J|A|C|S
ARTICLES
Published on Web 08/26/2006

Synthesis and Evaluation of Bioorthogonal Pantetheine Analogues for in Vivo Protein Modification

Jordan L. Meier, Andrew C. Mercer, Heriberto Rivera, Jr., and Michael D. Burkart*

Contribution from the Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0358

Received May 8, 2006; E-mail: mburkart@ucsd.edu

Abstract: In vivo carrier protein tagging has recently become an attractive target for the site-specific modification of fusion systems and new approaches to natural product proteomics. A detailed study of pantetheine analogues was performed in order to identify suitable partners for covalent protein labeling inside living cells. A rapid synthesis of pantothenamide analogues was developed and used to produce a panel which was evaluated for in vitro and in vivo protein labeling. Kinetic comparisons allowed the construction of a structure−activity relationship to pinpoint the linker, dye, and bioorthogonal reporter of choice for carrier protein labeling. Finally bioorthogonal pantetheine analogues were shown to target carrier proteins with high specificity in vivo and undergo chemoselective ligation to reporters in crude cell lysate. The methods demonstrated here allow carrier proteins to be visualized and isolated for the first time without the need for antibody techniques and set the stage for the future use of carrier protein fusions in chemical biology.

Introduction

Recent years have seen intense research effort focused toward the development of new methods for the study and manipulation of covalently modified proteins, with particular attention given to in vivo methodologies.[1] Fluorescent protein fusions[2] and antibody conjugates[3] provide powerful tools for protein imaging and manipulation. However drawbacks of these methods, such as structural perturbations sometimes induced by large fusions and general membrane impermeability of antibodies, have lead researchers to devise methods for the site-specific modification of proteins by small-molecule probes. Ideally these probes should be low molecular weight, covalent in nature, and possessed of fluorescence or affinity properties allowing for facile imaging and manipulation. We recently introduced one such technique, demonstrating cellular uptake and covalent modification of carrier protein fusions by pantetheine analogues.[4] These coenzyme A (CoA) precursors were shown to penetrate the cell membrane and be transformed into fully formed CoA derivatives via the endogenous CoA metabolic pathway, whereupon they were transferred to a carrier protein by the promiscuous phosphopantetheinyltransferase (PPTase; E.C. 2.7.7.7) Sfp (Figure 2). This advance allows carrier protein labeling, a technique first developed from cell lysates[5] and since demonstrated on the cell surface,[6] to be performed within the cell, opening the door for more sophisticated labeling systems. Recent developments have seen the trimming of the carrier protein domain down to just 11 amino acids,[7] offering a fusion tag of the size and flexibility to be competitive with contemporary tagging systems and further highlighting the importance of techniques for the labeling of intracellular carrier proteins.

Several strategies for site-specific labeling of proteins in vivo have been previously demonstrated. Examples include Bertozzi's manipulation of the sialic acid biosynthetic pathway for the introduction of keto and azido functionalized cell-surface glycoproteins,[8] Cravatt's introduction of azido/alkyne functionalities by covalent irreversible inhibition of protein active sites,[9] and Hsieh−Wilson's chemoenzymatic introduction of a keto-functionality for capture of *O*-GlcNAc-modified proteins.[10] In each of these examples the protein is not directly labeled with a fluorescence or affinity tag, but rather a unique and biologically inert chemical functionality is introduced. This functional-

(1) Bertozzi, C. R.; Prescher, J. A. *Nat. Chem. Biol.* **2005**, *1*, 13-21.
(2) (a) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509−544. (b) Lippincott-Schwartz, J.; Patterson, G. H. *Science* **2003**, *300*, 87−91.
(3) (a) Fritze, C. E.; Anderson, T. R. *Methods Enzymol.* **2000**, *327*, 3−16. (b) Massoud, T. F.; Gambhir, S. S. *Genes Dev.* **2003**, *17*, 545−580.
(4) Clarke, K. M.; Mercer, A. C.; La Clair, J. J.; Burkart, M. D. *J. Am. Chem. Soc.* **2005**, *127*, 11234−11235.
(5) (a) La Clair, J. J.; Foley, T. L.; Schegg, T. R.; Regan, C. M.; Burkart, M. D. *Chem. Biol.* **2004**, *11*, 195−201. (b) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 7754−7755. (c) Yin, J.; Liu, F.; Schinke, M.; Daly, C.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 13570−13571.
(6) (a) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2004**, *126*, 8896−8897. (b) Yin, J.; Lin, A. J.; Buckett, P. D.; Wessling-Resnick, M.; Golan, D. E.; Walsh, C. T. *Chem. Biol.* **2005**, *12*, 999−1006. (c) Vivero-Pol L.; George, N.; Krumm, H.; Johnsson, K.; Johnsson, N. *J. Am. Chem. Soc.* **2005**, *127*, 12770−12771.
(7) Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15815−15820.
(8) (a) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125−1128. (b) Yarema, K. J.; Mahal, L. K.; Bruehl, R.; Rodriguez, E. C.; Bertozzi, C. R. *J. Biol. Chem.* **1998**, *273*, 31168−31179. (c) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007−2010.
(9) (a) Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535−546. (b) Speers, A. E.; Adam, G. C.; Cravatt, B. F. *J. Am. Chem. Soc.* **2003**, *125*, 4686−4687. (c) Alexander, J. P.; Cravatt, B. F. *Chem Biol.* **2005**, *12*, 1179−1187.
(10) (a) Hwan-Ching, T.; Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2004**, *126*, 10500−10501. (b) Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *36*, 13132−13137.

10.1021/ja063217n CCC: $33.50

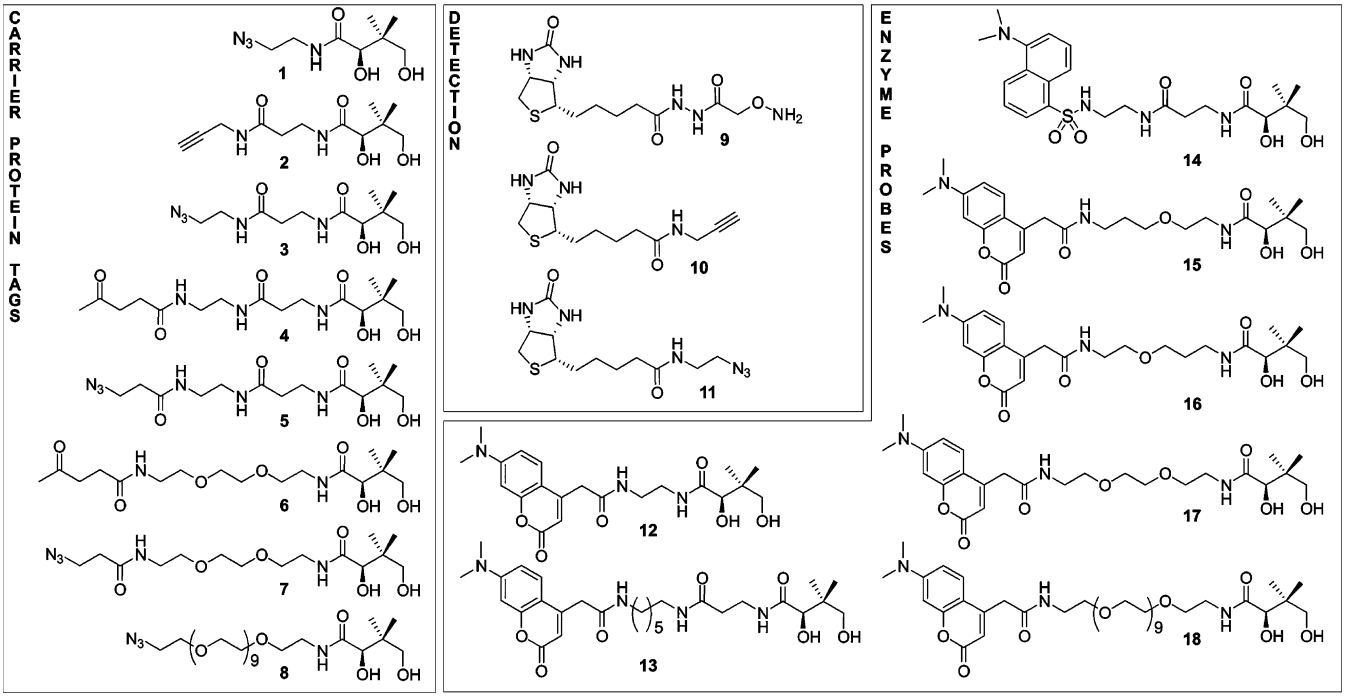


Figure 1. Structures of pantetheine analogues and biotin detection agents used in this study.

ity can then undergo reaction with exogenously delivered reporters to label the protein of interest for detection and/or isolation, depending on the nature of the reporter. Advantages of this two-step labeling process include (i) better uptake of smaller probes due to increased membrane permeability, (ii) increased incorporation of probes into native biosynthetic pathways due to a greater similarity to natural substrate, and (iii) the ability to conjugate a protein to virtually any reporter possessing reactivity with the bioorthogonal functionality.[1,9] Here we present a full study of simplified pantetheine analogues that harness the power of such bioorthogonal ligation reactions. First we optimize the synthesis of simplified pantetheine analogues via a one-step reaction with pantolactone. Next, the specificity of the CoA biosynthetic pathway is probed by a small panel of these simplified substrates. Finally, we validate the utility of this strategy by demonstrating and comparing the delivery of bioorthogonal chemical functionalities to carrier proteins in vitro and in vivo and using the newly tagged carrier proteins for two widely applied chemoselective ligations: the reaction of ketones and hydroxylamines to form oximes and the Cu(I)-catalyzed azide−alkyne [3 + 2] cycloaddition reaction ("click" chemistry). This new ability to manipulate a bioorthogonally tagged carrier protein in vivo promises to be a valuable tool for both new approaches to natural product proteomics as well as the study of novel intracellular carrier protein fusion systems.

Results and Discussion

Analogue Synthesis: Pantolactone-Ring Opening. In our efforts to address CoA biosynthesis with novel analogues, we had initially investigated the synthesis of pantetheine and phosphopantetheine derivatives that could be assembled in a manner analogous to peptide library synthesis.[11] This necessitated addressing the synthetic challenges associated with pantolactone, namely the lability of the α-proton following protection of the pantolactone secondary alcohol. At this point, the presumption was made that the identities of cystamine and β-alanine were necessary for turnover by the CoA metabolic pathway. However our own in vitro studies and recent work by Lee[12] questioned the degree of selectivity gained through specific interactions between the β-alanine moiety of pantothenate and PANK (used in this manuscript to denote all enzymes with pantothenate kinase activity; E.C. 2.7.1.33), the first enzyme in the CoA biosynthetic pathway and gatekeeper for downstream metabolism.[13] Further, we found that *E. coli* PANK (CoaA) catalyzed the phosphorylation of pantetheine and analogues with variations at the cystamine moiety almost as well as pantothenate itself.[14] Given this newly revealed permissiveness in the CoA biosynthetic pathway, we reasoned that simplified analogues of pantetheine could be used for both in vitro and in vivo applications. Elimination of the amide bond between cystamine and β-alanine significantly simplifies synthetic access to reporter-modified pantetheine analogues by reducing overall molecule polarity and solubility issues, eliminating time-consuming protection/deprotection steps of the 1,3-diol, and replacing the multiple peptide coupling and purification steps of previous syntheses[4] with a simple one-step nucleophilic ring-opening of pantolactone. With this in mind we chose to mimic the aletheine moiety (*N*-(β-alanyl)-β-aminoethanethiol) with more synthetically flexible poly(ethylene glycol) (PEG) linkers. In addition to the synthetic utility of this substitution, PEG spacers have the advantages of increasing the aqueous solubility of small molecule probes and distancing reporter labels from a labeled protein with the effect of both enhancing secondary detection properties and reducing any negative effect of reporter/protein interactions.[15] We sought to incorporate these

(11) Mandel, A. L.; La Clair, J. J.; Burkart, M. D. *Org. Lett.* **2004**, *6*, 4801−4803.
(12) Virga, K. G.; Zhang, Y. M.; Leonardi, R.; Ivey, R. A.; Hevener, K.; Park, H. W.; Jackowski, S.; Rock, C. O.; Lee, R. E. *Bioorg. Med. Chem.* **2006**, *14*, 1007−1020.
(13) Jackowski, S.; Rock, C. O. *J. Bacteriol.* **1981**, *148*, 926−932.
(14) Worthington, A. S.; Burkart, M. D. *Org. Biomol. Chem.* **2005**, *4*, 44−46.
(15) Kumar, V.; Aldrich, J. V. *Org. Lett.* **2003**, *5*, 613−616.

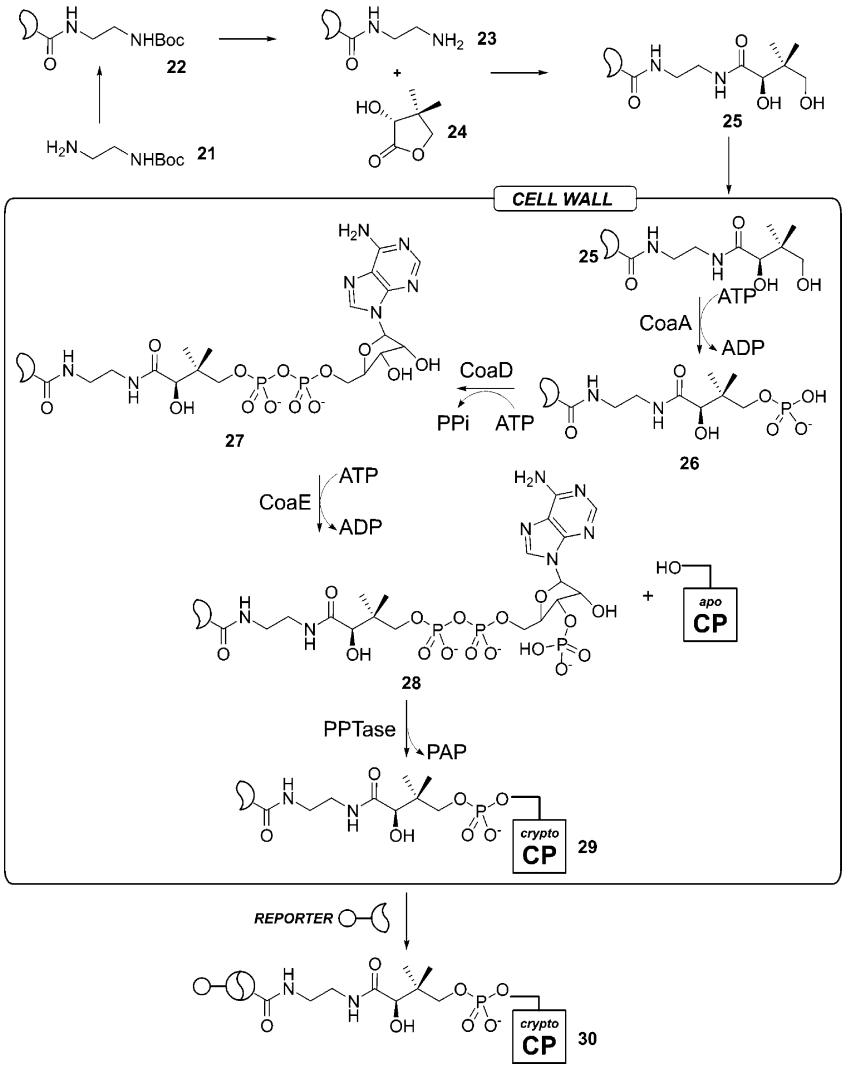


Figure 2. General strategy for in vivo labeling of carrier protein by pantetheine analogues. Virtually any monoprotected amine (**21**) can be transformed into a pantetheine analogue (**25**) by the three-step coupling/deprotection/ring-opening sequence. Cellular uptake and biosynthetic processing by CoaA, CoaD, and CoaE yield the CoA analogue **28**, which is then transferred to the carrier protein by a PPTase to yield bioorthogonally labeled carrier protein **28**. After cell lysis this carrier protein can now be conjugated to the reporter of choice via an appropriate chemoselective ligation reaction.

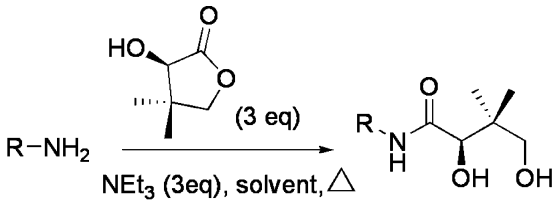
advantages into the design of an ideal, synthetically straightforward, biodetectible pantetheine analogue.

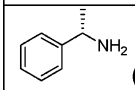
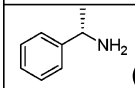
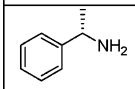
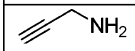
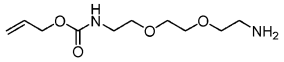
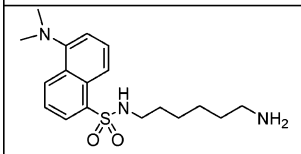
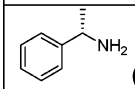
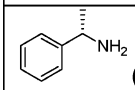
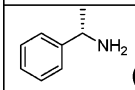
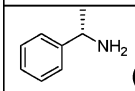
To quickly access a large selection of analogues it was deemed appropriate to first revise the current methodology for pantothenamide synthesis. Previous protocols calling for the base-promoted nucleophilic ring opening of pantolactone by an amine could be subject to racemization or hampered by long reaction times (>24 h).[12,16,17] To address these problems we turned to microwave-assisted organic synthesis. By using (*S*)-(−)-α-methylbenzylamine we were able to test a variety of conditions for their ability to open pantolactone with a fairly hindered chiral nucleophile and analyze enantiopurity by ^{1}H NMR (Table 1, see Supporting Information for ^{1}H NMR data).

The study showed pantolactone to be surprisingly robust to a variety of conditions, and reaction times could be reduced nearly 50-fold compared to previous preparations with a retention of optical purity. As expected from the hypothesized transition state of this reaction, protic solvents proved ideal for nucleophilic ring-opening, with ethanol providing the best balance of energy-absorbance and solubilization. Moving from our model system to usefully functionalized amines, it was shown that alkyne (**41**), PEG (**44**), and fluorophore (**20**) containing pantetheine analogues could be synthesized in good to moderate yields within 30 min using microwave assistance. Interestingly a very recent report also presented ethanol as the solvent of choice for this transformation under thermal conditions; however without the addition of any base these large-scale syntheses suffered from very long reaction times (72−120 h).[18] Accordingly our ideal microwave reaction conditions were also tested under simple reflux. Triethylamine proved to be a sufficient base, as replacement with Hunig's base showed no significant effect on the reaction outcome. Stronger bases were avoided. Again it was found that reflux of (*S*)-(−)-α-methylbenzylamine with excess pantolactone and triethylamine provided pantetheine analogues with no apparent racemization in excellent yields in 7−12 h. This alternative synthesis provides

(16) Dueno, E. E.; Chu, F.; Kim, S. I.; Jung, K. W. *Tetrahedron Lett.* **1999**, *40*, 1843−1846.
(17) (a) Michelson, A. M. *Biochim. Biophys. Acta* **1964**, *93*, 71−77. (b) Moffatt, J. G.; Khorana, H. G. *J. Am.* Chem. Soc. **1961**, *83*, 663−675.
(18) Krause, B. R. et al. *Synth. Commun.* **2006**, *36*, 365−391.

Table 1. Data Table for One-Step Synthesis of Pantetheine Analogues via Nucleophilic Ring-Opening of Pantolactone



Amine	Solvent	Temp (°C)	Time (hr)	Yield
 (**19**)	EtOH	160 (a)	0.5	91%
 (**19**)	DMF	165 (a)	0.5	63%
 (**19**)	THF	110 (a)	0.5	44%
 (**41**)	EtOH	160 (a)	0.5	42%
 (**44**)	EtOH	160 (a)	0.5	82%
 (**20**)	EtOH	160 (a)	0.5	75%
 (**19**)	EtOH	Reflux (b)	7	97%
 (**19**)	MeOH	Reflux (b)	7	84%
 (**19**)	CH_3CN	Reflux (b)	22	83%
 (**19**)	DME	Reflux (b)	12	30%

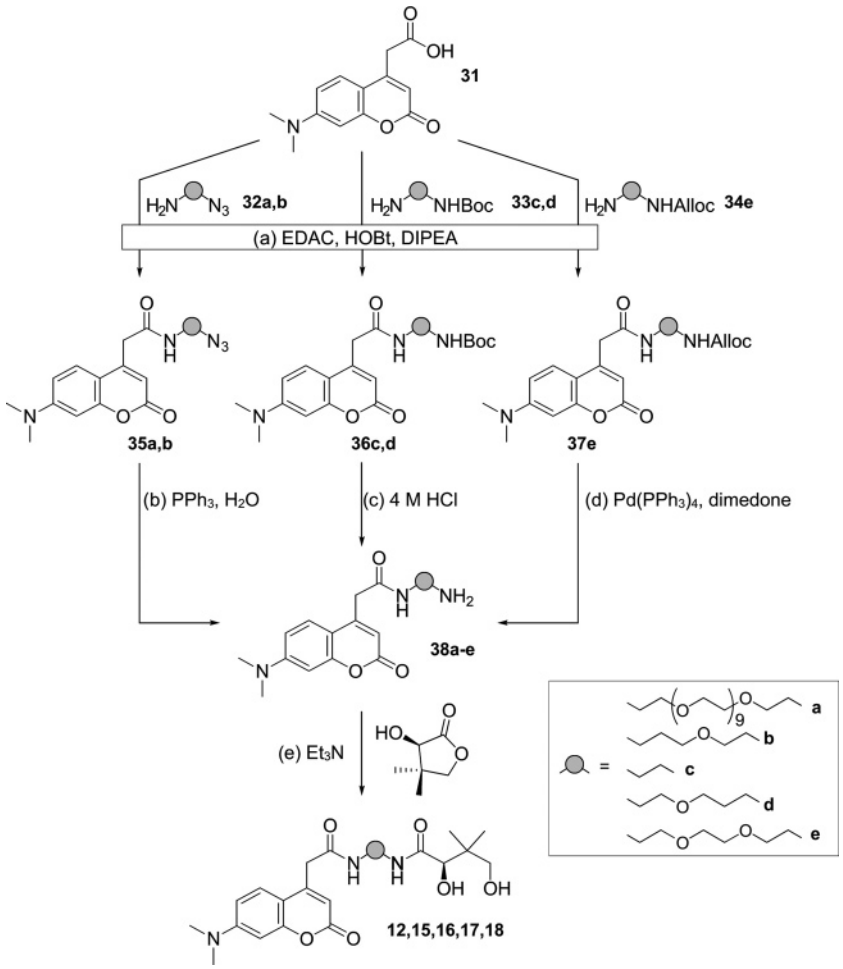
[a] Microwave-assisted. [b] Thermal condition.

another avenue for analogue preparation in cases where the reporter or linker is sensitive to decomposition under microwave conditions.[19]

Synthesis of Bioorthogonal and Fluorescent Pantetheine Analogues. The general strategy for chemoenzymatic synthesis of CoA analogues is depicted in Figure 2. First a monoprotected amine (depicted in this example by *N*-Boc ethylenediamine **21**) is conjugated to the biodetectible tag of choice by standard peptide coupling conditions. After deprotection, this amine can be conjugated to either pantolactone **24** through nucleophilic ring opening or pantothenic acid via EDAC mediated coupling. The newly formed pantetheine analogue is then processed via

(19) Microwave-assisted conjugation of 7-dimethylaminocoumarin-4-acetic acid containing amines to pantolactone resulted in the formation of unidentified decomposition products. For these couplings the classical condition (MeOH, NEt_3, reflux) was used.

Scheme 1. Synthesis of CoaA Probes[a]



[a] (a) EDAC (2 equiv), DIPEA (2 equiv), DMF, rt 12 h; (b) PPh_3 (1.2 equiv), 10:1 THF/H_2O, rt 12−24 h; (c) 4 M HCl/dioxanes, rt 24 h; (d) $Pd(PPh_3)_4$ (cat.), PPh_3 (2 equiv), dimedone (7 equiv), rt 12 h; (e) pantolactone (3 equiv), NEt_3 (3 equiv), MeOH, reflux 24−72 h.

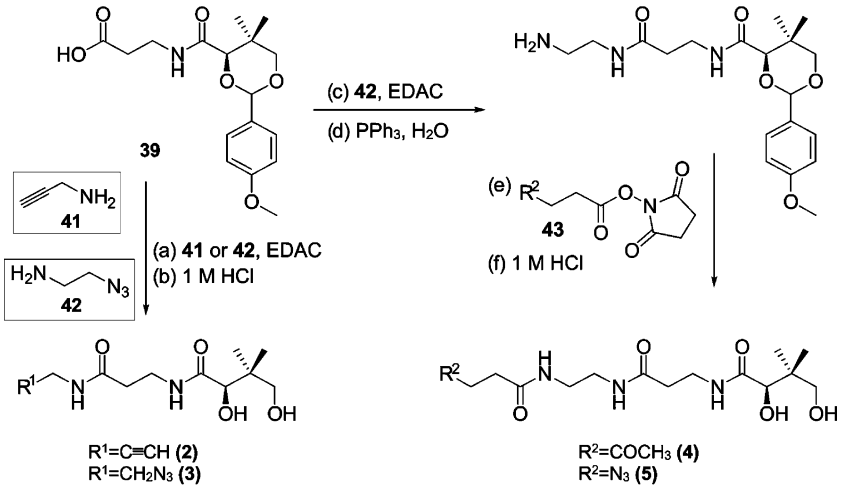
stepwise conversion by CoaA, a phosphopantetheine adenylytranserase (CoaD; E.C. 2.7.7.3), and a dephospho-CoA kinase (CoaE; E.C. 2.7.1.24) to form CoA analogue **28**, which is subsequently transferred to a conserved residue of the carrier protein by a PPTase to produce reporter-modified *crypto*-carrier protein **30**.

Analogues **12**−**18** were synthesized by this route in order to test the permissibility of CoA biosynthesis toward unnatural pantetheine analogues, particularly the effect of changes in the β-alanine/cystamine region (Scheme 1). We chose three parallel amino-protecting group strategies (azide, Boc, Alloc) based on the commercial availability (**32a**), simple synthesis from literature preparations (**33c**, **34e**), and orthogonal protecting group traits (**32b**, **33d**) of the specified diamines (Scheme 1). This strategy allowed compound **12** to be synthesized in two steps from *N*-Boc ethylenediamine conjugated 7-dimethylaminocoumarin-4-acetic acid (DMACA) **36c**.[20] Acid-catalyzed deprotection afforded the free amine, which performed nucleophilic ring opening of pantolactone to afford the final product in 85% yield. Compound **13** was synthesized as previously described,[4] while dansylated pantetheine **14** was synthesized by an analogous route. Compounds **15** and **16** were chosen to probe the effect of replacement of the strong H-bond accepting carbonyl and H-bond donating nitrogen of the natural substrate amide with weak H-bond accepting ether oxygens. Their synthesis made use of a common orthogonally protected diamine **48** (see Supporting Information), which underwent differential deprotection to give monoprotected diamines **32b** and **33d**. Subsequent EDAC mediated conjugation to dye **31**, azido/Boc deprotection by standard conditions, and nucleophilic ring opening of pantolactone afforded enzyme probes **15** and **16**. Our interest in incorporating the advantageous properties of PEG spacers in our library of pantetheine analogues lead us to synthesize PEG-linked-pantoic acid conjugate **17** in a 41% overall yield through an analogous dye conjugation/deprotection/nucleophilic ring-opening sequence starting from previously described *N*-Alloc diaminoethylene glycol **34e**. Compound **18** was chosen to test the limits of linker length in CoA biosynthesis and was easily attainable from the commercially available monoazido/monoamino terminal nonaethylene glycol **32a** through a similar series of reactions.

Compounds **1**−**8** sought to create a wide spectrum of pantetheine analogues incorporating some of the most commonly used bioorthogonal tags such as ketones, azides, and alkynes. Particular attention was paid to the azido moiety, where analogues replacing the β-Ala (**1**), cystamine (**2**), and thiol (**5**) regions of natural pantetheine with terminal azide moieties were

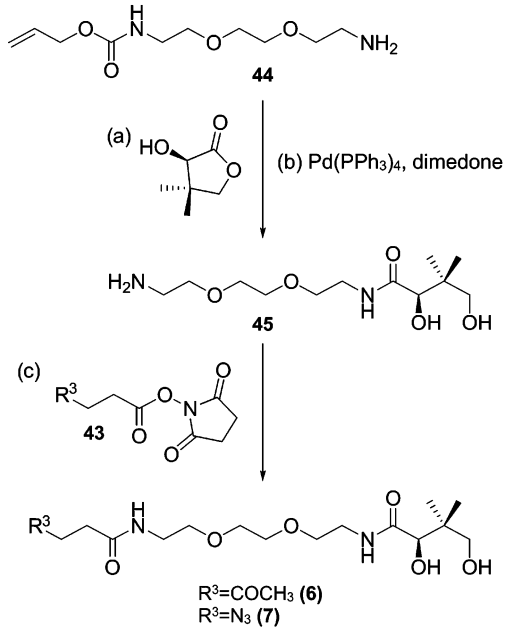
(20) La Clair, J. J.; et al. *ChemBioChem* **2006**, *7*, 409−416.

Scheme 2. Synthesis of β-Ala-Linked Bioorgthogonal Pantetheine Analogues[a]



[a] (a) **41** or **42**, EDAC (2 equiv), HOBt (2.5 equiv), DIPEA (2 equiv), DMF, rt 12 h; (b) 1 M HCl, 1:1 THF/H_2O, rt 1 h; (c) **42**, EDAC (2 equiv), HOBt (2.5 equiv), DIPEA (2 equiv), DMF, rt 12 h; (d) PPh_3 (2 equiv), 10:1 THF/H_2O, rt 12 h; (e) NHS ester **43a** or **43b**, NEt_3 (2 equiv), rt 8 h; (f) 1 M HCl, 1:1 THF/H_2O rt 1 h.

Scheme 3. Synthesis of PEG-Linked Bioorthogonal Pantetheine Analogues[a]



[a] (a) Pantolactone (3 equiv), NEt_3 (3 equiv), EtOH, 160 °C 0.5 h; (b) $Pd(PPh_3)_4$ (0.05 equiv), PPh_3, dimedone; (c) NHS ester **43a** or **43b**, NEt_3 (2 equiv), rt 4 h.

synthesized, in addition to experimentation with short (**7**) and long (**8**) PEG-linked azido-pantetheine analogues. Compound **1** was synthesized in one step from the nucleophilic ring-opening of pantolactone by 2-azidoethanamine (**42**). Compounds **2** and **3** (Scheme 2) were synthesized by standard peptide coupling of *p*-methoxybenzylideneacetal-protected (PMB) pantothenate **39** and the corresponding alkynyl and azido-amines, followed by acidic deprotection of the 1,3-diol. Compounds **4** and **5** again utilized **39** as a starting material, which was coupled to 2-azidoethanamine, deprotected to the corresponding amine, and coupled to the corresponding *N*-hydroxysuccinimidyl keto/azido (**43a,b**) ester. PEG linked pantetheine conjugates **6** and **7** (Scheme 3) were constructed in a similar fashion by nucleophilic ring-opening of pantolactone under microwave-conditions by *N*-Alloc protected diaminoethylene glycol **44** to give a common intermediate, followed by $Pd(PPh_3)_4$ mediated deprotection and coupling of the purified free amine to an keto/azido acid activated as the succinimidyl ester. Finally, analogue **8** was synthesized in one step through microwave-assisted nucleophilic ring-opening of pantolactone by monoazido/monoamino terminated nonaethylene glycol **32a**.

In Vitro Pathway Incorporation: Kinetics with CoaA. As mentioned earlier, phosphorylation of the primary hydroxyl group of pantothenate by the first protein in the CoA biosynthesis pathway, CoaA, is believed to be the rate-limiting step in vivo.[13] Due to its role as the gatekeeper of CoA biosynthesis, we assayed each of our new analogues for kinetic activity with CoaA (Table 2).

The assay was performed as previously described using the prototypical bacterial pantothenate kinase, CoaA from *E. coli*.[15] The reaction of CoaA with the natural substrate pantothenate gave values that conformed to those previously reported in the literature.[14,21] Pantothenate mimics **2** and **3** show k_{cat} and K_m values closely approaching those of the natural substrate. As seen in our previous studies pantetheine is also a substrate for CoaA, and pantetheine-resembling substrates **4**, **5**, **13**, and **14** show turnovers near equal to (**4**, **13**, **14**) or greater than (**5**) that of pantetheine. Compounds **5**, **13**, and **14** are processed particularly efficiently by CoaA. Conversely, compounds with PEG linkers between the pantoic acid moiety and the bioorthogonal terminus were poor substrates for CoaA. The length of the linker region was a strong factor in determining substrate suitability, with the longest compounds **8** and **18** proving such poor substrates that kinetic data could not be generated. Shorter PEG linked compounds (**7** and **17**) were viable substrates in the assay but turned over at a rate 2−5-fold less than that for derivatives containing β-Ala/diamine linkers. PEG linked pantoic acid conjugate **6** differs only from **7** by the exchange of an azide for an acyl substitutent but shows markedly decreased kinetic activity.

To isolate and investigate the effect of H-bond accepting heteroatoms in the linker region, pantetheine analogues **15** and

(21) Strauss, E.; Begley, T. P. *J. Biol. Chem.* **2002**, *277*, 48205−48209.

Table 2. Kinetic Parameters of *E. coli* CoaA with Natural Substrates and Pantetheine Analogues and Summary of in Vivo/in Vitro Results[a]

compound	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1} M^{-1}$) ($\times 10^3$)	in vitro	in vivo
pantothenate	31.27 ± 0.58	28.56 ± 1.77	18.25 ± 0.68	na	na
pantetheine[14]	19.2 ± 0.1	91 ± 10	3.53 ± 0.44	na	na
1	22.33 ± 1.54	692.3 ± 89.13	0.54 ± 0.07	++	−
2	31.45 ± 0.75	32.85 ± 2.54	15.96 ± 0.76	++	++
3	28.02 ± 0.42	43.53 ± 1.99	10.73 ± 0.32	++	++
4	20.64 ± 0.45	62.65 ± 3.89	5.49 ± 0.24	++	+
5	40.91 ± 1.12	71.89 ± 6.98	9.48 ± 0.52	++	++
6	1.60 ± 0.06	0.196 ± 0.09	136.4 ± 10.63	++	−
7	6.16 ± 0.44	53.76 ± 11.16	1.91 ± 0.27	++	−
8	na	na	na	+	−
12	11.56 ± 0.57	37.24 ± 5.82	5.17 ± 0.51	++	+
13	19.16 ± 1.41	28.40 ± 6.92	11.25 ± 1.65	++	++
14	17.26 ± 0.62	27.33 ± 3.28	10.53 ± 0.76	++	−
15	1.36 ± 0.05	0.76 ± 0.32	0.03 ± 0.04	++	−
16	1.00 ± 0.03	0.14 ± 0.18	0.12 ± 0.01	++	−
17	10.06 ± 0.43	51.18 ± 6.42	3.28 ± 0.28	++	−
18	na	na	na	+	−

[a] Kinetic data with CoaA and summary of in vivo/in vitro labeling efficiency of carrier proteins using reporter modified pantetheine analogues. Kinetic values for the natural substrate pantothenate and pantetheine are given for comparison. In vivo/in vitro labeling is based on data from gel shift, fluorescent gel, Western blot, and MALDI-MS data (see Figures 3,4,5, Supporting Information) and represented semiquantitatively as follows: ++ for excellent labeling of carrier protein, + for low but detectable labeling, and − for no detectable labeling of carrier protein.

16 were synthesized. These compounds were designed to replace the amide bond between β-Ala and cystamine of pantetheine with a single ether oxygen, allowing investigation of subtle substrate−enzyme interactions in the CoaA active site.[12] To our surprise these compounds were extremely poor substrates. While compounds with the aforementioned short PEG linkers (**7**, **17**) showed K_m values 2-fold higher than that for natural substrate pantothenate, and compounds with traditional β-Ala/diamine linkers showed values that were either the same (**13**, **14**) or 2-fold higher (**4**, **5**), the K_m values for **15** and **16** were 100-fold *lower*. Turnover for these compounds was proportionately low, indicative of tight binding. To test if these compounds were acting as inhibitors of CoaA, a competitive kinetic assay was set up using pantotenate as the substrate. The results (see Supporting Information) show that **15** acts as a noncompetitive inhibitor, suggesting that it may bind the allosteric regulation site of CoaA.[23] Investigation of these compounds as potential inhibitor scaffolds is ongoing. Azide **1**, which omitted entirely the β-Ala/carbon diamine or PEG linkages of other analogues, showed turnover within the range of the natural substrate; however the K_m was 20-fold higher than that for pantothenate suggesting deletion of an interaction involved in active site binding. Another interesting pantoic acid analogue **12**, which shortened the pantoic acid-reporter linker length to four carbons and reversed the carbonyl and amide β-Ala/cystamine linkage of natural pantetheine, showed lower turnover and catalytic efficiency than analogues containing the natural β-Ala/pantoic acid linkage (**5**, **13**, **14**).

In Vitro Pathway Incorporation: Gel Shift. The conversion of *apo*-ACP to *holo*-ACP or reporter-modified *crypto*-ACP causes a change in the mobility of the protein on a nondenaturing polyacrylamide gel.[12] To assay each compound for activity throughout the entire co-opted CoA pathway, we ran covalently modified carrier protein on a native PAGE gel and compared the mobility to *apo*-ACP (Figure 3a). Pantetheine analogues were reacted as previously reported[4] with the enzymes CoaA-E to create CoA analogues, followed by the addition of the PPTase Sfp and *apo*-ACP. As seen in Figure 3a, all of the compounds tested demonstrated some change in mobility with relation to *apo*-ACP. For most compounds full conversion to *crypto*-ACP is obtained; however analogues with long PEG linkers (**8**, **18**) give multiple bands on the gel that suggest *apo*-protein remains.

To confirm the results from this assay, we subjected the reaction mixtures to matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). The MALDI-MS data (Figure 4 and Supporting Information) confirm that all pantetheine analogues in our panel are indeed converted into CoA derivatives and transferred onto a carrier protein in vitro. *Apo*-ACP (Figure 4a) shows a characteristic peak with a mass of 8505 Da. Compound **13** was reacted with the CoA biosynthesis

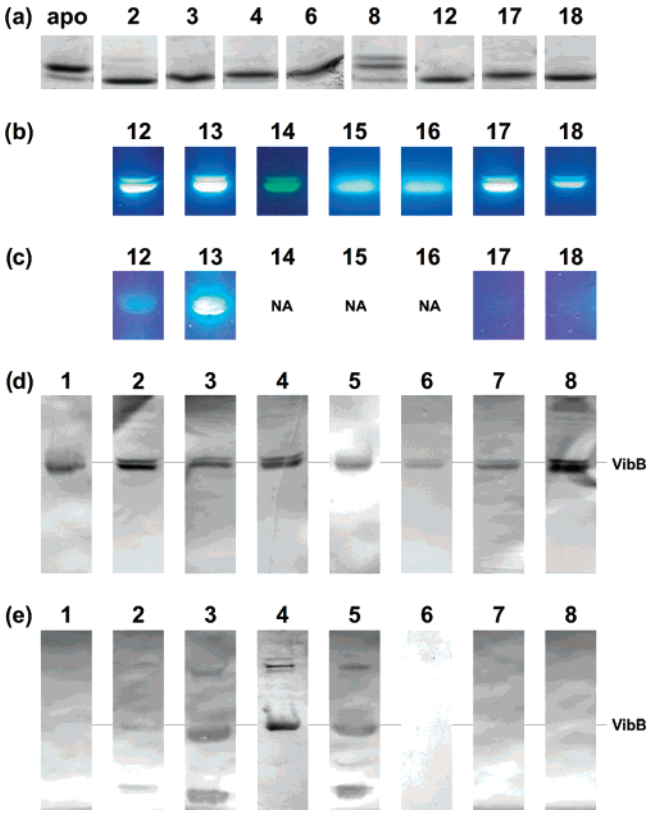


Figure 3. In vivo and in vitro activity of pantetheine analogue panel. (a) Analogues were assayed for gel shift after reaction with CoA biosynthesis enzymes (CoaA/D/E), PPTase (Sfp), and carrier protein (*E. coli* ACP). Conversion of *apo*-ACP to reporter modified *crypto*-ACP causes a change in the mobility of the protein on native PAGE. (b) In vitro modification of the carrier protein VibB by reaction with fluorescent pantetheine analogues, CoA biosynthesis enzymes (CoaA/D/E), and Sfp. (c) In vivo modification of carrier protein by incubation of fluorescent pantetheine analogues with *E. coli* overexpressing VibB and the PPTase Sfp. (d) In vitro modification of VibB by reaction with bioorthogonal pantetheine analogues, CoA biosynthesis enzymes (CoaA/D/E), and Sfp. Labeled carrier protein is visualized by chemoselective ligation to the appropriate biotin reporter (**9**−**11**) followed by SDS-PAGE, blotting onto nitrocellulose, and incubation with streptavidin-linked alkaline phosphatase. (e) In vivo modification of carrier protein by incubation of bioorthogonal pantetheine analogues with *E. coli* overexpressing VibB and the PPTase Sfp. Visualization as in (d).

(22) Mercer, A. C.; La Clair, J. J.; Burkart, M. D. *ChemBioChem* **2005**, *8*, 1335−1337.
(23) Ivey, R. A.; Zhang, Y.; Virga, K. G.; Hevener, K.; Lee, R. E.; Rock, C. O.; Jackowski, S.; Park, H. *J. Biol. Chem.* **2004**, *279*, 35622−35629.

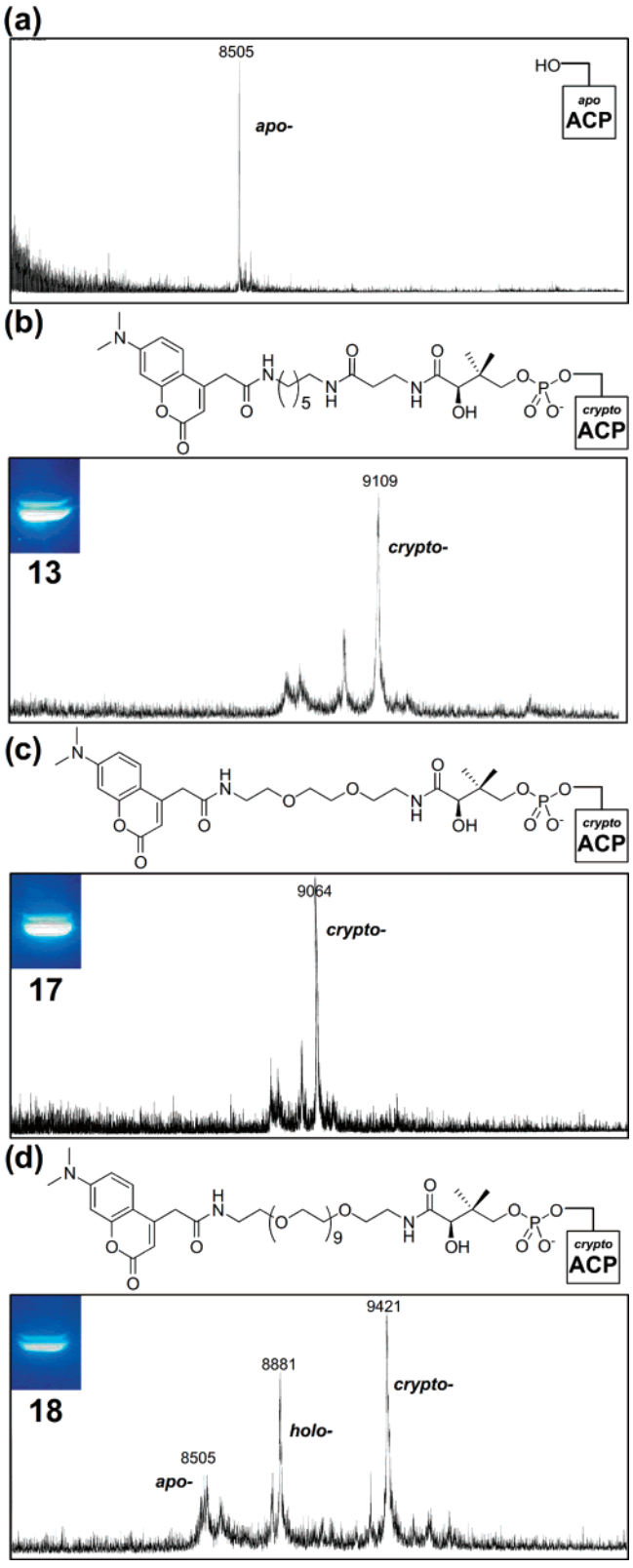


Figure 4. In vitro labeling of ACP. *Apo*-ACP (a) is reacted with pantetheine analogues, CoA biosynthetic enzymes (CoaA/D/E), and the PPTase Sfp. Fluorescent compounds **13** (b), **17** (c), and **18** (d) are all shown to be converted to CoA analogues and modify ACP by gel and mass spectral analysis.

enzymes, ACP, and Sfp as described above, and the reaction mix was analyzed by MALDI without further purification. As seen in Figure 4b, reaction with the CoA analogue of **13** causes the expected mass change of 604 mass units. PEG analogues **17** and **18** (Figure 4c,d) also show the expected mass shifts of 559 units and 916 units, respectively, corresponding to the formation of a *crypto*-carrier protein. Only in the case of the PEG-linked derivatives **8** and **18** does the conversion appear significantly incomplete, supporting the results of the gel shift assay.

In Vitro Pathway Incorporation: Biodetectability. Having confirmed that the pantetheine analogues were suitable substrates for the CoA biosynthesis enzymes and PPTase/carrier protein reaction, we next investigated the biodetectability of each analogue. The fluorescent analogues (**12**−**18**) were detected by UV visualization on PAGE gels as previously described.[4] Bioorthogonally tagged pantetheine analogues **1**−**8** were detected by chemoselective ligation to the appropriate alkoxyamine/azide/alkyne functionalized biotin followed by PAGE and visualization by western blotting and incubation with streptavidin-conjugated alkaline phosphatase. Keto-pantetheine compounds **4** and **6** were reacted overnight with biotin hydroxylamine (**9**) at room temperature, while pantetheine analogues with azide and alkyne functionalities were reacted with the corresponding alkynyl/azido (**10**/**11**) biotin following the procedure of Alexander.[9c] Inspection of the fluorescent gels (Figure 3b) and western blots (Figure 3d) confirmed the results of the gel shift assay and mass spectral data, indicating biodetectible covalent modification of carrier proteins in vitro for compounds **1**−**8** and **12**−**18**.

In Vivo Pathway Incorporation. Next we tested our library of pantetheine analogues for integration into the *E. coli* CoA pathway using our in vivo assay.[4] *E. coli* overexpressing the carrier protein VibB and the PPTase Sfp were incubated with 1 mM of each compound in 1 mL of culture. After 4 h of growth cells were pelleted, washed, and lysed. Lysate from these cultures was run on SDS-PAGE gels, and detection was carried out as described for the in vitro studies. Compounds **2**, **3**, **5**, **12**, and **13** demonstrated detectible modification of VibB in vivo (Figure 3c,e). Keto-pantetheine analogue **4** was also detectable but showed much weaker labeling than the similarly linked azido-analogue **5** (Figure 3e). The compounds most active in vivo show a strong correlation with the CoaA kinetic profile. To verify the results of the gels and blots, samples of crude lysate were assayed by MALDI-MS (Figure 5). For MALDI-MS analysis doubly transformed *E. coli* containing plasmids for the carrier protein Fren and the PPTase Sfp were used. As seen in Figure 5a compound **13** was taken up by the cell, processed into a CoA analogue and attached onto Fren. The *apo* peak can be seen at 8570 mass units. *Holo*-carrier protein is also visible at 8910 Da. This peak arises from the fact that natural CoA is available in the cell and readily ligated by PPTases to the overexpressed *apo*-Fren. Carrier protein modified with **13** can be seen at 9179 Da giving the expected 609 mass unit change. Similarly mass spectral analysis of cell lysate after incubation of bioorthogonal pantetheine analogue **2** with Fren/Sfp overexpressing *E. coli* shows an observable 315 Da shift of the known *apo* peak indicating formation of an alkyne-modified *crypto*-carrier protein. Subjection of the same crude cell lysate to click reaction conditions with biotin reporter **11** resulted in another 317 Da mass shift, indicating successful formation of biotinylated Fren via a Cu(I)-catalyzed [3 + 2] cycloaddition process.

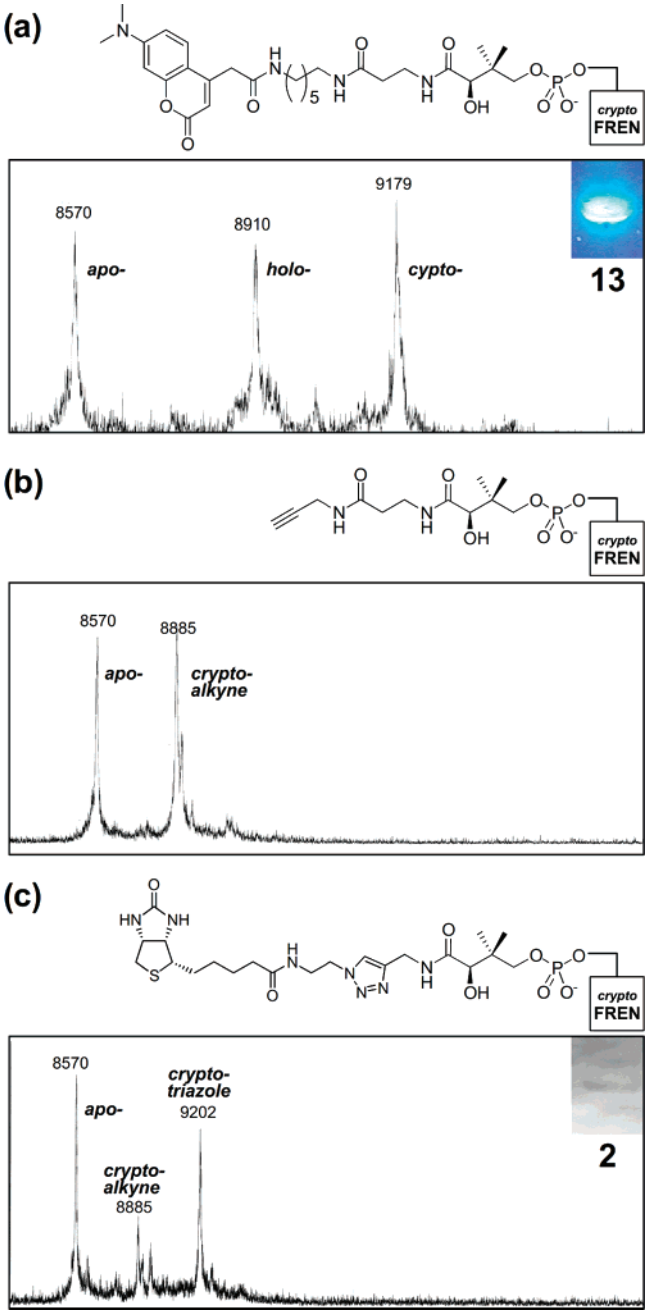


Figure 5. In vivo carrier protein labeling. The carrier protein Fren is labeled in vivo by incubation of *E. coli* overexpressing Fren and Sfp with (a) a fluorescent pantothenate analogue **13** and (b) a bioorthogonally tagged analogue **2**. Click reaction of alkyne-modified *crypto*-carrier protein with biotin reporter **11** affords triazole-linked biotinylated carrier protein (c), resulting in the expected shift in mass and allowing protein visualization by western blot.

Insights into CoaA Substrate Specificity from Kinetic, in Vivo, and in Vitro Analyses. Comparisons of the kinetic, in vitro, and in vivo assay results for different members of the panel yield important insights into the structure−activity relationships between *E. coli* PANK and pantetheine analogues. Despite the poor performance of compounds **1**, **6**−**8**, **12**, and **15**−**18** in the CoaA kinetic assay, all were shown to be converted to CoA analogues and loaded onto the carrier protein ACP by Sfp in vitro. This is both a testament to the incredible efficiency of Sfp in transferring unnatural CoA derivatives to *apo*-carrier proteins and a further demonstration of the utility of the chemoenzymatic approach to synthesis of unnatural CoA analogues. As mentioned above this finding greatly simplifies the synthetic task of constructing CoA derivatives for in vitro applications which utilize carrier protein tagging.[6] This allows access to virtually any reporter labeled-CoA analogue from a monoprotected amine in three steps, one of which can be expedited using microwave technology. Additionally the subtle substrate preferences Sfp has been shown to exhibit in systems incorporating multiple carrier proteins can be used for selective coding based on differential phosphopantetheinylation by unnatural CoA analogues,[22] adding another layer of complexity to in vitro and cell-surface carrier protein fusion systems.

Detailed analyses of the kinetic data demonstrate several important relationships. In general compounds containing a similar β-Ala linker region to that of natural pantothenate show the best kinetic profiles. In all compounds of the panel the pantoic acid region (C1−N5, numbering from the terminal primary hydroxyl of pantetheine) was conserved. Lee et al. indicated low binding of PANK inhibitors which lacked a proton-donating amide NH at the N8 position, pointing to a model of the PANK−ADP−pantothenate ternary complex in which the C-7 acid of pantothenate acts as an H-bond donor toward two key residues.[12,23] Our results verify the importance of this interaction. Compounds **1** and **3** contain the same 2-azidoethanamine-derived terminal azide, differing only in the β-Ala linkage of **3**. This change results in a 10-fold increase in K_m and near 20-fold increase in catalytic efficiency, indicating much better binding of the substrate with the H-bond donating β-Ala linker. Pantetheine analogues containing ethylene glycol based linkers incapable of acting as efficient H-bond donors (**6**−**8**, **17**−**18**) showed similarly poor kinetics compared with those containing β-Ala linkers. Compounds **15** and **16** provide perhaps the strongest evidence of the importance of this interaction, losing almost all substrate activity when reducing the strength of the electron pair donor and removing the H-bond donating NH completely from the pantetheine analogue. Interestingly these compounds show markedly different kinetics than PEG-linked pantetheine **17**, which differs by only two atoms, demonstrating the limitations of this model in predicting effects caused by alternate variables such as the reduced rotation around an sp^2 hybridized carbon at C8 and addition of an extra H-bond accepting heteroatom further down the linker. The kinetics of analogue **12**, in which the connectivity of the amide at C8 and N9 of the natural substrate pantetheine is reversed, indicate that there is some flexibility in the pocket around this position. Compound **12** shows good binding but slow turnover, with a catalytic efficiency poorer than that of any of the compounds with the H-bond donator in its natural position (**2**−**5**, **13**−**14**) but better than that of every compound in which the H-bond donating NH is absent (**1**, **6**−**8**, **15**−**18**). While the general trend toward an H-bond donor effect is large, other interactions resulting from the proximity of the aromatic coumarin-reporter molecule to the active site in this analogue must also be considered.

Several other trends which may be important for future design of carrier protein tags are of interest. In general, analogues terminating in alkynes show better kinetic parameters than those of azides, with azides showing better kinetic parameters than those of ketones. For PEG-based pantetheine derivatives, chain length proved an important factor, with longer chains showing

negligible activity with CoaA and poor in vivo protein tagging. Appending a different dye to the end of the pantetheine had no statistically significant impact on CoaA kinetics; however substitution of DMACA with dansyl lead to a complete loss of in vivo activity, suggesting a possible lack of a viable membrane-transport mechanism for dansyl-pantetheine analogues or an intracellular degradation process. It has been shown previously that DMACA is an excellent dye for in vivo applications.[20] The kinetics of β-Ala containing bioorthogonal pantetheines (**2**−**5**) compared with β-Ala containing pantetheines in which a fluorescent reporter that was directly appended (**13**−**14**) showed slightly better turnover and similar catalytic efficiency. On the whole the binding site of CoaA beyond the β-Ala moiety appears to be quite promiscuous, with little kinetic effect observed on substitution of the hexanediamine linker of **13** with a ethylenediamine linker or substitution of the ethylenediamine linker of **14** with a short (8-atom) PEG linker.[24]

Perhaps the most important conclusion that can be drawn from the assay results is that the in vivo activity of a pantetheine analogue has a direct correlation with kinetic activity with CoaA. Analogues **2**, **3**, **4**, **5**, **12**, and **13** were shown to be biodetectible in *E. coli* by Western blot analysis and fluorescence visualization. These show a k_{cat} between 11.5 and 40.9 min^{-1} and k_{cat}/K_m of 5.2×10^3−16.0×10^3 compared with the values of 19.2 min^{-1} (k_{cat}) and 3.5×10^3 s^{-1} M^{-1} (k_{cat}/K_m) for unmodified pantetheine (Table 2). In order for a pantetheine analogue to be processed into a CoA analogue in vivo, not only must an organism contain a promiscuous PANK with pantetheine kinase activity but also the pantetheine analogue must have comparable or better kinetics with PANK than pantetheine and (ideally) natural substrate pantothenate. Given that enterococci produce far more pantothenate than they require for primary metabolism,[25] any modified pantetheine analogue must make extremely efficient use of CoaA in order for CoA conversion and subsequent protein labeling to occur at detectible levels. Although the structural requirements for the import of pantetheine analogues have not been studied in detail, *E. coli* has been shown to have a pantothenate import system, the pantothenate permease (panF) symporter,[26] which also may exert some selectivity in the import of analogues and thus influence the ability of pantetheine analogues to be integrated into the CoA pathway in vivo. However previous studies showed that while overexpression of the panF gene resulted in elevated pantothenate uptake, a concurrent increase in CoA production was not observed, indicating PANK activity as the principal regulator of CoA biosynthesis. With that in mind these kinetic parameters should prove useful for the future design and assay of in vivo carrier protein tags.

Conclusions

The molecules described here have varied applications and provide insight valuable in the expanding field of proteomics. For in vitro applications and in vivo cell-surface labeling of carrier protein fusions, all pantetheine analogues studied in this manuscript are efficiently converted into CoA analogues and tethered to the protein via the four-step enzymatic pathway. This alternative methodology negates the purification steps necessary in the production of maleimide−CoA analogues from commercial sources, allows almost any variance of the chemical identity of the linker region, and provides an economical substitute for producing large amounts of CoA analogues in cases where large quantities of the desired CoA−maleimide−reporter conjugate may be prohibitively expensive. In addition the expansion of analogues with bioorthogonal reporters allows for increased detection and sensitivity. However despite these subtle advances, it is in the prospect of in vivo labeling that these tools become particularly important. The covalent modifications described herein have practical value in the study of in vivo activity of proteins and a place among the ever increasing myriad of proteomic techniques used to study them.

Of particular importance are the chemoselective ligation reactions demonstrated by the ketone, azide, and alkyne protein labels. These tools allow carrier proteins to be visualized and isolated for the first time without the expense and complication of antibody techniques. While our survey of bioorthogonal coupling partners was not exhaustive, the functionalities introduced should be applicable to other published methods. For instance, one can easily envisage analogues **3** and **5** being modified by Bertozzi's covalent Staudinger ligation with a reporter-conjugated triarylphosphine analogue for in vivo applications in which more stringent Cu(I)-catalyzed click chemistry conditions are not ideal.[27] It was to our disappointment that the PEG-incorporating pantetheine analogues proved nonamenable to in vivo protein labeling, most likely due to deletion of a crucial H-bonding interaction. Yet while PEG analogues are often useful for distancing reporter labels from the protein of interest for downstream modification, most likely they are not a necessity in this instance by virtue of the 4′-phosphopantetheine moiety. Indeed the 4′-phosphopantetheine is commonly believed to be appended to carrier proteins as a means to distance substrates and products from the protein core, a concept reinforced by recent structural studies of the fatty acid synthase.[28]

The system used in this study was limited to the *E. coli* CoA pathway and relied on the overexpression of a carrier protein PPTase pair to facilitate detection. However it has already been demonstrated that AcpS, a widely expressed PPTase, has the ability to transfer CoA analogues chemoenzymatically generated from panthothenamide type substrates to apo-ACP.[12] Immediate work will aim at lowering the detection limit to native protein levels in *E. coli*, a reasonable goal given the demonstrated ability of bioorthogonally labeled carrier protein to undergo ligation to affinity agents readily amenable to enrichment strategies. Also of interest to the growth of this methodology is the promiscuity of CoA biosynthetic enzymes in species other than *E. coli*. Little is known about the interspecies substrate specificity of CoaD and CoaE, although studies of the mechanism of inhibition of N-alkylpantothenamides have shown these enzymes capable of catalyzing the formation of unnatural CoA analogues which modify the fatty acid ACP in both *E. coli* and *S. aureus*,[21,29] despite a disparity in sequence homology between the PANK

(24) This comparison refers to kinetic data compiled for alternatively linked dansyl and DMACA−pantetheine analogues in ref 14.
(25) Rock, C. O.; Calder, R. B.; Karim, M. A.; Jackowski, S. *J. Biol. Chem.* **2000**, *275*, 1377−1383.
(26) Jackowski, S.; Alix, J. H. *J. Bacteriol.* **1990**, *172*, 3842−3848.
(27) Kohn M, Breinbauer R. *Angew. Chem., Int. Ed.* **2004**, *43*, 3106−3116.
(28) (a) Simon Jenni, S.; Leibundgut, M.; Maier, T.; Ban, N. *Science* **2006**, *311*, 1258−1262. (b) Maier, T.; Jenni, S.; Ban, N. *Science* **2006**, *311*, 1263−1267.
(29) Leonardi, R.; Chohnan, S.; Zhang, Y.; Virga, K. G.; Lee, R. E.; Rock, C. O.; Jackowski, S. *J. Biol. Chem.* **2005**, *280*, 3314−3322.

enzymes of the different species.[30] Studies in our own lab have shown the mammalian form of CoaD and CoaE, which exists as a bifunctional fusion, efficiently processes 4′-phosphopantetheine analogues into CoA analogues in vitro.[14] One system where this approach may be limited is in a subset of pathogenic bacteria which contain only the recently discovered CoaX type pantothenate kinase. This PANK isoform has been shown to be resistant to inhibition by N-alkylpantothenamides, possibly indicating an inability of pantetheine analogues to act as competitive substrates.[31] To date most studies of pantetheine analogues have focused on their properties of inhibitors; the new amenability of carrier protein systems to fusion tagging strategies and our demonstration of protein modification in living systems provide strong incentive for the reexamination of the substrate properties of this class of small molecules.

Here we have performed a detailed investigation of pantetheine analogues to identify suitable partners for covalent protein labeling inside living cells. A rapid synthesis of pantothenamide analogues was developed for this purpose and used to produce a panel which was evaluated for in vitro and in vivo protein labeling. Kinetic comparisons allowed the construction of structure−activity relationships to pinpoint the linker, dye, and bioorthogonal reporter of choice for protein labeling. Finally bioorthogonal pantetheine analogues were shown to target carrier proteins with high specificity in vivo and undergo chemoselective ligation to reporters in crude cell lysate. The paucity of site-specific protein labeling tools represents a major obstacle to the routine application of such small-molecule probes in vivo.[32] Our increasing understanding of the kinetic parameters and structural limitations of pantetheine analogues sets the stage for the future use of 4′-phosphopantetheine analogue labeling in chemical biology.

Acknowledgment. This manuscript is dedicated to Bill Fenical on the occasion of his 65th birthday. Funding was provided by the University of California, San Diego, Department of Chemistry and Biochemistry, NSF CAREER Award 0347681 and NIH RO1GM075797. J.M. was supported by NIH Training Grant T32DK007233. H.R. was supported as an NIH NIGMS PREP scholar. We thank Joseph Noel and Thomas Baiga of the Salk Institute for Biological Studies for the use of a microwave reactor and Jessica Alexander of the Scripps Research Institute for helpful discussions.

Supporting Information Available: Abbreviations used are given in ref 33. General procedures for microwave assisted ring-opening of pantolactone; experimental details for the synthesis of compounds **1**−**18**; kinetic, in vitro, and in vivo assay details; full author listings for refs 18 and 20; and ^{1}H and ^{13}C NMR spectra of all final compounds and synthetic intermediates. This material is available free of charge via the Internet at http://pubs.acs.org.

JA063217N

(30) Choudhry, A. E.; Mandichak, T. L.; Broskey, J. P.; Egolf, R. W.; Kinsland, C.; Begley, T. P.; Seefeld, M. A.; Ku, T. W.; Brown, J. R.; Zalacain, M.; Ratnam, K. *Antimicrob. Agents Chemother.* **2003**, *47*, 2051−2055.
(31) Brand, L. A.; Strauss, E. *J. Biol. Chem.* **2005**, *280*, 20185−20188.
(32) Chen, I.; Ting, A. Y. *Curr. Opin. Biotechnol.* **2005**, *16*, 35−40.
(33) Abbreviations: CoA, Coenzyme A; PPTase, phosphopantetheinyltransferase; *O*-GlcNAc, *O*-linked *β*-*N*-acetylglucosamine; *β*-Ala, *β*-alanine; PANK, pantothenate kinase; CoaA, ecoli PANK; CoaD, *E. coli* phosphopantetheine adenylytransferase; CoaE, *E. coli* dephospho-CoA kinase PEG, poly(ethylene glycol); Pantolactone, (D)-(−)-pantolactone; Boc, *tert*-butyl carbamate; Alloc, allyl-carbamate; DMACA, 7-dimethylaminocoumarin-4-acetic acid; CP, carrier protein; ACP, fatty acid synthase acyl carrier protein from *E. coli*; VibB, vibriobactin synthase carrier protein from *Vibrio cholerae*; Fren, frenolicin synthase carrier protein from *Streptomyces roseofulvus*; panF, pantothenate permease.