloramphenicol acetyltransferase encoded by Tn9. All the fusion proteins encoded by this chimeric gene became biotinylated in *E.coli* as long as the biotin ligase recognition sequence remained intact. The use of this approach was expanded to include higher eukaryotic hosts such as cultured mammalian cells and mice (48).

The size of PSTCD was 123 amino acids. To maximize the usefulness of this biotin tag in mammalian cells, deletion constructs were created to reveal the minimum domain of the PSTCD that could be biotinylated. A series of deletion constructs of the PSTCD was made by inverse PCR of the original 123-amino acid biotinylated domain fused to the GFP and knob proteins (49, 50). They found that only the 70- and 63-amino-acid N-terminally truncated forms of the PSTCD domain were still biotinylated when fused to GFP (51).

Although the truncated PSTCD could be used as a tag for the site-specific modification of protein, some features of this approach including the introduction of a bulky (>75 residues) biotinyl domain, and incomplete biotinylation of the target protein, are less than ideal. The first of these issues was resolved by the successful screening of peptide libraries for small peptides capable of accepting biotin from *E.coli* BirA. Beckett *et al.* (43) reported that a numbers of peptides that serving as substrates for BirA-catalyzed biotinylation were identified from screens of four biased peptide libraries fused to the *lac* repressor. Among these a 23-mer peptide, MAGGLNDIFEAQ**K**IEWHEDTGGS, was as effective as the natural substrate in BirA-catalyzed biotin transfer.

To establish the minimal target for BirA-catalyzed biotinylation, a series of truncations of this 23-mer were synthesized and a 14-mer, GLNDIFEAQKIEWH, efficiently mimicked the biotin acceptor function of the much larger protein domain normally recognized by BirA (43, 44). Although this sequence worked well, Cull *et al* (44) recommended using a slightly extended 15-mer, termed AviTag (GLNDIFEAQKIEWHE) because it is consistently biotinylated at a rate slightly better than that of the natural substrate.

Ting *et al.* (52) fused AviTag (acceptor peptide) to CFP targeted to the surface with a signal peptide and the transmembrane (TM) domain of the platelet-derived growth factor receptor (AP-CFP-TM). After transfection of this construct into HeLa cells, addition of BirA to the cell medium together with ATP and biotin, biotinylation was detectable using fluorescent dyes or quantum dots (QD) conjugated streptavidin. The main advantage of AP is its small size, which means it is unlikely to perturb folding,

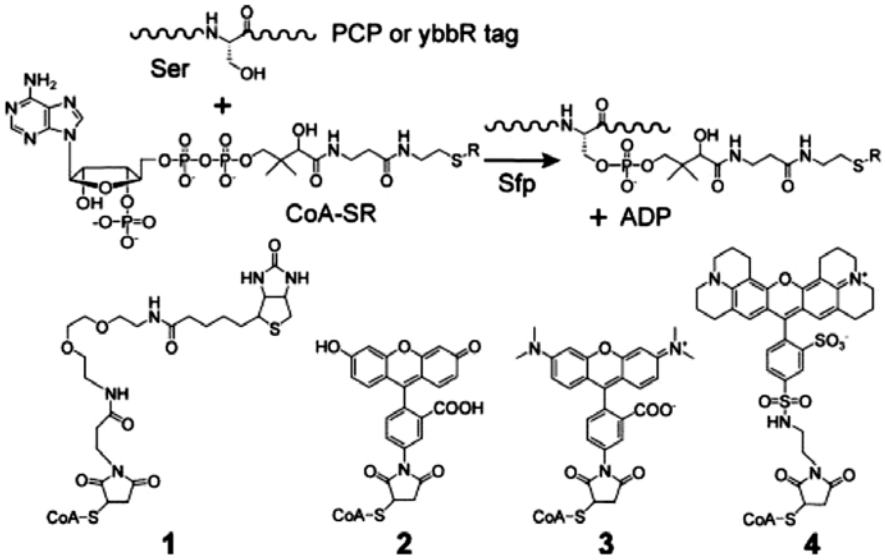


Figure 4. Sfp-catalyzed PCP or ybbR tag modification at a specific Ser residue by various small-molecule-CoA conjugates (CoA-SR): 1, biotin-CoA; 2, fluorescein-CoA; 3, tetramethylrhodamine-CoA; and 4, Texas red-CoA. Reproduced with permission from (67).

secretion, and protein–protein interactions during trafficking to the surface, compared with fluorescent proteins (GFP is 238 residues) (48, 53) and other proteins directing site-specific labeling such as O^6-alkylguanine-DNA alkyltransferase (54) or acyl carrier protein (55).

To expand the usefulness of enzymatic protein biotinylation, Ting *et al.* (56) screened the phage library and identified a 15-mer peptide sequence (termed AP2) which can be specifically recognized and biotinylated by Saccharomyces cerevisiae (yeast) biotin ligases (yBL), in the process discovering a new biotin ligase-AP pair, (BL2-AP2). This new pair has a similar specificity and kinetics to the original BirA-AP pair, but represents an orthogonal system, in that BirA does not recognize the new AP2 while conversely, the new BL2 does recognize the original AP. Such BirA-AP/BL2-AP2 orthogonal systems have been applied to perform two-color QD labeling on HeLa cells.

5.4. Sfp phosphopantethinyl transferase-based site-specific modification

The Sfp phosphopantetheinyl transferase covalently transfers 4'-phosphopantetheinyl (Ppant) groups from CoA to conserved serine residues on peptide carrier protein (PCP) and acyl carrier protein (ACP) domains in nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs) in *Bacillus subtilis* (57). The post-translational modification of PCP and ACP domains by Ppant is crucial for the activation of NRPS and PKS, because the flexible 20 Å Ppant group functions as a swinging arm providing the anchoring point for the attachment of the peptide or polyketide chain as it elongates along the biosynthetic assembly line (58).

Although the substrate for Sfp must contain CoA, the enzyme is otherwise quite tolerant of a range of small molecules attached to CoA. In fact, Sfp transfers small molecules of diverse structures from CoA to the conserved serine on the PCP or ACP domain (58, 59). Sfp-catalyzed post-translational modification of PCP or ACP takes advantage of the substrate promiscuity of Sfp and the small size of PCP and ACP domains (~80 residues) to provide an attractive method for one-step site-specific protein labeling with small molecules. The *E.coli* phosphopantetheinyl transferase AcpS is functionally similar to Sfp, although AcpS preferentially modifies ACP domains (60).

Sfp and AcpS have been used for site-specific protein labeling in cell lysates or on live cell surfaces. In these applications the target protein is expressed as a fusion to PCP or ACP at either the N or the C terminus, and Sfp is used to enzymatically attach a small-molecule probe - Ppant conjugate to the expressed PCP or ACP tag. This method has allowed the site-specific labeling of target proteins by small molecules of diverse structures and functionalities, such as biotin, fluorophores, porphyrins, sugars and peptides (Figure 4) (59, 61, 62). The labeling reaction proceeds with high efficiency and can be carried out in cell lysates or in culture media. A sufficient number of cell surface protein molecules can be labeled for imaging purposes within 15 min at submicromolar concentrations of the enzyme and CoA-conjugated small-molecule probes. Thus far, protein labeling catalyzed by Sfp and AcpS has been used for high-throughput printing of protein microarrays (63), site-specific attachment of small molecules to phage particles (64) and live cell imaging of cell surface proteins (55, 65, 66).

Most recently, the ybbR tag, a short (11-residue) peptide (DSLEFIASKLA), was found to be an efficient substrate for Sfp-catalyzed protein labeling, thereby replacing the full-length PCP or ACP domain for the construction of smaller fusions of the target protein. It was so named because part of its sequence was derived from the

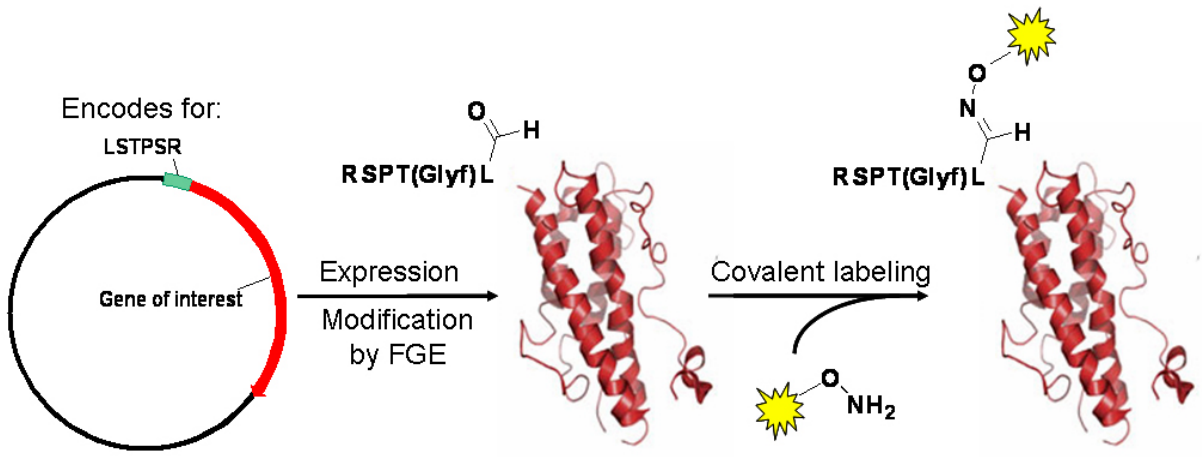


Figure 5. Nucleotides encoding the 6-amino-acid sequence can be appended to or inserted within a gene of interest. Upon expression, the encoded cysteine is modified to an aldehyde, which can be used as a chemical handle for a variety of applications such as fluorophore labeling. fGly, formylglycine. Reproduced with permission from (69).

ybbR ORF in the *B. subtilis* genome. The site of Sfp-catalyzed ybbR tag labeling was mapped to the underlined Ser residue, and the ybbR tag was found to have a strong tendency for adopting an α-helical conformation in solution. ybbR tag can be fused to the N or C terminus of target proteins or inserted in a flexible loop in the middle of a target protein for site-specific protein labeling by Sfp. The short size of the ybbR tag and its compatibility with various target proteins, the broad substrate specificity of Sfp for labeling the ybbR tag with small-molecule probes of diverse structures, and the high specificity and efficiency of the labeling reaction make Sfp-catalyzed ybbR tag labeling an attractive tool for expanding protein structural and functional diversities by posttranslational modification (67).

5.5. Formylglycine-generating enzyme-based site-specific modification Over a decade ago, Schmidt *et al.* (68) discovered that most sulfatases contain an aldehyde-bearing formylglycine residue within their active site that is required for catalytic activity. The formylglycine residue is produced by co- or post-translational modification of a conserved cysteine residue found within the "sulfatase motif" (Figure 5) (69). This motif contains a highly conserved CXPXR submotif, where X is usually serine, threonine, alanine or glycine. Formylglycine-generating enzyme (FGE) oxidizes cysteine to formylglycine and has recently been identified in eukaryotes (70) and prokaryotes. Interestingly, FGE recognizes its substrate mainly by its primary sequence. This means that a small peptide tag based on the sulfatase motif, called the “aldehyde tag”, might direct formation of formylglycine independent of its context. Furthermore, many organisms have endogenous FGE activities (71), suggesting that this post-translational modification system could serve as a general means for engineering proteins for site-specific labeling (72-74).

Carrico *et al.* (69) expressed recombinant proteins with either N- or C- terminal aldehyde tags in *E.coli*. The tags they used included a full-length sulfatase tag (LCTPSRGSLFTGR) and a 6-amino-acid tag (LCTPSR) containing only core conservative residues. These constructs were co-expressed with an additional FGE from *Mycobacterial tuberculosis* to maximize the formylglycine formation. After purification, a panel of proteins was robustly site-specific labeled with aminooxy- or hydrazide-functionalized probes, including fluorescent dyes, biotin, and poly (ethylene glycol) (PEG).

5.6. The hAGT based site-specific modification

Human O^6-alkylguanine-DNA alkyltransferases (hAGT) is a 207-amino acid enzyme. The physiological function of hAGT is the removal of alkyl groups from guanine bases of DNA to avoid mutations during cell division. O^6-Alkylguanine adducts are formed in DNA by many alkylating agents, including important environmental carcinogens such as nitrosamines. These adducts are repaired by a unique single step mechanism mediated by AGT (75-77). AGT transfers the alkyl group from O^6-alkylguanine to a cysteine residue located within the protein, and each molecule can act only once because the *S*-alkylcysteine formed is not regenerated. This transfer protects against the mutagenic and cytotoxic effects of alkylating agents AGT undergoes a classic suicide reaction as the cysteine residue in its active site forms a covalent and stable thioether bond with the alkyl group, allowing the release of the de-alkylated, original guanine base, and the content of AGT is a critical factor in determining cellular sensitivity to these agents (78, 79). The AGT family members share a highly conserved -PCHRV- amino acid sequence surrounding the cysteine acceptor site and contain a number of other conserved amino acids in the active site and DNA binding domain of the protein (Figure 6).

For AGT-based site-specific modification of protein (54), covalent labeling with a small molecule is achieved through hAGT, which irreversibly transfers the alkyl group from its substrate, O^6-alkylguanine-DNA, to one of its cysteine residues. This covalent and site-specific modification of AGT was studied in great detail. (80, 81). They observed that AGT can accept any chemical compound when it is linked to the guanine base through a benzyl group in different hosts such as E coli, yeast, and mammalian cell cultures.

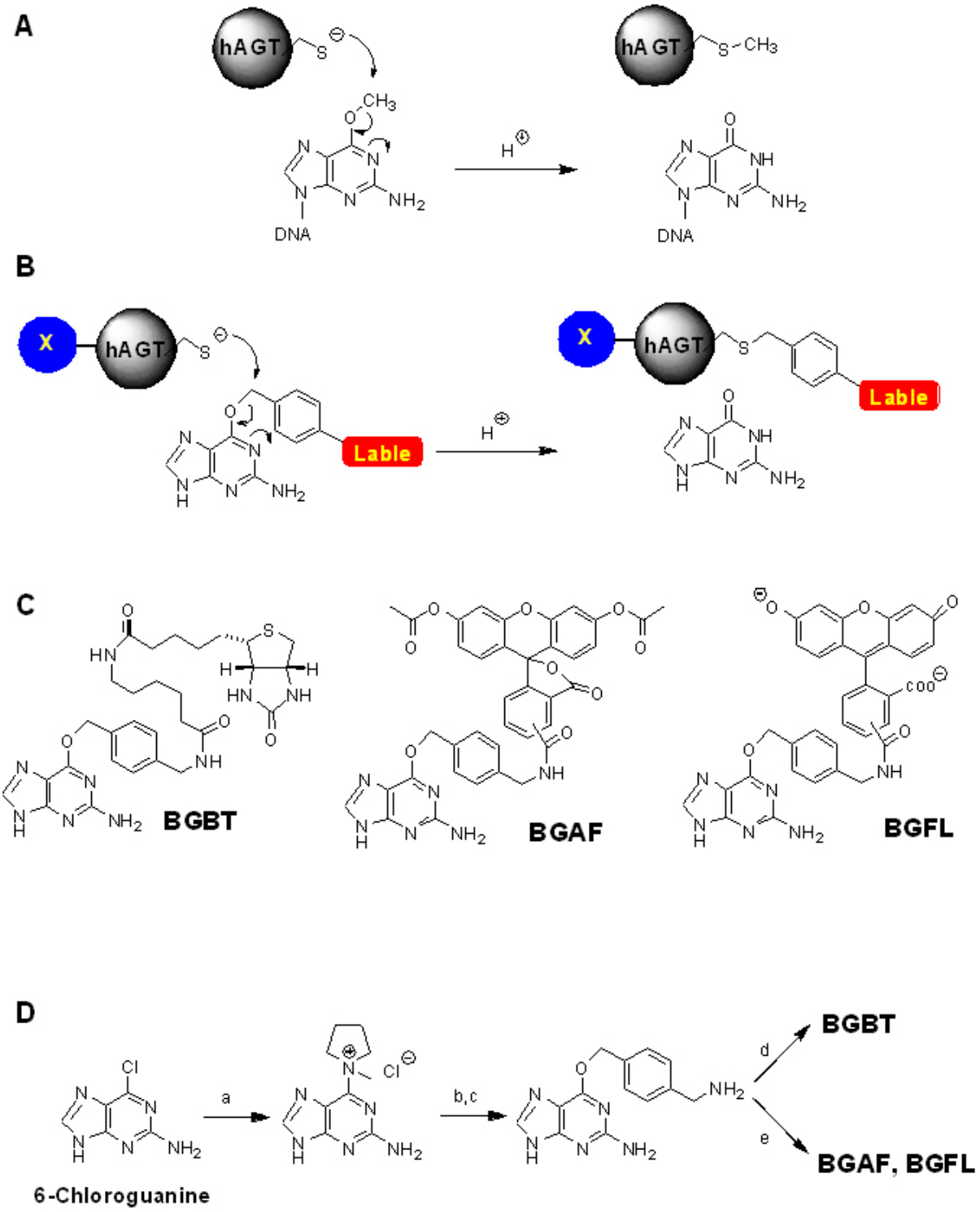


Figure 6. (A) Mechanism of DNA repair by hAGT; (B) Covalent labeling of an X-hAGT fusion protein using O^6-benzylguanine (BG) derivatives of type 1; (C) Structure of the BG derivatives BGBT, BGAF, and BGFL used in this work; (D) Scheme of synthesis of BGBT, BGAF, and BGFL. Reproduced with permission from (54).

The fact that the gene encoding hAGT can be expressed as a soluble protein in various host cells led to the development of the SNAP-tag™ technology (82, 83). Directed evolution experiments and protein engineering

with hAGT resulted in a smaller enzyme with much faster reaction kinetics and devoid of interactions with other biomolecules such as double stranded DNA. The reaction mechanism of the SNAP-tag™ system allows the covalent and specific attachment of target proteins to any desired compound or solid surface. For example, Tirat *et al.* (84) fused ubiquitin conjugating enzyme Rad6B to either N- or C- terminus of SNAP-tag™ using DNA recombinant technology. After labeling, the AGT-tagged Rad6B proteins were characterized by mass spectrometry, peptide map and a SDS-PAGE band shift assay, with all results demonstrating the high specificity of labeling by using the SNAP-tag™ technology. Furthermore, modified AGT-Rad6B fully retained its functional activity. They further tested an AGT substrate consisting of an Alexa555 fluorophore bound to benzylguanine without any spacer (85) and similar results were obtained. The major problem of this approach, however, is the rather long sequence of hAGT, which might have some effect on target protein conformation and function when used as a fusion tag.

5.7. Transglutaminase-based site-specific modification

The basic premise for this strategy lies in the specificity of the transglutaminase. Transglutaminase catalyzes the reaction:

$$R{-}CONH_2 + R'{-}NH_2 \overset{Ca^{2+}}{\rightleftharpoons} R{-}CONH{-}R' + NH_3$$

Here, R-$CONH_2$ represents the acceptor, a Gln residue in proteins, and R'-NH_2, the donor, an alkylamine; *in vivo*, the common amine donor is probably a protein-bound lysine side chain (86), and an amide bond cross-linkage is formed between the proteins. It appears that transglutaminase, especially guinea pig liver transglutaminase (TGase), has very stringent sequence specificity and structural requirements for the amine acceptor site around the Gln residues. According to the specific (87) determinant for proteins, substrate Gln residues for TGase must first satisfy an accessibility criterion, either by being in a highly flexible region of the polypeptide chain or by being clearly exposed to the solvent in a more structured region of the polypeptide chain. Due to the conformational requirement of transglutaminase-catalytic labeling, the direct Gln-specific modification or design of the mutation for the modification may be difficult to apply to most proteins.

Sato *et al.* (88) were able to site-specifically label IL-2 protein with dansylcadaverine by introducing a very short substrate sequence for TGase at the N-terminus of IL-2. The chimeric protein was expressed and then enzymatically labeled with the fluorophore only at a Gln site in the appended N-terminal sequence. A model system was constructed. The sequence selected for the preparation of the chimeric protein (named rTG1-IL-2) is PKPQQFM, derived from Substance P, which is known to be a very good substrate *in vitro* for TGase. The CD spectra and IL-2 bioactivity studies showed that the overall structure of rTG1-IL-2 was similar to that of rhIL-2. In fact, the addition of the full sequence of substance P (RPKPQQPGLM) to the N-terminus of hIL-2 revealed that almost 50% of the IL-2 bioactivity was lost on a molar basis. Thus, the study selected the essential part of the Substance P molecular consisting of the seven-residue sequence to avoid a conformational change of hIL-2.

5.8. Intein-mediated site-specific protein modification

Inteins are naturally occurring protein fragments that are capable of excising themselves out from a precursor protein and ligating the two flanking protein fragments (N-extein and C-extein), via a peptide bond, to form a mature protein. This process, termed intein-mediated protein slicing, occurs during the posttranslational stage of protein synthesis. Intein-mediated protein splicing is fascinating in that it is an autocatalytic event and requires neither cofactors nor auxiliary enzymes in the process (89).

Intein-mediated labeling reactions require two protein domains (90): the N-terminal domain which contains the protein sequence to be labeled (the target domain), and a C-terminal domain containing the intein (the intein domain). These domains are usually joined in a single fusion protein containing both the intein and target domains, a target-intein fusion protein. For those intein domains beginning with an N-terminal cysteine residue, the molecular basis of intein-mediated activation is a spontaneous reversible N→S transacylation step forming thioester that occurs at this cysteine residue. The ligation of probes to target domains has been achieved using a range of cysteine derivatives under appropriate conditions (91). Although the thioester is more reactive toward nucleophiles than a backbone amide, a direct reaction with peptides containing an N-terminal cysteine residue was unsuccessful, a result that has been rationalized in terms of steric hindrance from the bulky intein domain (92). The usage of several thiols including thiophenol (92), ethanethiol, and MESA (93) has been reported to enhance intein-mediated ligation reactions, presumably by trans-thioesterification to the less hindered intermediate. Under these conditions, cysteine derivatives are thought to form the kinetic product, then rearrange to the thermodynamically more stable amide (Figure 7) (89). The intein-mediated activation of the C-terminus of a protein domain has been shown to provide a general method for attaching a wide range of molecules, including proteins (49, 94), peptides (92), fluorescent labels (95, 96), carbohydrates (97), oligonucleotides (98), and affinity tags and metal chelators (99). Intein-mediated ligation has also been used for the assembly of selectively labeled proteins for NMR studies (99) and to prepare cyclic peptides and proteins (100, 101). For in-depth coverage of intern-mediated protein ligation and its application, readers are directed to a detailed review by Muralidharan and Muir (89). In our review, we will focus solely on the recent expanded applications of the IPL technology in the fields of bioimaging.

Tan *et al.* (102) developed strategies for site-specific biotinylation of proteins. In their strategy, the protein of interest was fused through its C-terminus to an intein, which had its other terminus fused to a chitin-binding domain (CBD). Upon expression in the host cell, the cell lysate containing the fusion protein was first loaded onto a column packed with chitin beads followed by the addition of a thiol cleaving reagent. The fusion protein underwent self-cleavage, catalyzed by the fused intein, and resulted in the protein having a reactive α-thioester group at the C-terminus. At this stage, the thiol side chain from the addition of cysteien-biotin attacked the α-thioester functionally and the resulting thioester-linked intermediate

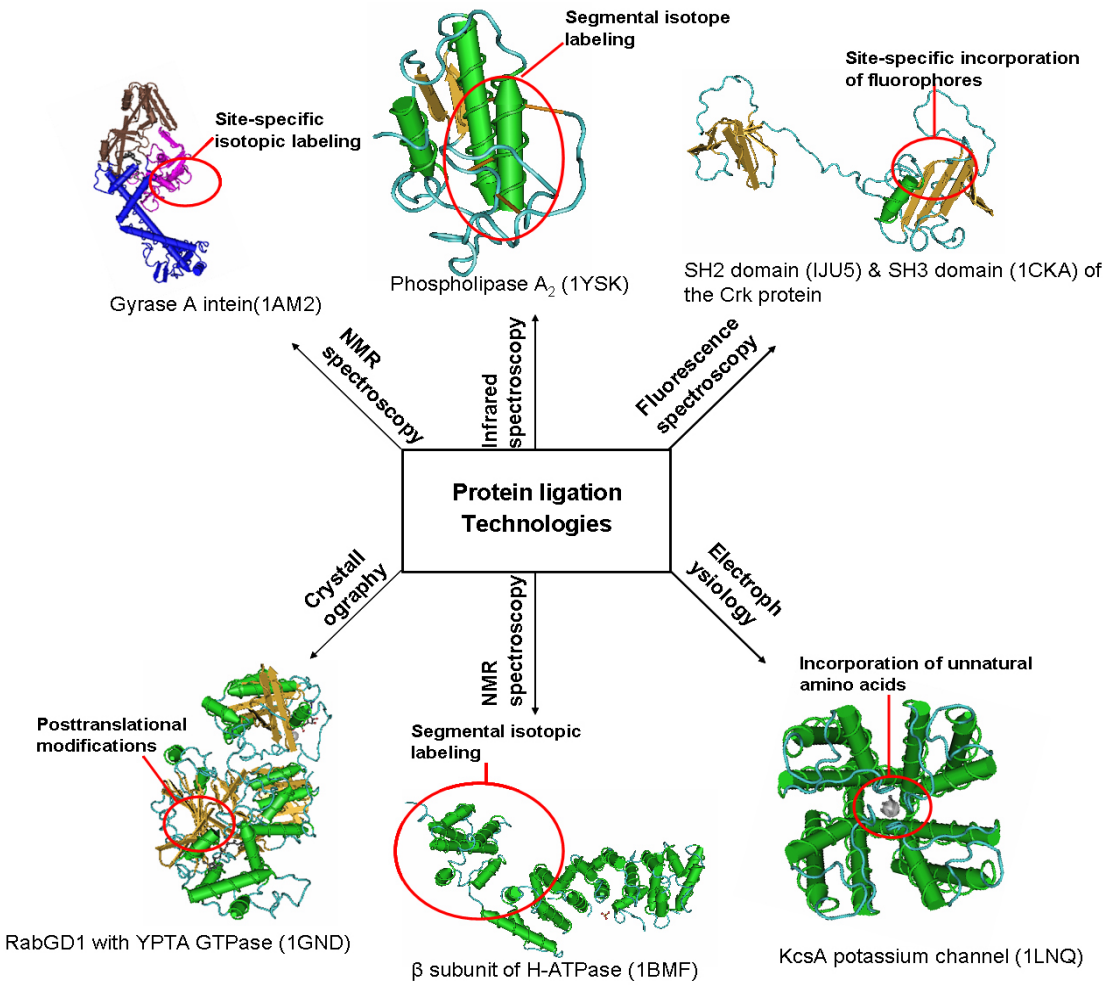


Figure 7. Selected applications of protein ligation technologies. Protein ligation techniques offer the ability to incorporate specific labels into proteins that have been used to assay structure and function, augmenting several traditional biophysical techniques. Reproduced with permission from (89).

underwent spontaneous rearrangement, allowing a native peptide bond to generate proteins that were site-specifically biotinyalated at their C-terminus. This strategy is simple yet efficient, and capable of biotinylating various proteins from different biological resources. This strategy has also been extended to *in vivo* biotinylation of proteins in bacterial and mammalian systems. The simple addition of the cell-permeable cysteine–biotin probe to the growth media followed by further incubation of the cells resulted in substantial biotinyaltion of the protein inside the cells. The efficiency of intein-mediated protein biotinylation depends greatly on the intein fused to target protein: as much as 2-10 times improvement in protein biotinylation may be achieved by the simple switch in the intein fused (102). However, the limitation of this strategy is the biotinylation of the endogenous protein in both mammalian cells and bacterial. From western blots, three other biotinylated proteins besides the targeted protein were detected from the lysate of mammalian cells and biotinylated acetyl-CoA carboxylase was found from bacterial lysate.

In fact, this intein-mediated protein biotinylation could be extended to cell-free protein system (103). A cell-free system has many potential advantages over traditional *in vivo* recombinant protein expression in cells. Cellular toxicity of overexpression of certain proteins, possible degradation by endogenous protease and formation of inclusion bodies of protein all might be avoided in a cell-free system. Although the cell-free biotinylation approach seems the most simple out of the three strategies mentioned here, the efficiency of protein expression and subsequent biotinylation depend on a few other factors including the DNA quality and concentration, as well as the nature of the target protein (104).

One potential problem with intein-mediated ligation chemistry is the possible oxidation of both the activating thiol reagent and the modified cysteine derivative to disulfides. The use of thiol reducing agents such as DTT is not possible, as these compounds effectively cleave and derivatize the target domain. This problem has been circumvented by the use of the water soluble phosphine

TCEP, which efficiently reduces disulfides to thiols. Another potential problem with intein-mediated protein modification is the relatively large size of intein. To date, most intein used were over 100 amino acid, for example Mxe GyrA intein , 198 aa (105); Sce Vma intein, 455 aa (106); Mtb RecA intein, 229 aa (107, 108); Mth RIRI intein, 134 aa (94); and Ssp DnaB intein (109), 154aa (109).

6. CONCLUSIONS AND PERSPECTIVES

Modern drug discovery strongly depends on the availability of target proteins in sufficient amounts and with desired properties. For some applications, proteins have to be equipped with additional functionalities for protein purification, fluorescent or radioactive labels for detection, glycosylation and phosphorylation for biological activity, and many more properties. This functionality can result from the expression of the protein as a fusion protein with an additional polypeptide, the labeling the protein with synthetic molecules, the incorporation of unnatural amino acids or a combination of these different approaches. For a protein modification technique to find widespread use among biologists, it must be practical, general and available. Appending a peptide tag to target protein fulfills all these requirements. Recombinant proteins have been expressed for decades as fusion proteins with an additional polypeptide for purification or characterization of target proteins. To extend the traditional use, the tag can also be used to label the fusion protein with synthetic probes in a site-specific and covalent manner.

This review summarizes current methods for site-specific modification of target protein. These methods have been widely used in proteomics, protein-protein interaction, protein purification and immobilization, and protein microarray studies, among others. However, only the cysteine-containing tag has been used for molecular imaging purpose. The AviTag can be foreseeably applied to produce site-specifically biotinylated proteins for labeling avidin-derivatized fluorophores, quantum dots, microbubbles, and chelators for optical, ultrasound, and radionuclide imaging. This strategy may also be extended by adding an ybbR tag onto the target protein to provide Sfp-catalyzed protein labeling of biotin, fluorophores, and metal chelators for molecular imaging applications. Furthermore, two or more orthogonal tags can be introduced into one protein and used to couple different detectable probes for multimodality imaging. In combining the convenience of expressing genetically tagged fusion proteins and the diversity of synthetic chemistry, site-specifically labeled protein probes will become ever more useful and popular in the molecular imaging area.

7. REFERENCES

1. Mellstedt, H.: Monoclonal antibodies in human cancer. *Medicamentos de actualidad*, 39 Suppl C, 1-16 (2003)
2. Radford, I. R.: Imatinib. Novartis. *Curr Opin Investig Drugs*, 3, 492-9 (2002)
3. Kelloff, G. J., K. A. Krohn, S. M. Larson, R. Weissleder, D. A. Mankoff, J. M. Hoffman, J. M. Link, K. Z. Guyton, W. C. Eckelman, H. I. Scher, J. O'Shaughnessy, B. D. Cheson, C. C. Sigman, J. L. Tatum, G. Q. Mills, D. C. Sullivan & J. Woodcock: The progress and promise of molecular imaging probes in oncologic drug development. *Clin Cancer Res*, 11, 7967-85 (2005)
4. Antos, J. M. & M. B. Francis: Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J Am Chem Soc*, 126, 10256-7 (2004)
5. Joshi, N. S., L. R. Whitaker & M. B. Francis: A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J Am Chem Soc*, 126, 15942-3 (2004)
6. Schlick, T. L., Z. Ding, E. W. Kovacs & M. B. Francis: Dual-surface modification of the tobacco mosaic virus. *J Am Chem Soc*, 127, 3718-23 (2005)
7. Bock, A., K. Forchhammer, J. Heider, W. Leinfelder, G. Sawers, B. Veprek & F. Zinoni: Selenocysteine: the 21st amino acid. *Mol Microbiol*, 5, 515-20 (1991)
8. Srinivasan, G., C. M. James & J. A. Krzycki: Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science*, 296, 1459-62 (2002)
9. Xie, J. & P. G. Schultz: A chemical toolkit for proteins--an expanded genetic code. *Nat Rev Mol Cell Biol*, 7, 775-82 (2006)
10. Ohno, S., M. Matsui, T. Yokogawa, M. Nakamura, T. Hosoya, T. Hiramatsu, M. Suzuki, N. Hayashi & K. Nishikawa: Site-selective post-translational modification of proteins using an unnatural amino acid, 3-azidotyrosine. *J Biochem (Tokyo)*, 141, 335-43 (2007)
11. Wingfield, P., P. Graber, A. R. Shaw, A. M. Gronenborn, G. M. Clore & H. R. MacDonald: Preparation, characterization and application of interleukin-1 beta mutant proteins with surface-accessible cysteine residues. *Eur J Biochem*, 179, 565-71 (1989)
12. Olafsen, T., C. W. Cheung, P. J. Yazaki, L. Li, G. Sundaresan, S. S. Gambhir, M. A. Sherman, L. E. Williams, J. E. Shively, A. A. Raubitschek & A. M. Wu: Covalent disulfide-linked anti-CEA diabody allows site-specific conjugation and radiolabeling for tumor targeting applications. *Protein Eng Des Sel*, 17, 21-7 (2004)
13. Wu, A. M., P. J. Yazaki, S. Tsai, K. Nguyen, A. L. Anderson, D. W. McCarthy, M. J. Welch, J. E. Shively, L. E. Williams, A. A. Raubitschek, J. Y. Wong, T. Toyokuni, M. E. Phelps & S. S. Gambhir: High-resolution microPET imaging of carcinoembryonic antigen-positive xenografts by using a copper-64-labeled engineered antibody fragment. *Proc Natl Acad Sci U S A*, 97, 8495-500 (2000)
14. Natarajan, A., C. Y. Xiong, H. Albrecht, G. L. DeNardo & S. J. DeNardo: Characterization of site-specific ScFv PEGylation for tumor-targeting pharmaceuticals. *Bioconjug Chem*, 16, 113-21 (2005)
15. Blankenberg, F. G., M. V. Backer, Z. Levashova, V. Patel & J. M. Backer: *In vivo* tumor angiogenesis imaging with site-specific labeled (99m)Tc-HYNIC-VEGF. *Eur J Nucl Med Mol Imaging*, 33, 841-8 (2006)
16. Backer, M. V., T. I. Gaynutdinov, V. Patel, A. K. Bandyopadhyaya, B. T. Thirumamagal, W. Tjarks, R. F. Barth, K. Claffey & J. M. Backer: Vascular endothelial growth factor selectively targets boronated dendrimers to tumor vasculature. *Mol Cancer Ther*, 4, 1423-9 (2005)
17. Backer, M. V., V. Patel, B. T. Jehning, K. P. Claffey & J. M. Backer: Surface immobilization of active vascular endothelial growth factor via a cysteine-containing tag. *Biomaterials*, 27, 5452-8 (2006)

18. Backer, M. V., Z. Levashova, V. Patel, B. T. Jehning, K. Claffey, F. G. Blankenberg & J. M. Backer: Molecular imaging of VEGF receptors in angiogenic vasculature with single-chain VEGF-based probes. *Nat Med*, 13, 504-9 (2007)
19. Shi, Y. & J. M. Berg: Specific DNA-RNA hybrid binding by zinc finger proteins. *Science*, 268, 282-4 (1995)
20. Michael, S. F., V. J. Kilfoil, M. H. Schmidt, B. T. Amann & J. M. Berg: Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proc Natl Acad Sci U S A*, 89, 4796-800 (1992)
21. Smith, J. J., D. W. Conrad, M. J. Cuneo & H. W. Hellinga: Orthogonal site-specific protein modification by engineering reversible thiol protection mechanisms. *Protein Sci*, 14, 64-73 (2005)
22. Folk, J. E.: Transglutaminases. *Annu Rev Biochem*, 49, 517-31 (1980)
23. Jager, M., E. Nir & S. Weiss: Site-specific labeling of proteins for single-molecule FRET by combining chemical and enzymatic modification. *Protein Sci*, 15, 640-6 (2006)
24. Stalteri, M. A., S. Bansal, R. Hider & S. J. Mather: Comparison of the stability of technetium-labeled peptides to challenge with cysteine. *Bioconjug Chem*, 10, 130-6 (1999)
25. Nedelman, M. A., D. J. Shealy, R. Boulin, E. Brunt, J. I. Seasholtz, I. E. Allen, J. E. McCartney, F. D. Warren, H. Oppermann, R. H. Pang & *et al.*: Rapid infarct imaging with a technetium-99m-labeled antimyosin recombinant single-chain Fv: evaluation in a canine model of acute myocardial infarction. *J Nucl Med*, 34, 234-41 (1993)
26. Bridger, G. J., M. J. Abrams, S. Padmanabhan, F. Gaul, S. Larsen, G. W. Henson, D. A. Schwartz, C. B. Longley, C. A. Burton & M. E. Ultee: A comparison of cleavable and noncleavable hydrazinopyridine linkers for the 99mTc labeling of Fab' monoclonal antibody fragments. *Bioconjug Chem*, 7, 255-64 (1996)
27. Mease, R. C. & C. Lambert: Newer methods of labeling diagnostic agents with Tc-99m. *Semin Nucl Med*, 31, 278-85 (2001)
28. George, A. J., F. Jamar, M. S. Tai, B. T. Heelan, G. P. Adams, J. E. McCartney, L. L. Houston, L. M. Weiner, H. Oppermann & A. M. Peters: Radiometal labeling of recombinant proteins by a genetically engineered minimal chelation site: technetium-99m coordination by single-chain Fv antibody fusion proteins through a C-terminal cysteinyl peptide. *Proc Natl Acad Sci U S A*, 92, 8358-62 (1995)
29. Berndorff, D., S. Borkowski, D. Moosmayer, F. Viti, B. Muller-Tiemann, S. Sieger, M. Friebe, C. S. Hilger, L. Zardi, D. Neri & L. M. Dinkelborg: Imaging of tumor angiogenesis using 99mTc-labeled human recombinant anti-ED-B fibronectin antibody fragments. *J Nucl Med*, 47, 1707-16 (2006)
30. Bogdanov, A., Jr., M. Simonova & R. Weissleder: Design of metal-binding green fluorescent protein variants. *Biochim Biophys Acta*, 1397, 56-64 (1998)
31. Tait, J. F., C. Smith & F. G. Blankenberg: Structural requirements for *in vivo* detection of cell death with 99mTc-annexin V. *J Nucl Med*, 46, 807-15 (2005)
32. Wang, Y., X. Liu & D. J. Hnatowich: An improved synthesis of NHS-MAG3 for conjugation and radiolabeling of biomolecules with (99m)Tc at room temperature. *Nat Protoc*, 2, 972-8 (2007)
33. Seetharaman, S., M. H. Sosabowski & J. R. Ballinger: Validation of alternative methods of preparing (99m) Tc-MAG3. *Appl Radiat Isot* (2007)
34. Engfeldt, T., A. Orlova, T. Tran, A. Bruskin, C. Widstrom, A. E. Karlstrom & V. Tolmachev: Imaging of HER2-expressing tumours using a synthetic Affibody molecule containing the (99m)Tc-chelating mercaptoacetyl-glycyl-glycyl-glycyl (MAG3) sequence. *Eur J Nucl Med Mol Imaging*, 34, 722-33 (2007)
35. Vanderheyden, J. L., G. Liu, J. He, B. Patel, J. F. Tait & D. J. Hnatowich: Evaluation of 99mTc-MAG3-annexin V: influence of the chelate on *in vitro* and *in vivo* properties in mice. *Nucl Med Biol*, 33, 135-44 (2006)
36. Hu, J., Y. X. Zhang, X. L. Lan, G. M. Qin, J. Zhang & Z. H. Hu: An imaging study using laminin peptide 99mTc-YIGSR in mice bearing Ehrlich ascites tumour. *Chin Med J (Engl)*, 118, 753-8 (2005)
37. Waibel, R., R. Alberto, J. Willuda, R. Finnern, R. Schibli, A. Stichelberger, A. Egli, U. Abram, J. P. Mach, A. Pluckthun & P. A. Schubiger: Stable one-step technetium-99m labeling of His-tagged recombinant proteins with a novel Tc (I)-carbonyl complex. *Nat Biotechnol*, 17, 897-901 (1999)
38. Griffin, B. A., S. R. Adams & R. Y. Tsien: Specific covalent labeling of recombinant protein molecules inside live cells. *Science*, 281, 269-72 (1998)
39. Adams, S. R., R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis & R. Y. Tsien: New biarsenical ligands and tetracysteine motifs for protein labeling *in vitro* and *in vivo*: synthesis and biological applications. *J Am Chem Soc*, 124, 6063-76 (2002)
40. Martin, B. R., B. N. Giepmans, S. R. Adams & R. Y. Tsien: Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat Biotechnol*, 23, 1308-14 (2005)
41. Gravel, R. A. & M. A. Narang: Molecular genetics of biotin metabolism: old vitamin, new science. *J Nutr Biochem*, 16, 428-31 (2005)
42. Nenortas, E. & D. Beckett: Purification and characterization of intact and truncated forms of the Escherichia coli biotin carboxyl carrier subunit of acetyl-CoA carboxylase. *J Biol Chem*, 271, 7559-67 (1996)
43. Beckett, D., E. Kovaleva & P. J. Schatz: A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci*, 8, 921-9 (1999)
44. Cull, M. G. & P. J. Schatz: Biotinylation of proteins *in vivo* and *in vitro* using small peptide tags. *Methods Enzymol*, 326, 430-40 (2000)
45. Smith, P. A., B. C. Tripp, E. A. DiBlasio-Smith, Z. Lu, E. R. LaVallie & J. M. McCoy: A plasmid expression system for quantitative *in vivo* biotinylation of thioredoxin fusion proteins in Escherichia coli. *Nucleic Acids Res*, 26, 1414-20 (1998)
46. Sibler, A. P., E. Kempf, A. Glacet, G. Orfanoudakis, D. Bourel & E. Weiss: *In vivo* biotinylated recombinant antibodies: high efficiency of labelling and application to the cloning of active anti-human IgG1 Fab fragments. *J Immunol Methods*, 224, 129-40 (1999)

47. Chapman-Smith, A., T. W. Morris, J. C. Wallace & J. E. Cronan, Jr.: Molecular recognition in a post-translational modification of exceptional specificity. Mutants of the biotinylated domain of acetyl-CoA carboxylase defective in recognition by biotin protein ligase. *J Biol Chem*, 274, 1449-57 (1999)
48. Lisenbee, C. S., S. K. Karnik & R. N. Trelease: Overexpression and mislocalization of a tail-anchored GFP redefines the identity of peroxisomal ER. *Traffic*, 4, 491-501 (2003)
49. Schwarzer, D. & P. A. Cole: Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr Opin Chem Biol*, 9, 561-9 (2005)
50. Leon-Del-Rio, A. & R. A. Gravel: Sequence requirements for the biotinylation of carboxyl-terminal fragments of human propionyl-CoA carboxylase alpha subunit expressed in Escherichia coli. *J Biol Chem*, 269, 22964-8 (1994)
51. Wojda, U., P. Goldsmith & J. L. Miller: Surface membrane biotinylation efficiently mediates the endocytosis of avidin bioconjugates into nucleated cells. *Bioconjug Chem*, 10, 1044-50 (1999)
52. Howarth, M., K. Takao, Y. Hayashi & A. Y. Ting: Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proc Natl Acad Sci U S A*, 102, 7583-8 (2005)
53. Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman & E. K. O'Shea: Global analysis of protein localization in budding yeast. *Nature*, 425, 686-91 (2003)
54. Keppler, A., S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel & K. Johnsson: A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol*, 21, 86-9 (2003)
55. George, N., H. Pick, H. Vogel, N. Johnsson & K. Johnsson: Specific labeling of cell surface proteins with chemically diverse compounds. *J Am Chem Soc*, 126, 8896-7 (2004)
56. Chen, I., Y. A. Choi & A. Y. Ting: Phage display evolution of a peptide substrate for yeast biotin ligase and application to two-color quantum dot labeling of cell surface proteins. *J Am Chem Soc*, 129, 6619-25 (2007)
57. Mofid, M. R., R. Finking, L. O. Essen & M. A. Marahiel: Structure-based mutational analysis of the 4'-phosphopantetheinyl transferases Sfp from Bacillus subtilis: carrier protein recognition and reaction mechanism. *Biochemistry*, 43, 4128-36 (2004)
58. Copp, J. N. & B. A. Neilan: The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Appl Environ Microbiol*, 72, 2298-305 (2006)
59. Mofid, M. R., R. Finking & M. A. Marahiel: Recognition of hybrid peptidyl carrier proteins/acyl carrier proteins in nonribosomal peptide synthetase modules by the 4'-phosphopantetheinyl transferases AcpS and Sfp. *J Biol Chem*, 277, 17023-31 (2002)
60. Belshaw, P. J., C. T. Walsh & T. Stachelhaus: Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science*, 284, 486-9 (1999)
61. Zhou, Z., P. Cironi, A. J. Lin, Y. Xu, S. Hrvatin, D. E. Golan, P. A. Silver, C. T. Walsh & J. Yin: Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. *ACS Chem Biol*, 2, 337-46 (2007)
62. Yin, J., A. J. Lin, D. E. Golan & C. T. Walsh: Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc*, 1, 280-5 (2006)
63. Yin, J., F. Liu, X. Li & C. T. Walsh: Labeling proteins with small molecules by site-specific posttranslational modification. *J Am Chem Soc*, 126, 7754-5 (2004)
64. Yin, J., F. Liu, M. Schinke, C. Daly & C. T. Walsh: Phagemid encoded small molecules for high throughput screening of chemical libraries. *J Am Chem Soc*, 126, 13570-1 (2004)
65. Vivero-Pol, L., N. George, H. Krumm, K. Johnsson & N. Johnsson: Multicolor imaging of cell surface proteins. *J Am Chem Soc*, 127, 12770-1 (2005)
66. Yin, J., A. J. Lin, P. D. Buckett, M. Wessling-Resnick, D. E. Golan & C. T. Walsh: Single-cell FRET imaging of transferrin receptor trafficking dynamics by Sfp-catalyzed, site-specific protein labeling. *Chem Biol*, 12, 999-1006 (2005)
67. Yin, J., P. D. Straight, S. M. McLoughlin, Z. Zhou, A. J. Lin, D. E. Golan, N. L. Kelleher, R. Kolter & C. T. Walsh: Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci U S A*, 102, 15815-20 (2005)
68. Schmidt, B., T. Selmer, A. Ingendoh & K. von Figura: A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. *Cell*, 82, 271-8 (1995)
69. Carrico, I. S., B. L. Carlson & C. R. Bertozzi: Introducing genetically encoded aldehydes into proteins. *Nat Chem Biol*, 3, 321-2 (2007)
70. Cosma, M. P., S. Pepe, I. Annunziata, R. F. Newbold, M. Grompe, G. Parenti & A. Ballabio: The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell*, 113, 445-56 (2003)
71. Landgrebe, J., T. Dierks, B. Schmidt & K. von Figura: The human SUMF1 gene, required for posttranslational sulfatase modification, defines a new gene family which is conserved from pro- to eukaryotes. *Gene*, 316, 47-56 (2003)
72. Sardiello, M., I. Annunziata, G. Roma & A. Ballabio: Sulfatases and sulfatase modifying factors: an exclusive and promiscuous relationship. *Hum Mol Genet*, 14, 3203-17 (2005)
73. Dickmanns, A., B. Schmidt, M. G. Rudolph, M. Mariappan, T. Dierks, K. von Figura & R. Ficner: Crystal structure of human pFGE, the paralog of the Calpha-formylglycine-generating enzyme. *J Biol Chem*, 280, 15180-7 (2005)
74. Roeser, D., A. Preusser-Kunze, B. Schmidt, K. Gasow, J. G. Wittmann, T. Dierks, K. von Figura & M. G. Rudolph: A general binding mechanism for all human sulfatases by the formylglycine-generating enzyme. *Proc Natl Acad Sci U S A*, 103, 81-6 (2006)
75. Lindahl, T., B. Sedgwick, M. Sekiguchi & Y. Nakabeppu: Regulation and expression of the adaptive response to alkylating agents. *Annu Rev Biochem*, 57, 133-57 (1988)
76. Mitra, S. & B. Kaina: Regulation of repair of alkylation damage in mammalian genomes. *Prog Nucleic Acid Res Mol Biol*, 44, 109-42 (1993)
77. Pegg, A. E., M. E. Dolan & R. C. Moschel: Structure, function, and inhibition of O6-alkylguanine-DNA alkyltransferase. *Prog Nucleic Acid Res Mol Biol*, 51, 167-223 (1995)

78. Crone, T. M., K. Goodtzova, S. Edara & A. E. Pegg: Mutations in human O6-alkylguanine-DNA alkyltransferase imparting resistance to O6-benzylguanine. *Cancer Res*, 54, 6221-7 (1994)
79. Gudtsova, K. V., G. V. Kukushkina & L. B. Gorbacheva: [Biochemical mechanisms of resistance to a new antineoplastic drug CRC 680578 from the nitrosourea class]. *Biokhimiia*, 56, 1509-21 (1991)
80. Margison, G. P. & M. F. Santibanez-Koref: O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *Bioessays*, 24, 255-66 (2002)
81. Pegg, A. E.: Repair of O (6)-alkylguanine by alkyltransferases. *Mutat Res*, 462, 83-100 (2000)
82. Gronemeyer, T., C. Chidley, A. Juillerat, C. Heinis & K. Johnsson: Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng Des Sel*, 19, 309-16 (2006)
83. Juillerat, A., T. Gronemeyer, A. Keppler, S. Gendreizig, H. Pick, H. Vogel & K. Johnsson: Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules *in vivo*. *Chem Biol*, 10, 313-7 (2003)
84. Tirat, A., F. Freuler, T. Stettler, L. M. Mayr & L. Leder: Evaluation of two novel tag-based labelling technologies for site-specific modification of proteins. *Int J Biol Macromol*, 39, 66-76 (2006)
85. Juillerat, A., C. Heinis, I. Sielaff, J. Barnikow, H. Jaccard, B. Kunz, A. Terskikh & K. Johnsson: Engineering substrate specificity of O6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells. *Chembiochem*, 6, 1263-9 (2005)
86. Yan, S. B. & F. Wold: Neoglycoproteins: *in vitro* introduction of glycosyl units at glutamines in beta-casein using transglutaminase. *Biochemistry*, 23, 3759-65 (1984)
87. Coussons, P. J., S. M. Kelly, N. C. Price, C. M. Johnson, B. Smith & L. Sawyer: Selective modification by transglutaminase of a glutamine side chain in the hinge region of the histidine-388----glutamine mutant of yeast phosphoglycerate kinase. *Biochem J*, 273 (Pt 1), 73-8 (1991)
88. Sato, H., M. Ikeda, K. Suzuki & K. Hirayama: Site-specific modification of interleukin-2 by the combined use of genetic engineering techniques and transglutaminase. *Biochemistry*, 35, 13072-80 (1996)
89. Muralidharan, V. & T. W. Muir: Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat Methods*, 3, 429-38 (2006)
90. Tan, L. P. & S. Q. Yao: Intein-mediated, *in vitro* and *in vivo* protein modifications with small molecules. *Protein Pept Lett*, 12, 769-75 (2005)
91. Muir, T. W., D. Sondhi & P. A. Cole: Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A*, 95, 6705-10 (1998)
92. Ayers, B., U. K. Blaschke, J. A. Camarero, G. J. Cotton, M. Holford & T. W. Muir: Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers*, 51, 343-54 (1999)
93. Dawson, P. E., T. W. Muir, I. Clark-Lewis & S. B. Kent: Synthesis of proteins by native chemical ligation. *Science*, 266, 776-9 (1994)
94. Evans, T. C., Jr., J. Benner & M. Q. Xu: The *in vitro* ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum. *J Biol Chem*, 274, 3923-6 (1999)
95. Muir, T. W.: Semisynthesis of proteins by expressed protein ligation. *Annu Rev Biochem*, 72, 249-89 (2003)
96. Muralidharan, V., J. Cho, M. Trester-Zedlitz, L. Kowalik, B. T. Chait, D. P. Raleigh & T. W. Muir: Domain-specific incorporation of noninvasive optical probes into recombinant proteins. *J Am Chem Soc*, 126, 14004-12 (2004)
97. Ma, J. & C. L. Cooney: Application of vortex flow adsorption technology to intein-mediated recovery of recombinant human alpha1-antitrypsin. *Biotechnol Prog*, 20, 269-76 (2004)
98. Lovrinovic, M., R. Seidel, R. Wacker, H. Schroeder, O. Seitz, M. Engelhard, R. S. Goody & C. M. Niemeyer: Synthesis of protein-nucleic acid conjugates by expressed protein ligation. *Chem Commun (Camb)*822-3 (2003)
99. Xu, R., B. Ayers, D. Cowburn & T. W. Muir: Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci U S A*, 96, 388-93 (1999)
100. Iwai, H., A. Lingel & A. Pluckthun: Cyclic green fluorescent protein produced *in vivo* using an artificially split PI-PfuI intein from Pyrococcus furiosus. *J Biol Chem*, 276, 16548-54 (2001)
101. Abel-Santos, E., C. P. Scott & S. J. Benkovic: Use of inteins for the *in vivo* production of stable cyclic peptide libraries in E. coli. *Methods Mol Biol*, 205, 281-94 (2003)
102. Tan, L. P., R. Y. Lue, G. Y. Chen & S. Q. Yao: Improving the intein-mediated, site-specific protein biotinylation strategies both *in vitro* and *in vivo*. *Bioorg Med Chem Lett*, 14, 6067-70 (2004)
103. Lue, R. Y., G. Y. Chen, Y. Hu, Q. Zhu & S. Q. Yao: Versatile protein biotinylation strategies for potential high-throughput proteomics. *J Am Chem Soc*, 126, 1055-62 (2004)
104. Wood, R. J., D. D. Pascoe, Z. K. Brown, E. M. Medlicott, M. Kriek, C. Neylon & P. L. Roach: Optimized conjugation of a fluorescent label to proteins via intein-mediated activation and ligation. *Bioconjug Chem*, 15, 366-72 (2004)
105. Southworth, M. W., K. Amaya, T. C. Evans, M. Q. Xu & F. B. Perler: Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques*, 27, 110-4, 116, 118-20 (1999)
106. Chong, S., F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus & M. Q. Xu: Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene*, 192, 271-81 (1997)
107. Wood, D. W., W. Wu, G. Belfort, V. Derbyshire & M. Belfort: A genetic system yields self-cleaving inteins for bioseparations. *Nat Biotechnol*, 17, 889-92 (1999)
108. Shingledecker, K., S. Q. Jiang & H. Paulus: Molecular dissection of the Mycobacterium tuberculosis RecA intein: design of a minimal intein and of a trans-splicing system involving two intein fragments. *Gene*, 207, 187-95 (1998)
109. Mathys, S., T. C. Evans, I. C. Chute, H. Wu, S. Chong, J. Benner, X. Q. Liu & M. Q. Xu: Characterization of a self-splicing mini-intein and its conversion into

autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene*, 231, 1-13 (1999)

Abbreviations: PET: positron emission tomography, SPECT:single photon emission computed tomography, mMRI: molecular magnetic resonance imaging , CEU:contrast enhanced ultrasound, FRET: fluorescence resonance energy transfer, Cys: cysteine, His: histidine, DOTA: 1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetraacetic acid, RTP: reversible thiol protection, MRTP: metal-dependent reversible thiol protection , RRTP: redox-dependent reversible thiol protection, FN: fibronectin, VEGF: vascular endothelial growth factor, TMR: tetramethylrhodamine-5-maleimide, IAF: 5-iodoacetamide fluorescence, GBP: glucose-binding protein, MBP: maltose-binding protein, TGase: Transglutaminase, Gln: glutamine, Pro: praline, Lys: lysine, Phe: phenylalnine, BHS: biotin holoenzyme synthetases, BirA: Escherichia coli enzyme biotin ligase, yBL: yeast biotin ligase\, Ppant: phosphopantetheinyl, PCP: peptide carrier protein, ACP: acyl carrier protein, NRPSs: nonribosomal peptide synthetases, PKSs: polyketide synthases, FGE: Formylglycine-generating enzyme, hAGT: Human O6-alkylguanine-DNA alkyltransferases, CBD: chitin-binding domain

Key Words: Site-Specific Modification, Fusion Protein, Molecular Imaging, Cys-Tag, Chelator-Tag, Biotin Ligase (Bira), Sfp Phosphopantetheinyl Transferase, Formylglycine-generating enzyme, Human O^6-alkylguanine-DNA alkyltransferases (hAGT), Transglutaminase, Intein, Review

Send correspondence to: Xiaoyuan Chen, PhD, the Molecular Imaging Program at Stanford (MIPS), Department of Radiology and Bio-X Program, Stanford University School of Medicine, 1201 Welch Rd, P095, Stanford, CA 94305, Tel: 650-725-0950, Fax: 650-736-7925, E-mail: shawchen@stanford.edu

http://www.bioscience.org/current/vol13.htm