JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

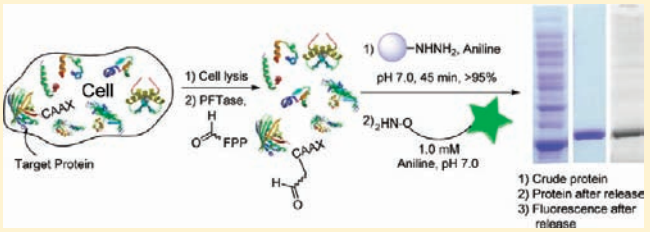
Chemoenzymatic Reversible Immobilization and Labeling of Proteins without Prior Purification

Mohammad Rashidian, James M. Song, Rachel E. Pricer, and Mark D. Distefano*

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55454, United States

Supporting Information

ABSTRACT: Site-specific chemical modification of proteins is important for many applications in biology and biotechnology. Recently, our laboratory and others have exploited the high specificity of the enzyme protein farnesyltransferase (PFTase) to site-specifically modify proteins through the use of alternative substrates that incorporate bioorthogonal functionality including azides and alkynes. In this study, we evaluate two aldehyde-containing molecules as substrates for PFTase and as reactants in both oxime and hydrazone formation. Using green fluorescent protein (GFP) as a model system, we demonstrate that the purified protein can be enzymatically modified with either analogue to yield aldehyde-functionalized proteins. Oxime or hydrazone formation was then employed to immobilize, fluorescently label, or PEGylate the resulting aldehyde-containing proteins. Immobilization via hydrazone formation was also shown to be reversible via transoximization with a fluorescent alkoxyamine. After characterizing this labeling strategy using pure protein, the specificity of the enzymatic process was used to selectively label GFP present in crude *E. coli* extract followed by capture of the aldehyde-modified protein using hydrazide−agarose. Subsequent incubation of the immobilized protein using a fluorescently labeled or PEGylated alkoxyamine resulted in the release of pure GFP containing the desired site-specific covalent modifications. This procedure was also employed to produce PEGylated glucose-dependent insulinotropic polypeptide (GIP), a protein with potential therapeutic activity for diabetes. Given the specificity of the PFTase-catalyzed reaction coupled with the ability to introduce a CAAX-box recognition sequence onto almost any protein, this method shows great potential as a general approach for selective immobilization and labeling of recombinant proteins present in crude cellular extract without prior purification. Beyond generating site-specifically modified proteins, this approach for polypeptide modification could be particularly useful for large-scale production of protein conjugates for therapeutic or industrial applications.



INTRODUCTION

Site-specific chemical modification of proteins is important for many applications in biology and biotechnology. It can facilitate studies of proteins with respect to their structure, folding, and interaction with other proteins in both biochemical and cellular investigations. In particular, in many biotechnology applications, the oriented (i.e., site-specific) covalent attachment of proteins to surfaces is important because it ensures homogeneous surface coverage and accessibility to the active site of the protein.[1−8] Protein immobilization is an important first step for many applications including construction of biosensors and protein microarrays, development of immunoassay methods, and employment of enzymes in biotechnology procedures.[9−11] Similarly, site-specific protein labeling is essential for a variety of applications ranging from introduction of fluorophores for biophysical studies to preparation of protein−polymer conjugates for medical applications.[12−17]

Importantly, the structural sensitivity of polypeptides calls for chemical transformations that proceed under mild conditions and are compatible with all functional groups present therein. However, such modification is challenging because of the large number of reactive functional groups typically present in polypeptides. Although many existing chemical reactions are applicable in principle, development of new methods for site-specific modification of proteins that function under mild conditions is an area of intense research.[18] While a number of reactions suitable for protein modification have been developed,[19−22] to date, the Cu(I)-catalyzed click reaction has been the most widely employed bioorthogonal process.[20] Although highly useful, that reaction employs Cu(I) which is toxic to cells and can in some cases erode biological activity. To address those issues, copper-free variations of the click reaction have been developed that function based on inclusion of electron-withdrawing substituents and/or ring strain into alkyne-containing reagents. While highly promising, these new reagents are not generally commercially available, are difficult to synthesize, and manifest low aqueous solubility.[23−25]

As an alternative, oxime- and hydrazone-based reactions have found wide application in the conjugation of biomolecules on account of the absence of aldehyde or ketone groups in proteins and their orthogonal reactivity with aminooxy or hydrazine derivatives to give stable hydrazones or oximes.[26−33] While reactions between aldehydes and ketones with alkoxy-

Received: December 2, 2011
Published: March 21, 2012

   dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455−8467

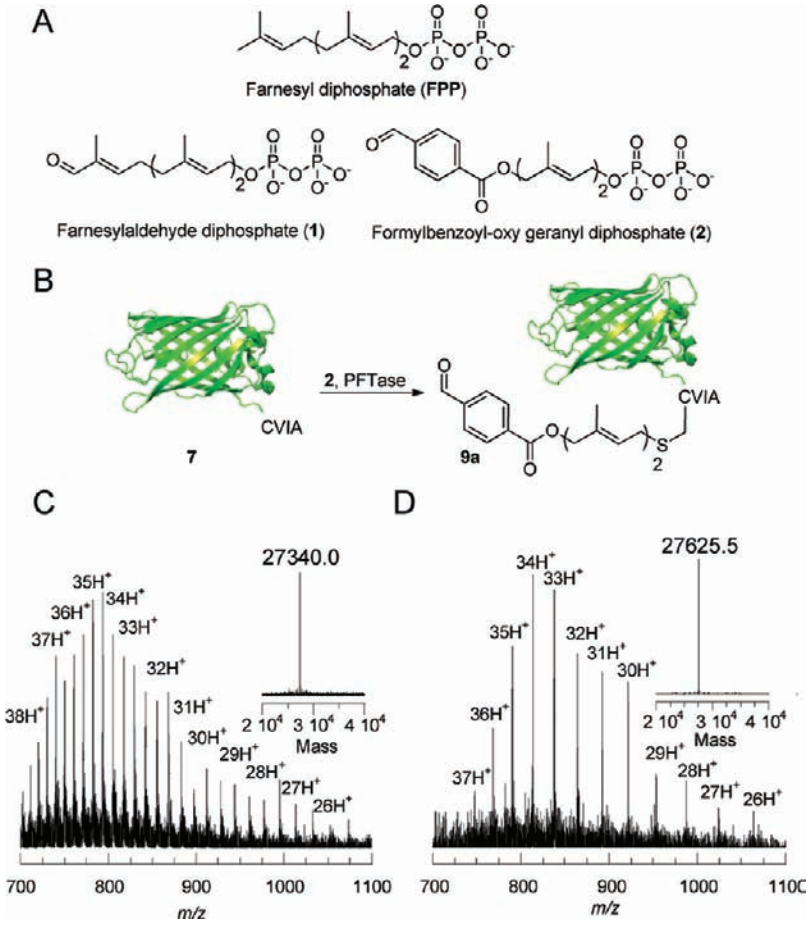


Figure 1. (A) Structures of farnesyl diphosphate, farnesyl aldehyde diphosphate (**1**), and formylbenzoyl-oxy geranyl diphosphate (**2**). (B) Schematic representation of prenylation of a protein containing a CAAX-box positioned at its C-terminus (GFP−CVIA, **7**) with aldehyde-containing analogue **2** to yield the prenylated product **9a**. (C) ESI MS analysis of **7** with the deconvoluted mass spectrum shown in the inset. (D) ESI MS analysis of **9a** with the deconvoluted mass spectrum shown in the inset.

amines or hydrazides are generally slow, they can be significantly accelerated by addition of aniline.[27,34] This has resulted in a number of exciting applications ranging from the site-specific glycosylation of proteins[29] to the fluorescent labeling of bacteria and mammalian cells.[35] Given the utility of oxime and hydrazone formation, a number of methods have been developed to introduce aldehydes and ketones into proteins. Chemical approaches include transamination in the presence of sodium glyoxylate and copper sulfate[36] or using pyridoxal-5-phosphate,[37] while these have proved to be powerful methods, they are not applicable to all N-termini and cannot always be driven to completion. Enzymatic methods include the action of formylglycine-generating enzyme[38] or nonsense suppression approaches that permit incorporation of aldehydes into internal positions within proteins.[39]

Recently, our laboratory and others have exploited the high specificity of the enzyme protein farnesyltransferase[40] (PFTase) to site-specifically modify peptides and proteins.[32,41−43] In nature, PFTase catalyzes the transfer of a farnesyl isoprenoid group from farnesyl diphosphate (FPP, Figure 1) to a sulfur atom present in a cysteine residue. That residue must be located in a tetrapeptide sequence (denoted as a CAAX-box) positioned at the C-terminus of a protein or peptide to be a PFTase substrate. Interestingly, CAAX-box sequences such as CVIA can be appended to the C-termini of many proteins, rendering them efficient substrates for PFTase. Since PFTase can tolerate many simple modifications to the isoprenoid substrate,[3,44−48] it can be used to introduce a diverse range of functionality into proteins at their C-termini. Chemoselective reaction with the resulting functionalized protein can then be used for a wide range of applications. Since the "AAX" residues from a CAAX-box sequence can be removed by treatment with carboxypetidase after prenylation, net addition to the protein in this labeling method can be limited to a single prenylcysteine residue.[49]

In an initial communication,[32] we reported that compound **1**, an aldehyde-containing analogue of FPP, can be incorporated into a purified protein substrate using PFTase and that the resulting aldehyde-functionalized protein can be immobilized or fluorescently labeled via oxime formation. In this study, we followed up on those initial observations by comparing the properties of α,β-unsaturated aldehyde **1** with aryl aldehyde **2** in terms of their efficiency as PFTase substrates and reactants in both oxime and hydrazone formation. Using green fluorescent protein (GFP[50]) as a model system, we demonstrate that the purified protein can be enzymatically modified with **1** or **2**. Oxime or hydrazone formation was then employed to immobilize, fluorescently label, or PEGylate the resulting aldehyde-functionalized proteins. Immobilization via hydrazone formation was also shown to be reversible via transoximization with a fluorescent alkoxyamine. After characterizing this labeling strategy using pure protein, the specificity of the

dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455−8467

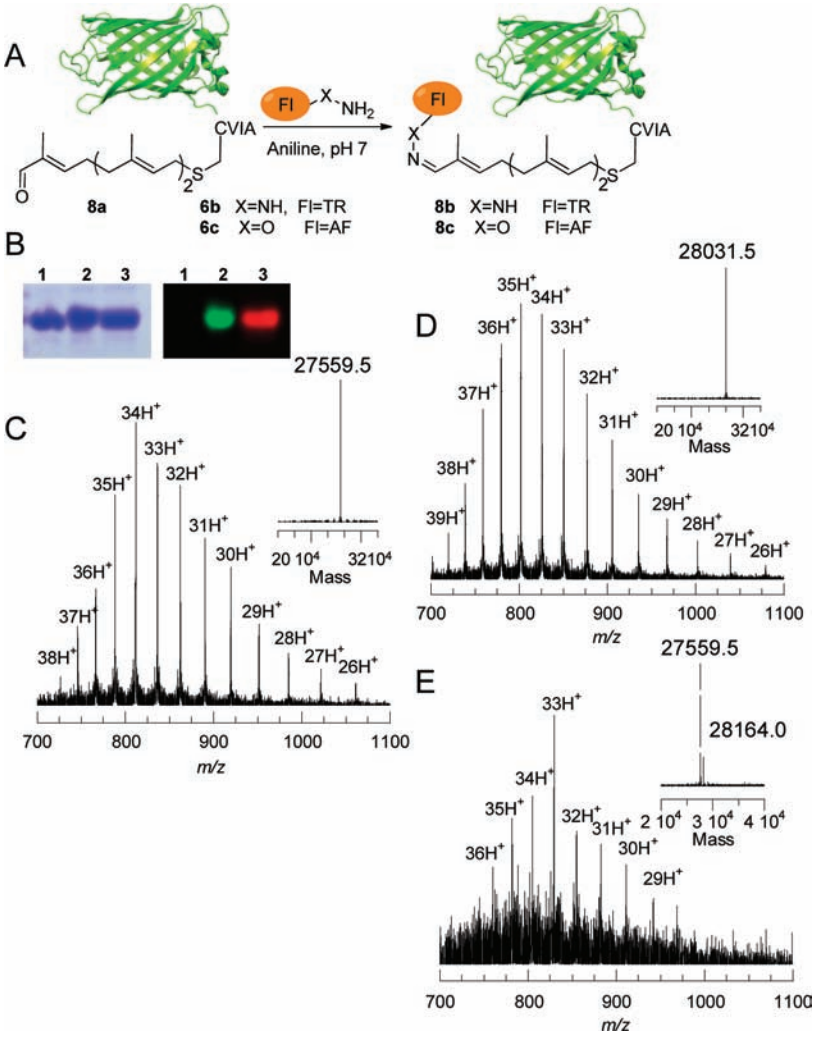


Figure 2. (A) Schematic representation of oxime and hydrazone ligations of **8a** to yield **8b** and **8c**. (B) Fluorescence (right) and Coomassie blue staining (left) images of a gel loaded with **8a** labeled with alexafluor **6c** and Texas red **6b** via oxime and hydrazone ligations, respectively, showing covalent attachment of fluorophores to the protein: lane 1, GFP−CVIA **7**; lane 2, **8c**; lane 3, **8b**. ESI MS spectra of **8a** (C) and hydrazone/oxime ligation products **8b** and **8c**, showing full conversion for oxime (D) and ~20% for hydrazone (E) ligations with the deconvoluted mass spectra shown in the insets.

enzymatic process was used to selectively label GFP present in crude *E. coli* extract followed by capture of the aldehyde-modified protein using hydrazide−agarose. Subsequent incubation of the immobilized protein using a fluorescently labeled or PEGylated alkoxyamine resulted in the release of pure GFP containing the desired site-specific covalent modifications. This procedure was also employed to produce PEGylated glucose-dependent insulinotropic polypeptide[51] (GIP), a protein with potential therapeutic activity for diabetes.[52]

EXPERIMENTAL SECTION

Enzymatic Studies of FPP Analogues 1 and 2 Using a Continuous Fluorescence Assay. Enzymatic reaction mixtures contained Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (20 mM), $ZnCl_2$ (10 μM), 2.4 μM *N*-dansyl-GCVIA (**3**), 0.04% (w/v) *n*-dodecyl-ß-D-maltoside, 80 nM PFTase, and varying concentrations of either **1** or **2** (0−50 μM) in a final volume of 250 μL. Reaction mixtures were equilibrated at 30 °C for 5 min, initiated by addition of PFTase, and monitored for an increase in fluorescence (λ_{ex} = 340 nm, λ_{em} = 505 nm) for approximately 10 min. The initial rates of formation of products were obtained as slopes in IU/min using least-squares analysis. Corrections were applied to all of the rate calculations based on the difference between the fluorescence intensity of the prenylated product and the starting peptide. Assuming 100% conversion, the difference corresponds only to the fluorescence of the total amount of product. The slope was then divided by the fluorescence difference followed by multiplying by the total concentration of peptide (2.4 μM), which then gives the rate of formation of product in μM/s. It should be noted that the K_M values reported here are actually apparent K_M values, since the measurements were performed at only a single peptide concentration. The data were fit to a Michaelis−Menten model using a nonlinear regression program to determine k_{cat} and K_M.

Enzymatic Synthesis of 4a and 5a. Enzymatic reactions (26 mL) contained Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (20 mM), $ZnCl_2$ (10 μM), DTT (5.0 mM), **3** (2.4 μM), PFTase (80 nM), and either **1** or **2** (30−50 μM). To ensure complete disulfide reduction of the peptide, all reagents except substrates and enzyme were premixed and incubated for 2 h at 4 °C. With all reagents mixed, the reaction was initiated by addition of enzyme and the resulting mixture was incubated at 30 °C for 1 h. Reaction progress was monitored by UV absorbance (λ = 340 nm, absorbance of the dansyl chromophore) using analytical RP-HPLC. The following conditions were employed: flow rate, 1 mL·min^{-1}; 500 μL injection loop; gradient 0−100% B in 30 min; solvent A, NH_4HCO_3 (25 mM in H_2O); solvent B, CH_3CN. After 1 h, the reaction was purified using a Waters Sep-Pak

dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455−8467

Plus reversed-phase C_{18} Environmental Cartridge. The cartridge was first washed with solvent B (10 mL) and then equilibrated with solvent A (20 mL). The crude enzymatic reaction mixture was applied to the cartridge, and a gradient elution was performed in the following sequence: 10 mL of solvent A, 10 mL of solvent C (20% solvent B, 80% solvent A), 10 mL of solvent D (40% solvent B, 60% solvent A), and 10 mL of solvent E (60% solvent B, 40% solvent A). Fractions (1 mL per tube) were collected, and product elution was monitored using a hand-held UV lamp. The green-fluorescent product was clearly visible, the brightest fraction was selected, and its purity was confirmed by RP-HPLC. LC-MS analysis of the purified products gave ions of 913.5 and 979.4 as the predominant species, which are consistent with $[M + H]^+$ for **4a** and **5a**, respectively.

Oxime Ligation between Peptide−Aldehyde 4a and 5a and Aminooxy Alexafluor-488 (6c). Coupling reactions contained 3−5 μM **4a** or **5a**, 200 μM alexafluor-488 (**6c**), PB (0.1 M, pH 7.0), and aniline (100 mM) in a final volume of 500 μL. Reactions were performed at room temperature and initiated by addition of aniline (100 mM). LC-MS analysis of the reaction mixture after 3−4 h gave ions of 1384.6 and 1450.6 as the predominant species, which are consistent with $[M + H]^+$ for **4c** and **5c**, respectively.

Hydrazone Ligation between Peptide−Aldehydes 4a and 5a and Texas Red Hydrazide (6b). Coupling reactions contained 3−5 μM **4a** or **5a**, 200 μM Texas red (**6b**), PB (0.1 M, pH 7.0), and aniline (100 mM) in a final volume of 500 μL. Reactions were performed at room temperature and initiated by addition of aniline (100 mM). LC-MS analysis of the reaction mixture after 1 h gave ions of 758.46 and 791.45 as the predominant species, consistent with $[M + 2H]^{2+}$ for hydrazones **4b** and **5b**, respectively.

Enzymatic Incorporation of Compounds 1 and 2 into GFP−CVIA (7).[53] Enzymatic reaction mixtures (10 mL) contained Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (30 mM), $ZnCl_2$ (10 μM), DTT (5.0 mM), **7** (2.4 μM), either **1** or **2** (30−50 μM), and PFTase (80−200 nM). After incubation at 30 °C for 2 h for **1** and overnight for **2**, the respective reaction mixtures were concentrated using an Amicon Centriprep centrifugation device (10 000 MW cutoff). Next, the excess of **1** or **2** was removed through a NAP-5 (Amersham) column using Tris·HCl (50 mM, pH 7.5) as the eluant. The subsequent protein concentration was calculated by UV absorbance at 488 nm (ε = 55 000 $M^{-1}\cdot cm^{-1}$).

Coupling Reaction between Aldehyde-Labeled GFP−CVIA (8a and 9a) with Alexafluor-488 (6c). Alexafluor-488 (**6c**) (3.2 μL of 3.2 mM solution in DMSO) was added to 42 μL of **8a** or **9a** (stock solution of 60 μM in PB). PB (2 M, pH 7.0, 2.5 μL) was added, and the reaction was initiated by adding aniline (100 mM) and allowed to proceed for 3−5 h at room temperature. The mixture was then purified by a NAP-5 column to remove excess dye. LC-MS analysis of the sample showed only oxime-ligated protein, and no free aldehyde was detected, indicating a complete reaction in both cases.

Immobilization of 9a onto Hydrazide Agarose Beads. Hydrazide agarose beads (Thermo Scientific, hydrazide loading 16 μmol/mL) (300 μL) were washed with PB (0.1 M, pH 7.0, 3 × 500 μL). PB (30 μL, 1 M, pH 7.0) was added to the beads followed by addition of **9a** (200 μL, 87 μM). Immobilization was initiated by adding aniline (2 μL, 100 mM). For controls, GFP−CVIA (**7**) was added instead of **9a**. The solution was centrifuged; then the GFP UV-absorbance of the supernatant was measured as a function of time (488 nm, ε = 55 000 $M^{-1}\cdot cm^{-1}$). After 2 h, the solution was centrifuged and the beads were washed thoroughly with PB (0.3 M, pH 7.3, 3 × 300 μL) and KCl (1 M, 3 × 300 μL) to remove nonspecifically bound proteins and stored in pH 7.5 Tris buffer at 4 °C.

Release of Immobilized GFP from Beads Using Hydroxylamine. The GFP beads were incubated in PB (0.3 M, pH 7.0) with hydroxylamine (200 mM) and aniline (100 mM), and the resulting mixture was vortexed. The solution was centrifuged, and then the UV absorbance of the GFP in the supernatant was measured as a function of time (488 nm, ε = 55 000 $M^{-1}\cdot cm^{-1}$).

Coupling Reaction between Aldehyde-Labeled GFP−CVIA (8a and 9a) with Alexafluor-488 (6c). Alexafluor-488 (**6c**) (3.2 μL of 3.2 mM solution in DMSO) was added to 42 μL of **8a** or **9a** (stock solution of 60 μM in PB). PB (2 M, pH 6.7, 2.5 μL) was added, and the reaction was initiated by adding 100 mM aniline and allowed to proceed for 5−6 h at room temperature. The mixture was then purified using a NAP-5 column to remove excess dye. LS-MS analysis of the sample showed only oxime-ligated protein, and no free aldehyde was detected, indicating a complete reaction in both cases.

Coupling Reaction between Aldehyde-Labeled GFP−CVIA (8a and 9a) with Texas Red Hydrazide (6b). Texas red hydrazide (**6b**) (7 μL of 1.9 mM solution in DMSO) was added to 100 μL of **8a** and **9a** (stock solution of 40 μM in PB). Reaction was initiated by adding 100 mM aniline and allowed to proceed for 1 h at room temperature. The mixture was then purified using a NAP-5 column to remove excess **6b**. LC-MS analysis of the sample showed the presence of both hydrazone-ligated proteins **8b** and **9b** and free aldehydes **8a** and **9a**. The ratios of free aldehydes to their respective hydrazone products were ∼4, indicating only ∼20% completion within this range of reactant concentrations. The mass spectrum in Figure 2E, which corresponds to conversion of compound **8a** to **8b**, was made by superimposing the two separate mass spectra of both the free aldehyde protein **8a** and the hydrazone-ligated protein **8b** present in the product mixture, based on their relative intensities. The aldehyde- and hydrazone-functionalized proteins have different retention times and thus show two different peaks in the corresponding LC chromatograms.

FRET Studies Between GFP−Aldehyde 9a and Texas Red Hydrazide (6b). Texas Red hydrazide (**6b**) (7 μL of 1.9 mM solution in DMSO) was added to 90 μL of **9a** (stock solution of 60 μM in PB, 0.1 M, pH 6.7). Reaction was initiated by adding 0.9 μL of aniline (100 mM) followed by vortexing the solution and allowing it to proceed for 1 h at room temperature. The mixture was then purified using a NAP-5 column to remove excess **6b**. The protein solution was diluted to 1 nM, and its fluorescence was measured. For the first control, the same amount of GFP−CVIA (**7**) fluorescence was measured and compared with that of **9b**. As a second control, the ligated protein (**9b**) was heated for a few minutes to denature the protein and the resulting fluorescence was measured to verify that FRET required both the protein and the Texas red fluorophores.

FRET Studies between GFP−Aldehyde 9a and Aminooxy−TAMRA (6d). Aminooxy−TAMRA (**6d**) (100 μM) was added to 90 μL of **9a** (stock solution of 50 μM in Tris·HCl (100 mM, pH 7.0). Reaction was initiated by adding 0.9 μL of aniline (100 mM) followed by vortexing the solution and allowing it to proceed for 3 h at room temperature. The mixture was then purified using a NAP-10 column to remove excess **6d**. The collected solution appeared red and not green, suggesting that efficient FRET was occurring between the GFP and the TAMRA fluorophores. The fluorescence of the solution was measured and compared with two other controls. For the first control, the same amount of GFP−CVIA (**7**) fluorescence was measured and compared with that of **9e**. As a second control, the same amount of TAMRA fluorescence was measured. The two controls confirmed that FRET was occurring between the protein and the TAMRA fluorophores.

Crude Prenylation, Immobilization, and Subsequent Labeling and Release of GFP−CVIA. A pellet of cells expressing GFP−CVIA was suspended in buffer (20 mM Tris·HCl pH 7.5, 1 mM EDTA), sonicated, and clarified by centrifugation. The GFP concentration present in the crude soluble protein mixture was calculated by UV absorbance at 488 nm. Next, prenylation was performed by adding PFTase (200 nM), **2** (50 μM), Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (30 mM), $ZnCl_2$ (10 μM), and DTT (5.0 mM) to a solution of **7** to achieve a final concentration of 2.0 μM in the crude mixture. After overnight incubation at 30 °C, the reaction mixture was filtered and concentrated using an Amicon Centriprep centrifugation device (10 000 MW cutoff). Next, excess **2** was removed through a NAP-5 (Amersham) column using Tris·HCl (50 mM, pH 7.5) as the eluting solvent. The subsequent GFP concentration in the crude mixture was calculated by UV absorbance at 488 nm and determined to be 30 μM. Immobilization was performed as described above. The beads were washed thoroughly

with PB (0.3 M, pH 7.3) and KCl (3 × 300 μL, 1 M) to remove nonspecifically bound proteins followed by incubation with aminooxy fluorophore **6c** (1 mM) and aniline (100 mM) overnight with constant agitation. The supernatant was then analyzed via SDS-PAGE and in-gel fluorescence analysis to confirm the labeling and release of the protein from the beads.

Coupling Reaction between Aldehyde-Labeled GFP−CVIA (9a) with Aminooxy PEG (10). Aminooxy PEG (**10**) (1.5 mg, MW 10 kDa) was added to 100 μL of **9a** (stock solution of 10 μM in 50 mM Tris·HCl). PB (pH 7) was added to a final concentration of 0.1 M. Reaction was initiated by adding 100 mM aniline and allowed to proceed for 1 h at room temperature. SDS-PAGE analysis of the sample was used to confirm covalent attachment of the PEG **10** to aldehyde **9a**. Excess **10** and PB were removed using a zip-tip protocol followed by MALDI MS analysis of the sample to characterize the product and demonstrate that no free aldehyde was present, indicating complete reaction.

PEGylation from Immobilized GFP Beads. Immobilization was performed as described above. Beads were washed thoroughly with PB (0.3 M, pH 7.3, 3 × 300 μL) and KCl (1 M, 3 × 300 μL) to remove nonspecifically bound proteins. Next, the beads were incubated with aminooxy PEG **10** (2 mM) and aniline (100 mM) overnight while vortexing the solution. SDS-PAGE analysis of the supernatant indicated the successful PEGylation and release of the aldehyde−GFP from the hydrazide beads.

Enzymatic Prenylation of GIP-CVIM (12a) with Aldehyde Substrate 2. Enzymatic reaction mixtures (10 mL) contained Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (30 mM), $ZnCl_2$ (10 μM), DTT (5.0 mM), **12a** (2 μM), **2** (50 μM), and PFTase (200 nM). After incubation at 30 °C overnight, significant precipitate was present in solution, suggesting possible precipitation of GIP was occurring upon prenylation. The precipitate was separated from the solution by centrifugation at 12 000*g* for 15 min, washed with 25 mM $(NH_4)_2CO_3$ buffer to remove excess **2**, and centrifuged again. MALDI-MS analysis of the precipitate (dissolved in H_2O, 0.5% TFA, v/v) confirmed the prenylation of GIP with **2**, while the solution showed neither starting GIP **12a** nor the prenylated material **12b**.

Coupling Reaction between Aldehyde-Labeled GIP (12b) with Aminooxy PEG (13). A small amount of the precipitate **12b** was dissolved in H_2O containing 0.5% TFA (v/v), and aminooxy PEG (**13**) was added to a final concentration of 200 μM. Reaction was allowed to proceed for 2 h at room temperature. Excess **13** was removed using a zip-tip protocol. MALDI-MS analysis of the sample was employed to characterize the product and demonstrate that no free aldehyde was present, indicating complete reaction.

Prenylation of GIP (12a) in Crude *E. coli* Extract. *E. coli* extract containing GIP-CVIM (**12a**) was subjected to enzymatic prenylation by incubating it in the presence of PFTase (200 nM), **2** (50 μM), Tris·HCl (50 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (30 mM), $ZnCl_2$ (10 μM), and DTT (5.0 mM). After overnight incubation, the precipitate was separated from the solution by centrifugation at 12 000*g* for 15 min, washed with 25 mM $(NH_4)_2CO_3$ buffer to remove excess **2**, and then centrifuged again. MALDI-MS analysis of the precipitate, performed as described above, was used to confirm the prenylation of GIP with **2**.

Immobilization and Subsequent PEGylation and Release of Resin-Bound GIP. Immobilization was performed as described above for GFP except that GIP was dissolved in H_2O containing 0.5% TFA (v/v). No aniline catalyst was added to the solution in this case. After incubation for 1 h, the beads were washed thoroughly with H_2O (3 × 300 μL) to remove nonspecifically bound proteins from the beads. Next, the beads were incubated with aminooxy PEG **13** (~400 μM) in H_2O containing 0.5% TFA overnight with constant agitation of the solution. Excess **13** was removed using a zip-tip protocol. MALDI MS analysis of the sample was employed to confirm the presence of the desired product and to assess the purity of the PEGylated GIP (**14**).

General Procedure for MALDI Analysis of Protein Samples. A zip-tip (C_4 column) was first washed with 10 μL of solvent A (CH_3CN containing 0.1% TFA; v/v) and then equilibrated with solvent B (H_2O containing 0.1% TFA; v/v). Sample (10 μL) was then adsorbed onto the C_4 matrix via repeated cycles of aspiration and ejection (5−10 cycles) using a pipettor. Next, the zip-tip was washed 5 × 10 μL with solvent B and the proteins eluted with 2 μL of a mixture of solvent A and B (75:25). Next, 0.7 μL of the eluted material was added to a MALDI plate, and 0.7 μL of matrix was added on top of the sample plate to form crystals. A saturated solution of sinapinic acid (3,5-dimethoxy-4-hydroxy-cinnamic acid) was used as the matrix.

RESULTS AND DISCUSSION

Comparison of Alkyl and Aryl Aldehydes As Substrates for PFTase. To examine the ability of PFTase to be used in a protein modification strategy employing oxime and hydrazone formation, we first wanted to explore the range of aldehydes that could be accepted as alternative substrates for PFTase. Thus, compound **2**, containing an aryl aldehyde, was designed and synthesized in five steps from geraniol (Scheme S2, Supporting Information). In brief, THP-protected geraniol was initially oxidized at C-8 to a terminal alcohol,[54] followed by acylation with formylbenzoic acid using EDC as the coupling reagent. The THP group was removed, and the alcohol was converted to the corresponding allylic bromide using CBr_4 and PPh_3. Subsequent displacement with $[(n\text{-}Bu)_4N]_3HP_2O_7$ followed by purification via ion-exchange chromatography and RP-HPLC yielded product in which the desired aldehyde was almost completely transformed to the corresponding carboxylic acid; therefore, a direct phosphorylation strategy using $(HNEt_3)_2HPO_4$ and CCl_3CN as the activating reagent was employed. Subsequent purification by RP-HPLC produced the desired aldehyde analogue **2** in 5.4% overall yield, whose structure was confirmed by ^1H NMR, ^{31}P NMR,[55] and HR-ESI-MS. Aldehyde **1** was prepared in six steps starting from farnesol as previously described[32] with several modifications that significantly improved the overall yield to 1.3%. Despite these improvements, synthesis of **2** was significantly more efficient primarily due to the selectivity in the SeO_2 oxidation step. Preparation of **1** proceeds via THP-protected farnesol, whereas synthesis of **2** uses THP−geraniol. Selective oxidation of the alkene at C-6 (over the electron-poor C-2 alkene) in geraniol is facile compared to preferential oxidation of the alkene at C-10 in farnesol due to competing reaction with the C-6 olefin, which exhibits comparable reactivity. Hence, the reaction cannot be driven to completion, resulting a significantly reduced yield (compare 56% for geraniol oxidation to 23% for farnesol oxidation).

Initially, prenylation reactions containing *N*-dansyl-GCVIA (**3**), **2**, and PFTase were monitored by HPLC and LC-MS/MS. As observed previously with **1**, a new species with longer retention time appeared in the reaction mixture containing **2**. LC-MS analysis of that compound gave an $[M + H]^+$ peak at 979.4 Da, consistent with the proposed structure of peptide **5a** (Supporting Information). Next, kinetic analysis of the incorporation of analogue **2** by PFTase was performed using a continuous fluorescence-based enzyme assay as previously carried out with **1**. Varying concentrations of **2** were incubated with the fluorescent peptide substrate, *N*-dansyl-GCVIA, and PFTase; the rates of those enzymatic reactions were determined and shown to obey saturation kinetics. Steady-state kinetic parameters for prenylation reactions with the two aldehyde analogues are summarized in Table 1 with additional details provided in the Supporting Information section (Figure S2). Comparison of the catalytic efficiencies for these alternative substrates indicates that both compounds have reduced efficiency relative to FPP, manifesting k_{cat}/K_M values of

Table 1. Steady-State Kinetic Parameters of Substrates, and HPLC Retention Times for Prenylated Peptide Products

compound[a]	k_{cat} (s^{-1})	K_M (μM)	$(k_{cat}/K_M)_{rel}$[a]	R_t (min)
FPP	0.52	1.71	1	
1	0.133 ± 0.003	1.87 ± 0.17	0.23	
2	0.015 ± 0.001	1.02 ± 0.16	0.05	
N-dansyl-GC(Far) VIA				21.5
4a				18.9
5a				18.6

[a] V_{rel} refers to k_{cat}/K_M with respect to FPP

0.23 and 0.05, respectively (relative to FPP). We found that decreases in k_{cat} constituted the major reason for the diminished catalytic efficiency of the analogues; k_{cat} for aldehyde **1** was 4-fold lower, while k_{cat} for aldehyde **2** was 35-fold lower (compared to k_{cat} for FPP). No significant differences were observed in the K_M values for the different analogues. Thus, in summary, while **1** is the superior alternative substrate, **2** is easier to prepare making these two compounds functionally interchangeable.

Preparation and Reactivity of PFTase-Mediated Aldehyde-Functionalized Peptides. In our earlier work with **1**, experiments with the aldehyde-functionalized peptide **4a** focused on oxime-forming reactions. Here, we sought to expand the scope of possible chemistry to include hydrazone formation as well. Accordingly, large-scale (26 mL) reactions containing *N*-dansyl-GCVIA (**3**), **1** or **2**, and PFTase were performed and the products isolated after purification via solid-phase extraction. The resulting material was subsequently used to evaluate ligation reactions between Texas red hydrazide **6b**/alexafluor-488 aminooxy **6c** and aldehyde-containing peptides **4a** and **5a** in the presence of aniline.[27] Kinetic analysis of oxime formation showed that in the range of 2−4 μM **4a** and **5a** ligations at pH 7 were essentially complete within 3−4 h. LC-MS analysis of the reaction mixture resulted in $[M + H]^+$ peaks being observed at 1384.6 and 1450.6 Da, consistent with production of oximes **4c** and **5c**, respectively (Supporting Information). In contrast, hydrazone formation with **4a** and **5a** in the same concentration range of reagents showed only 30−50% completion but within 30−60 min (a significantly shorter time frame). LC-MS analysis of the reaction mixtures resulted in $[M + 2H]^{2+}$ peaks observed at 758.46 and 791.45 Da, consistent with formation of hydrazones **4b** and **5b**, respectively (Supporting Information). Overall, these experiments with aldehyde-containing peptides **4a** and **5a** suggest that hydrazone ligations have the advantage over oxime-forming reactions of reaching equilibrium at a higher rate but at the cost of lower conversion to the conjugated products due to their lower association constants.

Preparation and Reactivity of PFTase-Mediated Aldehyde-Functionalized Proteins. With the ability of aldehyde analogues **1** and **2** to be incorporated by PFTase and their subsequent derivatization via oxime and hydrazone ligations established in a peptide model system, we next evaluated the utility of the aldehyde analogues for selective protein modification. Accordingly, aldehydes **1** and **2** were incubated with GFP−CVIA (**7**) in the presence of PFTase for 2 h and overnight at 30 °C, respectively. Those reaction times were based on our earlier observations that peptide substrate **3** could be prenylated with aldehyde analogues **1** and **2** in less than 1 and 4 h, respectively. Concentration by ultra-centrifugation followed by size-exclusion chromatography to remove unreacted substrates yielded aldehyde-functionalized GFP−CVIA **8a** and **9a**. Reaction completion was confirmed by LC-MS analysis (Figure 1C and 1D) in which none (in the case of **8a**) or very small amounts (in the case of **9a**) of free GFP−CVIA (**7**) could be detected in comparison to the large peaks for prenylated GFPs. Deconvolution of the LC-MS data from the purified protein products showed species at 27 559.0 and 27 625.5 Da, consistent with the structures of aldehyde−GFPs **8a** and **9a**. In general, LC-MS analysis of GFP and its congeners has proved to be quite powerful for studying these reactions. As noted above, in a preliminary communication[32] we had shown that aldehyde−GFPs **8a** could be derivatized to produce oxime-linked products. Here, it was desired to expand those experiments to include hydrazone formation and compare the relative reactivity between the two different aldehyde donors, **8a** and **9a**, containing α,β-unsaturated and aryl−aldehydes, respectively. To fluorescently label those aldehyde-functionalized proteins, we chose Texas red hydrazide (**6b**) and Alexaflour-488 aminooxy (**6c**) for their excellent quantum yields and high visible light absorption. Thus, aldehyde−GFPs **8a** and **9a** were incubated separately with alkoxyamine **6c** at pH 7 and room temperature. Kinetic analysis, performed via LC-MS measurements, showed that the reaction required 3−4 h to proceed to completion. At that point, no detectable unmodified protein−aldehydes (**8a** and **9a**) were observed. Gratifyingly, the deconvoluted MS data indicated the presence of species at 28 032.0 and 28 120.5 Da, consistent with the proposed oximes **8c** and **9c**. In-gel fluorescence analysis performed under denaturing conditions confirmed covalent attachment of aminooxy **6c** to the aldehyde-containing proteins (Figure 2B). Unprenylated GFP−CVIA (**7**) failed to show any labeling with alkoxyamine **6c**, further confirming that the ligations require the presence of the enzymatically introduced aldehyde functionality and that the ligation reaction is truly bioorthogonal. Overall, the oxime ligation reactions appear to be highly efficient since no unligated aldehyde−GFPs (**8a** or **9a**) were observed upon LC-MS analysis (Figure 2C and 2D) of the ligation reaction mixtures.

Aldehyde-functionalized GFPs **8a** and **9a** were also each incubated with hydrazide **6b** at pH 7 and room temperature under the same conditions employed in the aforementioned oxime ligations. LC-MS analysis of aldehyde-functionalized GFP-containing reactions after 1 h showed approximately 20% conversion of aldehydes **8a** and **9a** to their respective hydrazones **8b** and **9b** (Figure 2C and 2E); more extensive reaction times did not result in the appearance of additional hydrazone product, suggesting that the reaction had reached equilibrium within 1 h. These results are in good agreement with those from hydrazone ligations for aldehyde-functionalized peptides **4a** and **5a** described above.

Application to FRET Analysis of Labeled GFP. To demonstrate the utility of this method for applications beyond simple protein labeling, we next investigated the ability of Texas red-labeled GFP (**9b**) to undergo fluorescence resonance energy transfer (FRET). After performing the ligation reaction between aldehyde **9a** and Texas red hydrazide **6b** at room temperature for 1 h, excess fluorophore was removed via size-exclusion chromatography. A strong fluorescent signal at 640 nm (emission wavelength of Texas red) was observed upon excitation at 488 nm (excitation wavelength of GFP), indicative of FRET between Texas red and GFP due to their close

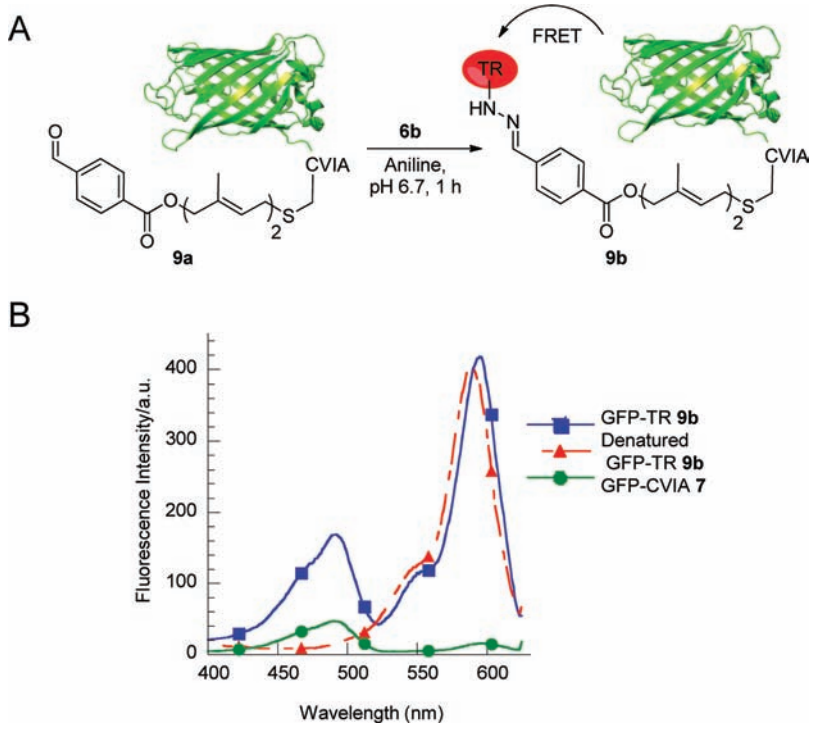


Figure 3. (A) Schematic representation of the fluorescent labeling of **9a** via hydrazone ligation. Conjugated protein was expected to show FRET between Texas red and GFP−CVIA. (B) Excitation spectra obtained by monitoring at 640 nm: (■) **9b**, (▲) denatured **9b**, and (●) **7**. All three samples had equal concentrations of the chromophores.

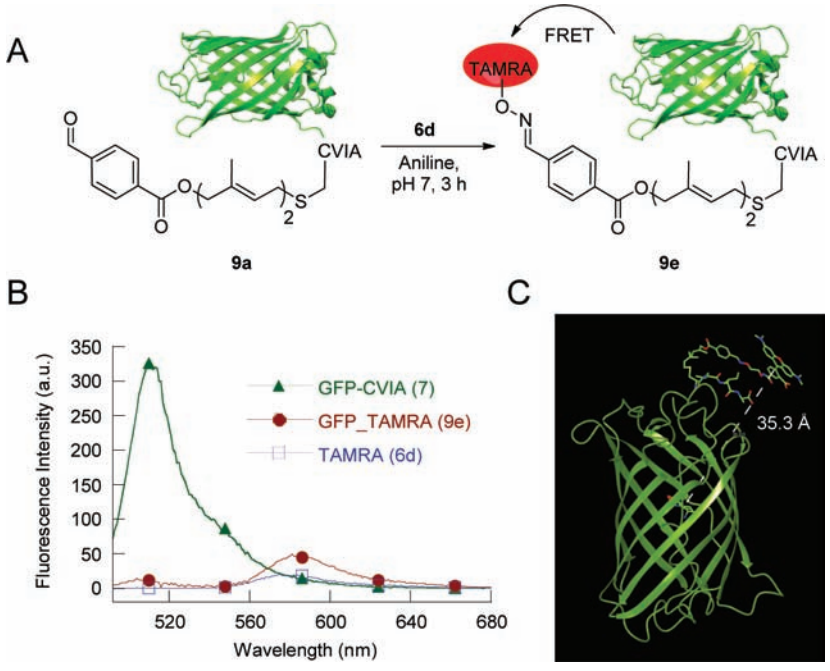


Figure 4. (A) Schematic representation of the fluorescence labeling of **9a** via oxime ligation. Conjugated protein was expected to show FRET between TAMRA and GFP. (B) Emission spectra obtained by excitation at 488 nm: (●) **9e**, (▲) GFP−CVIA (7), and (■) **6d**. All three samples had equal concentrations of the chromophores. (C) Molecular model of GFP−TAMRA **9e** conjugate.

proximity, resulting from covalent attachment of the fluorophore to the protein (Figure 3). When the hydrazone-ligated protein was denatured, no FRET was observed upon excitation at 488 nm (Figure 3, red spectrum) and only a small background peak was observed upon excitation of GFP that had not been modified with Texas red (Figure 3, green spectrum), further confirming that FRET was occurring between the aldehyde-functionalized protein and the fluorophore. While the above results appeared promising, the FRET efficiency could not be calculated from the experimental data due to incomplete hydrazone ligation reaction. Hence, aminooxy−TAMRA **6d** was ligated with aldehyde−protein **9a**. As expected, LC-MS analysis of the oxime formation reaction mixture showed a peak at 28 111 Da consistent with the structure of TAMRA-labeled

dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455−8467

GFP **9d** and showed no **9a**, in good agreement with the high efficiency observed in previous oxime ligation reactions. Emission spectra of **9e**, monitored at 488 nm excitation, showed FRET, while the same amount of GFP−CVIA (**7**) and fluorophore **6d** showed a substantially larger emission band at 510 nm and a smaller band at 580 nm, respectively. Energy was transferred from donor (GFP) to acceptor (TAMRA) with an efficiency greater than >96%, and the distance was calculated to be 37 Å (Figure 6B), consistent with a distance of 35 Å calculated for the model GFP−TAMRA (Figure 6C) and measured from the GFP chromophore to the TAMRA fluorophore (see the Supporting Information for a description of the modeling).

Reversible Immobilization of Purified Aldehyde-Functionalized GFP Using Hydrazide-Modified Agarose Beads. Next, we examined two additional applications for the aldehyde-functionalized proteins described here. First, their utility in protein immobilization was examined (Figure 5).

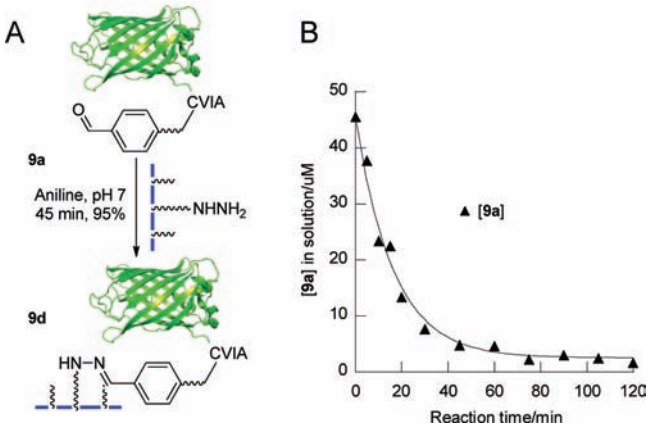


Figure 5. (A) Schematic representation of immobilization of **9a** onto hydrazide-functionalized agarose beads to yield **9d**. (B) Kinetic analysis of immobilization of **9a** onto hydrazide-functionalized agarose beads. Reaction was carried out at room temperature in the presence of 100 mM aniline and excess beads. UV absorbance of GFP in the supernatant was measured at different times showing >95% immobilization in ~45 min. Data was fit to a simple exponential process.

Hydrazide-functionalized agarose beads were incubated with aldehyde−GFP **9a** at room temperature in the presence of 100 mM aniline. The immobilization reaction was followed by monitoring the UV absorbance at 488 nm of the supernatant as a function of time. Results from those measurements showed that equilibrium was reached in approximately 45 min; in that time, the beads became highly fluorescent (Figure 6A); less fluorescent beads were observed in the absence of aniline catalyst, and no fluorescent beads were seen using GFP lacking the aldehyde moiety (Figure 6C). On the basis of the amount of aldehyde−GFP **9a** remaining in the supernatant, the efficiency of covalent immobilization was calculated to be greater than 95% (Figure 5B), an impressive result for site-specific protein immobilization. Next, oxime ligation using hydroxylamine in the presence of aniline was employed to remove the covalently immobilized hydrazone−GFP **9d**. Hydroxylamine (200 mM) was incubated with **9d** in the presence of aniline (100 mM) at room temperature, and the UV absorbance at 488 nm of the supernatant was measured as a function of time. In this case, analysis of the results showed that

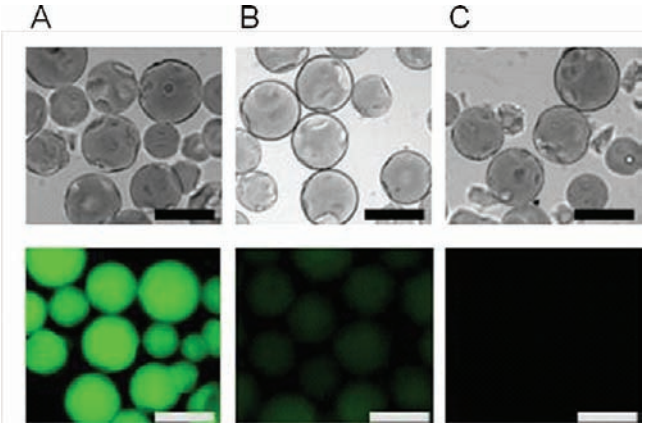


Figure 6. Immobilization onto and subsequent release of **9a** from hydrazide-functionalized agarose beads: (A) immobilization reaction mixture in the presence of aniline, (B) release of **9d** from agarose beads via oxime ligation with hydroxylamine in the presence of aniline for ~3 h, and (C) control immobilization reaction containing unmodified GFP−CVIA **7**. Immobilization reaction was carried out in the presence of protein (54 μM), aniline (100 mM), and PB (100 mM, pH 7). Release of hydrazone−GFP **9d** from agarose beads was carried out in the presence of hydroxylamine (200 mM), aniline (100 mM), and PB (200 mM, pH 7). Bright-field images are on the top, and fluorescent microscope images are on the bottom. Scale bars in the lower right-hand corners represent 200 μm.

in approximately 3 h 80% of the immobilized GFP was released from the beads, and accordingly, the beads became significantly less fluorescent (Figures 6B and 7). For comparison, the

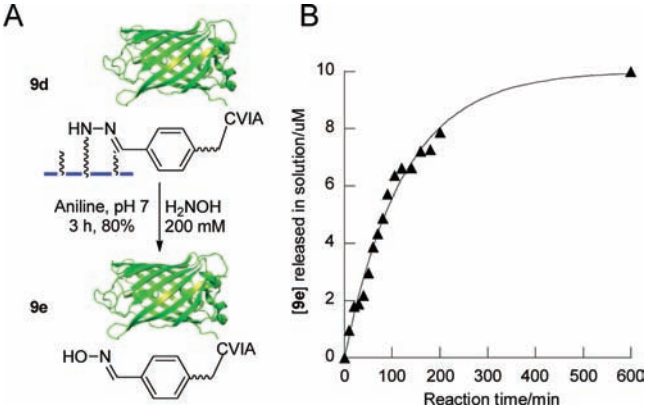


Figure 7. (A) Schematic representation of the release of immobilized GFP **9d** to yield **9e** from agarose beads via oxime ligation with hydroxylamine. (B) Kinetic analysis of the release of **9e** from agarose beads by oxime ligation. Reaction was carried out at room temperature in the presence of 100 mM aniline and 200 mM of hydroxylamine. UV absorbance of GFP in the supernatant was measured over time, which showed approximately 80% release of **9e** in 3 h. Analysis of the hydrolytic stability of **9d** in the absence of hydroxylamine and aniline showed no detectable release of GFP on the same time scale. Data was fit to a simple exponential decay process.

hydrolytic stability of immobilized GFP in the absence of hydroxylamine and aniline was also analyzed, and the results showed that the hydrazone bond in pH 7.5 Tris buffer was completely stable for 48 h with no detectable release of GFP.[56] This achievement highlights a significant advantage of this chemistry over click chemistry and other irreversible methods since it can be used to efficiently covalently immobilize proteins

dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455−8467

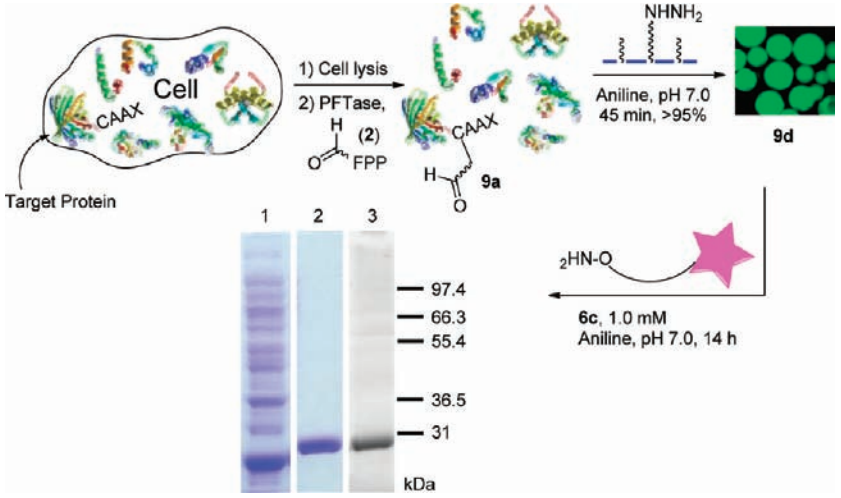


Figure 8. Chemoenzymatic site-specific tagging of proteins by aldehyde–FPP analogs by PFTase followed by capture of the aldehyde-functionalized protein in the crude cell lysate via hydrazide functionalized beads. Prenylation in the crude extract was confirmed by LC-MS analysis. Immobilized protein was then released into the solution or fluorescently labeled by addition of hydroxylamine or an aminooxy–fluorophore using aniline as the catalyst. SDS-PAGE analysis: lane 1, crude *E. coli* lysate containing **9a** visualized by Coomassie blue staining; lane 2, **9c** released from hydrazide beads after treatment with **6c** and visualized by Coomassie blue staining; lane 3, **9c** released from hydrazide beads after treatment with **6c** and visualized by gel fluorescence analysis.

onto solid surfaces and then release them under mild conditions without protein denaturation. Addition of aniline catalyzes the hydrolysis of hydrazone to hydrazide and aldehyde. Since oxime formation has a larger equilibrium constant than hydrazone formation,[57] the presence of hydroxylamine and aniline drives the equilibrium from hydrazone toward oxime formation and free hydrazide.

Enzymatic Modification, Immobilization, and Labeling in Crude Extract. An important feature of the labeling method described here is that it uses an enzymatic process for introduction of aldehyde groups into proteins. Due to the specificity of that biocatalytic process and the fact that there are no endogenous proteins in *E. coli* that contain a C-terminal CAAX box sequence, we reasoned that it should be possible to selectively functionalize proteins present in crude extract without purification. Additionally, once modified, it should also be possible to immobilize aldehyde-containing proteins and release them with an alkoxyamine that includes a fluorophore or PEG chain. In that way, a single protein present in *E. coli* crude extract could be modified, immobilized, and labeled without purification. To explore this, *E. coli* cells expressing GFP–CVIA were grown, lysed, and subjected to enzymatic prenylation using PFTase and substrate **2**. LC-ESI/MS analysis of the reaction mixture was employed to confirm introduction of the aldehyde functionality into GFP–CVIA **7** in the crude cell lysate. The reaction mixture was then concentrated, and excess **2** was removed via size-exclusion column chromatography (NAP-5 column). Aldehyde–GFP **9a** was then selectively immobilized from the crude cell lysate onto hydrazide-functionalized beads using aniline as the catalyst. Immobilization was followed by measuring the GFP absorbance present in solution and judged to be complete within 45 min, at which time the beads became highly fluorescent and the supernatant solution became almost colorless. Next, the beads were washed to remove any nonspecifically bound proteins and then treated with aminooxy fluorophore **6c** in the presence of 100 mM aniline overnight. SDS-PAGE analysis of the supernatant solution showed a single band (Figure 8, lane 2) migrating with an apparent mass of 29 kDa slightly higher than that of the starting GFP (due to addition of the aminooxy moiety) consistent with release of GFP. In-gel fluorescence analysis (Figure 8, lane 3) suggested that the released protein was labeled with the fluorophore **6c**; LC-MS analysis of the released protein provided additional evidence for formation of **9c**.

Application to Protein PEGylation. Attachment of polyethylene glycol (PEG) chains to proteins is the most widely used method for improving the pharmacokinetics of polypeptide-based therapeutic agents.[58–60] Current methods for PEGylation are generally nonselective and can result in a mixture of protein–PEG positional isomers with variable biological activity.[61] Site-specific methods offer a useful alternative approach for circumventing this problem of heterogeneity. Given our success in being able to incorporate a fluorescent label into a protein via the capture and release strategy described above, we decided to evaluate the utility of this approach for preparation of a PEGylated protein. Thus, aldehyde-functionalized GFP **9a** was first prepared from purified GFP **7** using PFTase as described above and treated with aminooxy-PEG-10 000 **10** to produce the protein–PEG conjugate **11**. Analysis of that material by MALDI-MS (Figure 9) showed an increase in molecular mass from 27.6 (for **9a**) to 38 kDa for **11**; the broader peak observed for **11** is consistent with attachment of a polydisperse polymer to a monodisperse protein. It is also important to note that no species resulting from addition of multiple PEG chains were observed, consistent with the selective nature of the chemistry employed here. Analysis of the PEGylation reaction mixture by SDS PAGE revealed a decrease in the electrophoretic mobility of **11** (Figure 10, lane 3) compared to the starting protein **9a** (Figure 10, lane 2). As noted in the MALDI MS, a wider band was observed for **11** relative to **9a**, again consistent with the polydisperse nature of the protein–PEG conjugate. With production of the PEGylated product clearly established, we next focused on generating the same material from **9a** that had not been purified chromatographically. Thus, **7** was prenylated with **2** using PFTase in crude *E. coli* extract followed by capture using hydrazide-functionalized agarose. After washing the

dx.doi.org/10.1021/ja211308s | *J. Am. Chem. Soc.* 2012, 134, 8455–8467

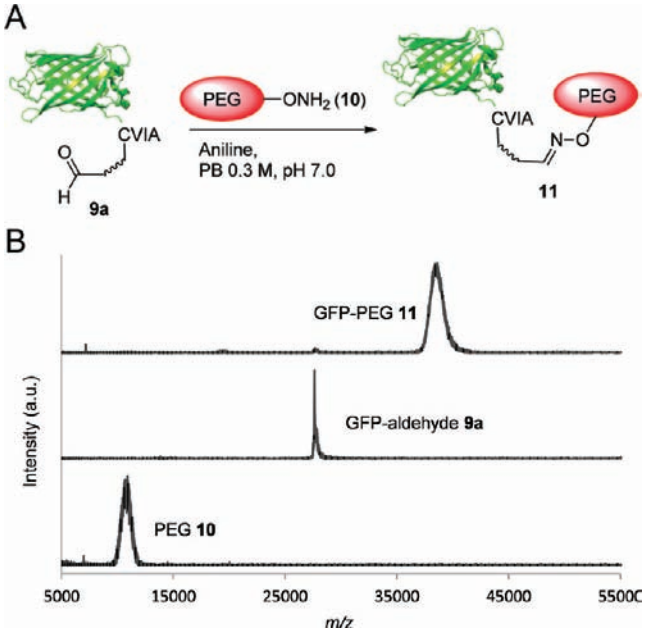


Figure 9. (A) Generation of site-specifically C-terminal PEGylated GFP from pure **9a**. (B) MALDI analysis of PEGylated GFP **11**. Lower panel is the MALDI spectrum of pure PEG **10**, middle panel is the MALDI spectrum of pure **9a**, and top panel is the MALDI spectrum of the oxime PEGylated GFP **11**, which confirms complete conversion. Reaction was performed using **9a** (10 μM) and **10** (100 μM) for 2 h. Excess of **10** was removed via a zip-tip protocol prior to MALDI analysis.

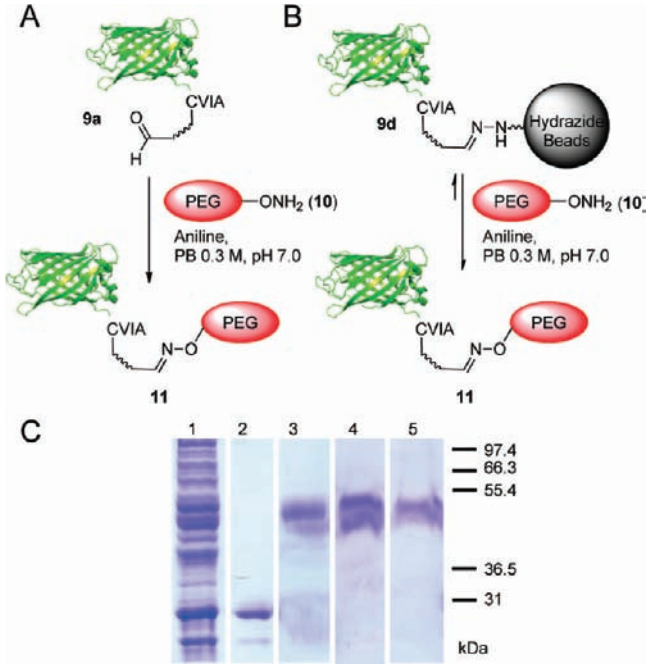


Figure 10. Use of PFTase-catalyzed protein modification for site-specific PEGylation from purified protein or crude cell lysate. (A) Generation of site-specific C-terminal PEGylated protein from pure **9a**. (B) PEGylation and release of immobilized **9d** from hydrazide beads using PEG **10**. (C) SDS PAGE analysis of PEGylated GFP (**11**) from purified **9a** or from immobilized protein **9d**. In the case of the crude cell lysate, **7** was chemoenzymatically and site-specifically tagged by aldehyde-containing analog **2** via PFTase-catalyzed reaction, followed by capture of the resulting aldehyde-functionalized protein from the lysate using hydrazide functionalized beads. Immobilized protein was then released back into solution and simultaneously site-specifically PEGylated by addition of aminooxy-PEG **10** using aniline as a catalyst. SDS-PAGE analysis: lane 1, crude *E. coli* lysate containing **9a**; lane 2, purified **9a**; lane 3, **11** produced by PEGylation of pure **9a** with **10**; lane 4, **11** prepared from **9d** (obtained using purified **9a**) and subsequently released with **10**; lane 5, **11** prepared from **9d** (obtained using **9a** present in crude lysate) and subsequently released with **10**.

material to remove nonspecifically bound proteins, the desired PEGylated protein (**11**) was eluted via treatment with **10** in the presence of aniline. SDS PAGE analysis showed the presence of a single band (Figure 10, lane 5) that comigrated with the authentic product prepared from pure **9a** (Figure 10, lane 3).

PEGylation of Glucose-Dependent Insulinotropic Polypeptide (GIP). The incretin, glucose-dependent insulinotropic polypeptide (GIP), is secreted from intestinal K-cells in response to nutrient ingestion and acts to augment insulin secretion in the pancreas. GIP has been proposed as a potential therapeutic agent for treatment of type 2 diabetes based on its stimulation of insulin secretion in the presence of elevated glucose levels;[62,63] however, efforts to bring GIP forward as a drug have been hampered due to its short circulating half-life. Recently, a modified form of GIP functionalized with a C-terminal mini-PEG group has shown resistance to proteolytic degradation while preserving biological activity in an obese rat model system.[52] Accordingly, having established the utility of our method for C-terminal site-specific modification described above with a model protein, GFP, we decided to demonstrate its utility for preparing a PEGylated form of GIP, a polypeptide with clear therapeutic potential. Thus, purified GIP−CVIM (**12a**), a form of GIP engineered to contain a C-terminal CAAX box (in this case CVIM[64]) was prenylated with analog **2** under conditions established above for GFP and subsequently PEGylated using a small aminooxy-functionalized PEG containing three ethylene glycol units (**13**). This shorter PEGylation reagent was employed since it is similar in length to what has previously been shown to be effective for increasing GIP stability in serum. MALDI MS analysis (Figure 11) confirmed successful prenylation and PEGylation of GIP; as noted above with GFP, both the enzymatic prenylation and the subsequent chemical PEGylation proceed with essentially complete conversion. Next, we employed the capture and release strategy developed above for GFP to prepare PEGylated GIP without prior purification. GIP−CVIM (**12a**), present in crude *E. coli* extract, was prenylated with **2**, and the resulting aldehyde-functionalized polypeptide **12b** was captured on hydrazide beads. The beads were washed extensively and then treated with aminooxy−PEG **13**, resulting in oxime formation and release into solution. MALDI MS analysis of the eluted material showed only the presence of PEGylated−GIP (**14**), indicating a high degree of specificity in the capture and release (Figure 12). Thus, this general method allows facile and effective purification of site-specifically PEGylated GIP from the crude cell extract. Overall, these experiments conclusively demonstrate how a protein, present in crude extract, can be selectively modified, labeled with a fluorophore or PEG polymer, and released in pure form via a simple process that requires no significant chromatographic steps. Given the specificity of the PFTase-catalyzed reaction coupled with the ability to introduce a CAAX-box onto almost any protein, this method shows great potential as a general approach for selective immobilization and labeling of recombinant proteins present in crude cellular extract without prior purification.

dx.doi.org/10.1021/ja211308s | *J. Am. Chem. Soc.* 2012, 134, 8455−8467

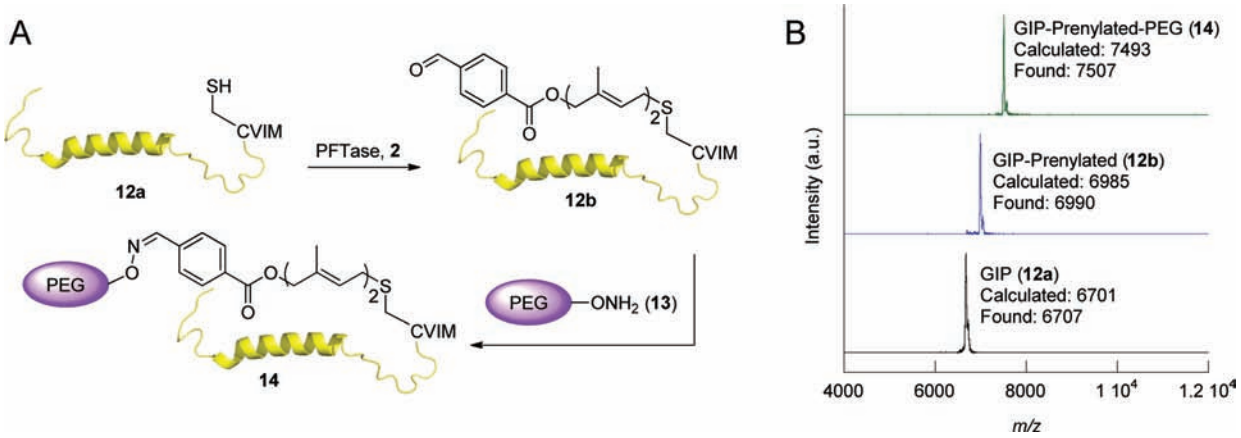


Figure 11. (A) Schematic representation of prenylation of glucose-dependent insulinotropic polypeptide (GIP) containing a CAAX-box positioned at its C-terminus (GIP-CVIM, **12a**) with aldehyde-containing analogue **2** to yield the prenylated product **12b**, which is then site-specifically PEGylated using a short chain aminooxy-PEG (**13**). (B) MALDI MS analysis of prenylation and PEGylation of GIP **12a**. MALDI MS spectra (from the top to the bottom) of oxime PEGylated GIP **14**, the prenylated aldehyde labeled GIP **12b**, and pure **12a**, respectively.

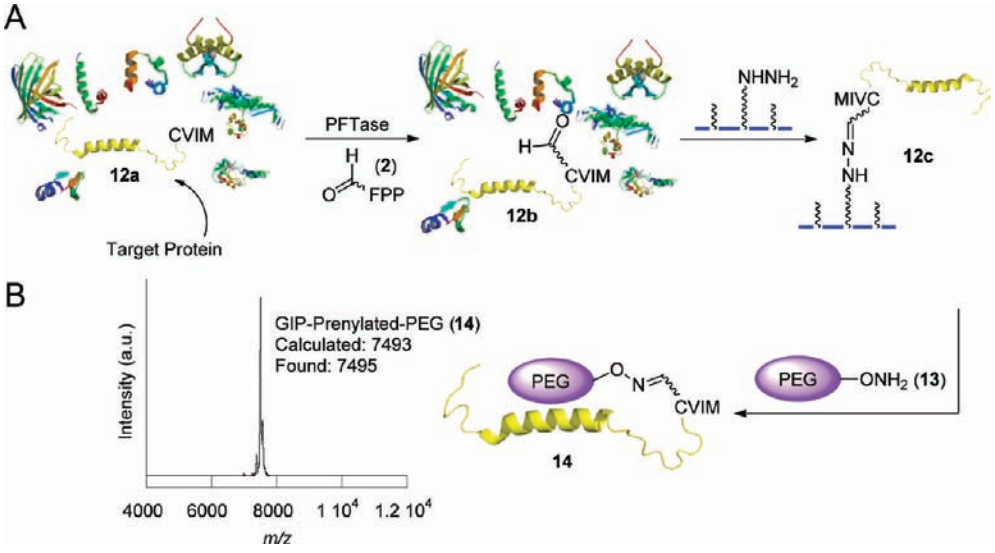


Figure 12. Use of PFTase-catalyzed protein modification for site-specific PEGylation of GIP **12a** from crude cell lysate. (A) Chemoenzymatic site-specific tagging of GIP **12a** by aldehyde–FPP analog **2** in the crude cell lysate via PFTase followed by capture of the aldehyde-functionalized polypeptide **12b** via hydrazide functionalized beads. Immobilized polypeptide was then released back into solution and simultaneously site-specifically PEGylated by addition of aminooxy-PEG **13**. (B) MALDI analysis of the released material confirmed formation and release of the pure PEGylated GIP (**14**) into the solution.

Beyond generating site-specifically modified proteins, this approach could greatly reduce the cost of producing PEGylated polypeptides for therapeutic applications due to the streamlined nature of the process.

■ CONCLUSION

In this work, we demonstrated that PFTase can be used to introduce aldehyde functionality near the C-terminus of a protein and that the resulting aldehyde-functionalized proteins can then be modified in a plethora of ways via aniline-catalyzed hydrazone or oxime ligation under mild conditions. We show that if the concentration of the aldehyde-functionalized protein is relatively high (>50 μM), hydrazone ligation is more efficient compared with oxime ligation due to its faster kinetics, whereas if the protein concentration is in the low micromolar range, oxime ligation is more advantageous due to its larger equilibrium constant. An important feature of the chemistry reported here is its reversible nature that can be harnessed to permit efficient release of proteins; covalent immobilization using hydrazone ligation of an aldehyde-containing protein can be followed by subsequent oxime formation to release the polypeptide without denaturation. Using synthetically modified alkoxyamines, a variety of new functionality ranging from fluorescent groups to PEG chains can be appended onto proteins. A second key feature of this approach concerns the enzymatic method for aldehyde incorporation. By capitalizing on the selectivity of the enzymatic process, the initial protein functionalization can be performed using unpurified protein substrates. The resulting modified protein can then be captured via hydrazone formation and released via oxime formation to produce a variety of pure, site-specifically modified protein conjugates. Such a streamlined approach for polypeptide modification could be particularly useful for large-scale production of protein conjugates for therapeutic or industrial applications. It should also be noted that the ligation chemistry described herein and the Cu(I)-catalyzed click reaction are

dx.doi.org/10.1021/ja211308s | J. Am. Chem. Soc. 2012, 134, 8455–8467

orthogonal. This opens up the possibility of performing multiple modifications on proteins using different bioorthogonal chemistries. Given that CAAX-box sequences can be appended to the C-terminus of almost any protein, the method reported here should be useful for a variety of applications in protein chemistry.

■ ASSOCIATED CONTENT

Supporting Information

Synthetic procedures and additional data and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

E-mail: diste001@umn.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Joseph Dalluge for helpful discussions regarding mass spectrometry. This work was supported by the National Institutes of Health (GM058842 and GM084152), the University of Minnesota, and the Minnesota Supercomputer Institute.

■ REFERENCES

(1) Zhu, H.; Bilgin, M.; Bangham, R.; Hall, D.; Casamayor, A.; Bertone, P.; Lan, N.; Jansen, R.; Bidlingmaier, S.; Houfek, T.; Mitchell, T.; Miller, P.; Dean, R. A.; Gerstein, M.; Snyder, M. *Science* **2001**, *293*, 2101−2105.
(2) Luk, Y.-Y.; Tingey, M. L.; Dickson, K. A.; Raines, R. T.; Abbott, N. L. *J. Am. Chem. Soc.* **2004**, *126*, 9024−9032.
(3) Weinrich, D.; Lin, P.-C.; Jonkheijm, P.; Nguyen, U. T. T.; Schröder, H.; Niemeyer, C. M.; Alexandrov, K.; Goody, R.; Waldmann, H. *Angew. Chem., Int. Ed.* **2010**, *49*, 1252−1257.
(4) Zhong, M.; Fang, J.; Wei, Y. *Bioconjugate Chem.* **2010**, *21*, 1177−1182.
(5) Lin, P.-C.; Ueng, S.-H.; Tseng, M.-C.; Ko, J.-L.; Huang, K.-T.; Yu, S.-C.; Adak, A. K.; Chen, Y.-J.; Lin, C.-C. *Angew. Chem., Int. Ed.* **2006**, *45*, 4286−4290.
(6) Goldstein, D. C.; Thordarson, P.; Peterson, J. R. *Aust. J. Chem.* **2009**, *62*, 1320.
(7) Wong, L. S.; Khan, F.; Micklefield, J. *Chem. Rev.* **2009**, *109*, 4025−4053.
(8) Lin, P.-C.; Weinrich, D.; Waldmann, H. *Macromol. Chem. Phys.* **2010**, *211*, 136−144.
(9) Wu, B.-Y.; Hou, S.-H.; Huang, L.; Yin, F.; Zhao, Z.-X.; Anzai, J.-I.; Chen, Q. *Mater. Sci. Eng., C* **2008**, *28*, 1065−1069.
(10) Yamaguchi, H.; Miyazaki, M.; Honda, T.; Briones-Nagata, M. P.; Arima, K.; Maeda, H. *Electrophoresis* **2009**, *30*, 3257−3264.
(11) Lee, Y.; Lee, E. K.; Cho, Y. W.; Matsui, T.; Kang, I.-C.; Kim, T.-S.; Han, M. H. *Proteomics* **2003**, *3*, 2289−2304.
(12) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proc. Natl. Acad. Sci.* **2004**, *101*, 9955−9959.
(13) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 7754−7755.
(14) Chen, I.; Ting, A. Y. *Curr. Opin. Biotechnol.* **2005**, *16*, 35−40.
(15) Torchilin, V. P.; Lukyanov, A. N. *Drug Discovery Today* **2003**, *8*, 259−266.
(16) Discher, D. E. *Science* **2002**, *297*, 967−973.
(17) Rizzi, S. C.; Hubbell, J. A. *Biomacromolecules* **2005**, *6*, 1226−1238.
(18) Kurpiers, T.; Mootz, H. D. *Angew. Chem., Int. Ed.* **2009**, *48*, 1729−1731.
(19) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346−16347.
(20) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004−2021.
(21) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518−13519.
(22) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. *Angew. Chem., Int. Ed.* **2008**, *47*, 2832−2835.
(23) Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 11486−11493.
(24) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci.* **2007**, *104*, 16793−16797.
(25) Lallana, E.; Fernandez-Megia, E.; Riguera, R. *J. Am. Chem. Soc.* **2009**, *131*, 5748−5750.
(26) Baskin, J. M.; Dehnert, K. W.; Laughlin, S. T.; Amacher, S. L.; Bertozzi, C. R. *Proc. Natl. Acad. Sci.* **2010**, *107*, 10360−10365.
(27) Cordes, E. H.; Jencks, W. P. *J. Am. Chem. Soc.* **1962**, *84*, 826−831.
(28) Dirksen, A.; Yegneswaran, S.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2010**, 2023−2027.
(29) Zeng, Y.; Ramya, T. N. C.; Dirksen, A.; Dawson, P. E.; Paulson, J. C. *Nat. Methods* **2009**, *6*, 207−209.
(30) Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc. Natl. Acad. Sci.* **2004**, *101*, 13132−13137.
(31) Yi, L.; Sun, H.; Wu, Y.-W.; Triola, G.; Waldmann, H.; Goody, R. S. *Angew. Chem., Int. Ed.* **2010**, *49*, 9417−9421.
(32) Rashidian, M.; Dozier, J. K.; Lenevich, S.; Distefano, M. D. *Chem. Commun.* **2010**, *46*, 8998.
(33) Park, S.; Yousaf, M. N. *Langmuir* **2008**, *24*, 6201−6207.
(34) Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602−15603.
(35) Rayo, J.; Amara, N.; Krief, P.; Meijler, M. M. *J. Am. Chem. Soc.* **2011**, *133*, 7469−7475.
(36) Dixon, H. B. F. *J. Protein Chem.* **1984**, *3*, 99−108.
(37) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307−5311.
(38) Rush, J. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 12240−12241.
(39) Wang, L.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2005**, *44*, 34−66.
(40) Turek, T. C.; Gaon, I.; Distefano, M. D.; Strickland, C. L. *J. Org. Chem.* **2001**, *66*, 3253−3264.
(41) Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D. *ChemBioChem* **2007**, *8*, 98−105.
(42) Gauchet, C.; Labadie, G. R.; Poulter, C. D. *J. Am. Chem. Soc.* **2006**, *128*, 9274−9275.
(43) Duckworth, B. P.; Chen, Y.; Wollack, J. W.; Sham, Y.; Mueller, J. D.; Taton, T. A.; Distefano, M. D. *Angew. Chem.* **2007**, *119*, 8975−8978.
(44) Labadie, G. R.; Viswanathan, R.; Poulter, C. D. *J. Org. Chem.* **2007**, *72*, 9291−9297.
(45) Subramanian, T.; Liu, S.; Troutman, J. M.; Andres, D. A.; Spielmann, H. P. *ChemBioChem* **2008**, *9*, 2872−2882.
(46) Kale, T. A.; Hsieh, S. J.; Rose, M. W.; Distefano, M. D. *Curr. Top. Med. Chem.* **2003**, *3*, 1043−1074.
(47) Mu, Y.; Gibbs, R. A.; Eubanks, L. M.; Poulter, C. D. *J. Org. Chem.* **1996**, *61*, 8010−8015.
(48) Placzek, A. T.; Gibbs, R. A. *Org. Lett.* **2011**, 3576−3579.
(49) Wollack, J. W.; Silverman, J. M.; Petzold, C. J.; Mougous, J. D.; Distefano, M. D. *ChemBioChem* **2009**, *10*, 2934−2943.
(50) Yakhnin, A. V.; Vinokurov, L. M.; Surin, A. K.; Alakhov, Y. B. *Protein Expression Purif.* **1998**, *14*, 382−386.
(51) Inagaki, N.; Seino, Y.; Takeda, J.; Yano, H.; Yamada, Y.; Bell, G. I.; Eddy, R. L.; Fukushima, Y.; Byers, M. G.; Shows, T. B.; Imura, H. *Mol. Endocrinol.* **1989**, *3*, 1014−1021.
(52) Gault, V.; Kerr, B.; Irwin, N.; Flatt, P. *Biochem. Pharmacol.* **2008**, *75*, 2325−2333.
(53) Duckworth, B. P.; Xu, J.; Taton, T. A.; Guo, A.; Distefano, M. D. *Bioconjugate Chem.* **2006**, *17*, 967−974.

(54) Gaon, I.; Turek, T.; Distefano, M. *Tetrahedron Lett.* **1996**, *37*, 8833−8836.
(55) Lenevich, S.; Distefano, M. D. *Anal. Biochem.* **2011**, *408*, 316−320.
(56) Kalia, J.; Raines, R. T. *Angew. Chem., Int. Ed.* **2008**, *47*, 7523−7526.
(57) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581−7584.
(58) Jevsevar, S.; Kunstelj, M.; Porekar, V. G. *Biotechnol. J.* **2010**, *5*, 113−128.
(59) Veronese, F. M.; Mero, A.; Pasut, G. In *PEGylated Protein Drugs: Basic Science and Clinical Applications*; Veronese, F. M., Ed.; Birkhäuser Basel: Basel, 2009; pp 11−31.
(60) Veronese, F. M. *Biomaterials* **2001**, *22*, 405−417.
(61) Roberts, M. J.; Bentley, M. D.; Harris, J. M. *Adv. Drug Delivery Rev.* **2002**, *54*, 459−476.
(62) Flatt, P. R. *Diabet. Med.* **2008**, *25*, 759−764.
(63) Elahi, D.; McAloon-Dyke, M.; Fukagawa, N. K.; Meneilly, G. S.; Sclater, A. L.; Minaker, K. L.; Habener, J. F.; Andersen, D. K. *Regul. Pept.* **1994**, *51*, 63−74.
(64) Friday, B. B.; Adjei, A. A. *Biochim. Biophys. Acta* **2005**, *1756*, 127−144.