Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase


Departments of *Biological Chemistry and Molecular Pharmacology and †Microbiology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115; and ‡Department of Chemistry, University of Illinois, 600 South Mathews Avenue, Urbana, IL 61801

Contributed by Christopher T. Walsh, September 2, 2005

An 11-residue peptide with the sequence DSLEFIASKLA was identified from a genomic library of Bacillus subtilis by phage display as an efficient substrate for Sfp phosphopantetheinyl transferase-catalyzed protein labeling by small molecule–CoA conjugates. We name this peptide the “ybbR tag,” because part of its sequence is derived from the ybbR ORF in the B. subtilis genome. The site of Sfp-catalyzed ybbR tag labeling was mapped to the underlined Ser residue, and the ybbR tag was found to have a strong tendency for adopting an α-helical conformation in solution. Here we demonstrate that the ybbR tag can be fused to the N or C termini of target proteins or inserted in a flexible loop in the middle of a target protein for site-specific protein labeling by Sfp. The short size of the ybbR tag and its compatibility with various target proteins, the broad substrate specificity of Sfp for labeling the ybbR tag with small-molecule probes of diverse structures, and the high specificity and efficiency of the labeling reaction make Sfp-catalyzed ybbR tag labeling an attractive tool for expanding protein structural and functional diversities by posttranslational modification.

Abbreviations: CD, circular dichroism; EGFP, enhanced GFP; AGT, O6-alkylguanine–DNA alkyltransferase; PCPs, peptide carrier proteins; NRPS, nonribosomal peptide synthetase; ACP, acyl carrier protein; PKS, polyketide synthase; Ppant, 4-phosphopantetheinyl; CoA, coenzyme A; PCP, peptide carrier protein; FRET, fluorescence resonance energy transfer; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; Ppant, 4-phosphopantetheinyl.

*To whom correspondence should be addressed. E-mail: christopher.walsh@hms.harvard.edu.

© 2005 by The National Academy of Sciences of the USA
Fig. 1. Sfp-catalyzed PCP or ybbR tag modification at a specific Ser residue by various small-molecule–CoA conjugates (CoA-SR): 1, biotin–CoA; 2, fluorescein–CoA; 3, tetramethylrhodamine–CoA; and 4, Texas red–CoA.

display, and the N-terminal half of the tag was derived from an ORF of unknown function named ybbR in the B. subtilis genome. We have shown that the ybbR tag can be fused to either the N or C terminus or inserted within a flexible loop region of the target protein, and the ybbR-tagged proteins can be labeled by Sfp with small-molecule probes of diverse structures with high efficiency and specificity.

Materials and Methods
Detailed experimental procedures for the synthesis of CoA conjugates 1–4, cloning, Fourier-transform MS, circular dichroism (CD) measurements, and protein labeling are in supporting information, which is published on the PNAS web site.

Peptide-Labeling Kinetics. To test whether a specific peptide was the substrate of Sfp-catalyzed biotin–CoA modification, 200 μM biotin–CoA and 1 μM Sfp were incubated with 100 μM peptide in a 100-μl solution of 10 mM MgCl₂/50 mM Hepes, pH 7.5, for 30 min at 37°C. Control reactions were also run in parallel with either Sfp or biotin–CoA excluded from the reaction. Reactions were then quenched by adding 30 μl of 4% trifluoroacetic acid (TFA) and analyzed by analytical HPLC with a reverse-phase C₁₈ column by using a gradient of 0–60% CH₃CN in 0.1% TFA over 30 min and monitored at 220 nm. Peptide-labeling reactions were also carried out at various pHs ranging from 5.0 to 8.5 with various buffering reagents (sodium acetate, 50 mM, pH 5.0; Mes, 50 mM, pH 6.0; Hepes, 50 mM, pH 7.0; Heps, 50 mM, pH 8.0; TrisHCl, 50 mM, pH 8.5) to test the effect of pH on the rate of Sfp-catalyzed peptide labeling. For the determination of kinetic parameters for Sfp-catalyzed peptide labeling at a saturating concentration of biotin–CoA or fluorescein–CoA, Sfp was added to a final concentration of 1 μM in 10 mM MgCl₂/50 mM Hepes, pH 7.5, buffer with varying concentrations of the peptide ranging from 2 to 500 μM while holding the concentrations of the biotin–CoA or fluorescein–CoA conjugate constant at 150 μM. For the determination of kinetic parameters at saturating concentration of the peptide, the peptide concentration was held at 500 μM and the biotin–CoA or fluorescein–CoA concentration was varied from 2 to 200 μM. The reaction was allowed to proceed at 37°C for 5 min and was quenched and analyzed by HPLC as described above. HPLC peak areas were integrated, and the product concentration was calculated as a percent of the total peak area. Initial velocity data were fit to the Michaelis–Menten equation by the computer software KALEIDOGRAFHY (Synergy Software, Reading, PA). The kinetic parameters for the Sfp-catalyzed PCP labeling were carried out in the same buffer (10 mM MgCl₂/50 mM Hepes, pH 7.5) in the presence of 0.1 μM Sfp by either varying the concentration of PCP from 0.5 to 50 μM at a constant biotin–CoA concentration of 150 μM or varying the biotin–CoA concentration from 2 to 200 μM at a constant PCP concentration of 20 μM. The reaction was allowed to proceed at 37°C for 5 min before quenching by the addition of 30 μl of 4% TFA to a 100-μl reaction mixture. The reaction was then analyzed by analytical HPLC with a reverse-phase C₁₈ column by using a gradient of 30–50% CH₃CN in 0.1% TFA/H₂O over 30 min and was monitored at 280 nm.

Results
Identification of Truncated ybbR Proteins as Substrates of Sfp. We previously showed that PCP displayed on the surface of M13 phages can be specifically labeled with biotin by Sfp-catalyzed biotin–CoA transfer (15). As part of our effort to identify proteins subjected to Sfp-catalyzed posttranslational modification in the B. subtilis proteome, a genomic library of B. subtilis was displayed on the surface of M13 phages as pIII fusion proteins and selected for Sfp-catalyzed biotin–Ppant modification. Phages displaying proteins recognized by Sfp for posttranslational modification were covalently labeled with biotin–Ppant and selected by binding to immobilized streptavidin. The proteins subjected to Sfp-catalyzed posttranslational modification were then identified by sequencing the selected phage clones. The detailed phage selection results will be published elsewhere.

Besides known ACPs and PCPs, we found that truncated forms of the predicted ybbR protein (23) (residues 1–484 for full length) corresponding to residues 95–278 (JY565), 111–278 (JY503), 214–278 (JY530), and 229–278 (JY529) were selected multiple times (Fig. 2), and Sfp-catalyzed biotin labeling of phage-displayed ybbR truncates was confirmed by ELISA. JY529, the shortest ybbR-
truncated protein selected by phage display, was cloned into the pET21b expression vector, and the 49-residue fragment was expressed as a fusion to a C-terminal 6× histidine tag (Y529-His). The purified protein was incubated with Sfp and biotin–CoA, and the biotin labeling of Y529-His was confirmed by Western blot analysis probed with a streptavidin–horseradish peroxidase conjugate, whereas control reactions with either no Sfp or no biotin–CoA added showed no biotin labeling on the Western blot (Fig. 3). This suggests that truncated ybbR proteins were selected by phage display due to their posttranslational modification by Sfp-catalyzed biotin–CoA labeling. The site of Ppant modification in Y529-His was mapped by Fourier-transform MS (FTMS) to Ser-274 (underlined), as shown by the HPLC trace 1 with a product peak at 21 min (Fig. 5A). Peptides ybbR11–13, failed to be the substrate of Sfp (data not shown), suggesting that the full length ybbR protein in the B. subtilis proteome may not be modified by Sfp.

The activities of Sfp with peptide ybbR13 and biotin–CoA as substrates were determined at various pHs ranging from 5.0 to 8.5, and it was found that Sfp showed the highest activity between pH 7.0 and 8.0, thus all subsequent kinetic measurements were done at pH 7.5. Fig. 5B shows a typical Michaelis–Menten plot for the Sfp-catalyzed ybbR peptide modification reaction. Detailed kinetic analysis was carried out for the Sfp-catalyzed modification of peptides ybbR11–13 by biotin–CoA and fluorescein–CoA (Table 1). Peptides ybbR11–13 have the same nine residues (LEFIASKLA) C-terminal to Ppant-modified Ser-274 and the same number of residues beyond Ser-274 as in Sfp active peptides ybbR11–13, failed to be the substrate of Sfp (data not shown), suggesting that the full length ybbR protein in the B. subtilis proteome may not be modified by Sfp.

Characterization of Short ybbR Peptides as Substrate of Sfp-Catalyzed Ppant Modification. Y529-His showed no revealing sequence homology with ACPs or PCPs, the known substrates of Sfp. The site of Ppant modification in Y529-His, Ser-274, was very close to the C terminus of the protein, in contrast to Ppant-modified Ser in ACPs or PCPs, which is in the middle of the 80- to 90-residue protein with 40–45 residues to either end of the protein. We thus tested whether the C-terminal peptide of Y529-His encompassing the Ser-274 residue could be recognized by Sfp as a substrate. Short peptides corresponding to the flanking sequence of Ser-274 in Y529-His were synthesized (Fig. 4A) and incubated with Sfp and biotin–CoA, followed by HPLC analysis. Upon incubation with Sfp and biotin–CoA for 30 min at 37°C, >75% of peptide ybbR13 (DSLEFIASKLA) was found to be modified by biotin–CoA, as shown by the HPLC trace 1 with a product peak at 21 min (Fig. 5A). The product formation depends on the presence of both biotin–CoA and Sfp, because no product was formed with either biotin or no biotin–CoA for 30 min at 37°C, suggesting that the full length ybbR protein in the B. subtilis proteome may not be modified by Sfp.
may be due to the larger surface area of PCP that can interact with 19.1 min/H9262 solution of 100 a biotin CoA. (trace 2) or biotin–CoA (trace 3) excluded from the reaction mixture. A 100- a biotin CoA. ( trace 2) or biotin–CoA (trace 3) excluded from the reaction mixture. A 100- min/H9262 biotin CoA. (trace 2) or biotin–CoA (trace 3) excluded from the reaction mixture. A 100- min/H9262 biotin CoA. (trace 2) or biotin–CoA (trace 3) excluded from the reaction mixture. A 100-

Fig. 5. Kinetic analysis of Sfp-catalyzed peptide ybbR13 modification by biotin CoA. (a) HPLC traces of the ybbR13 peptide-labeling reaction with biotin CoA and Sfp added (trace 1) and the control reactions with either Sfp (trace 2) or biotin–CoA (trace 3) excluded from the reaction mixture. A 100-μM solution of 100 μM peptide in 10 mM MgCl2/50 mM Hepes, pH 7.5, was incubated with 200 μM biotin–CoA/1 μM Sfp at 37°C for 30 min before HPLC analysis. (b) Michaelis–Menten plot for the measurement of kinetic parameters of Sfp-catalyzed ybbR13 labeling at a saturating concentration of biotin–CoA (150 μM).

tides ybbR11–13 have a 30-fold higher \( K_m \) and a similar \( k_{cat} \), which may be due to the larger surface area of PCP that can interact with Sfp. Similar values of \( K_m \) and \( k_{cat} \) for biotin–CoA were found at saturating concentrations of ybbR13 peptide (500 μM) and PCP (20 μM), respectively, suggesting that the binding of either PCP or peptide as substrates to Sfp did not affect the binding and turnover of biotin–CoA. YbbR13 peptide was modified by fluorescein–CoA at a similar rate as biotin–CoA with a \( K_m \) of 69.9 μM and a \( k_{cat} \) of 19.1 min\(^{-1}\) (Table 1), suggesting that Sfp retains its substrate promiscuity for the modification of the short-peptide substrates instead of PCP and does not differentiate various small-molecule probes conjugated to CoA, which would be a very desirable property for site-specific protein labeling using the ybbR tags.

CD Characterization of the ybbR Peptides. When Ser-274 in ybbR11–13 was found to be the key element for Sfp recognition and could be mapped to helix II in the known structures of PCP and ACP, we tested whether the ybbR peptides could form an α-helix in solution to be recognized by Sfp as substrates. The ybbR peptides were dissolved in 5 mM potassium phosphate buffer, pH 7.5, with 30% 2,2,2-trifluoroethanol (TFE), and the peptide conformation was measured by CD spectroscopy (see supporting information). The CD spectra of Sfp active peptides ybbR11, ybbR12, and ybbR13 all showed two minima at 208 and 222 nm and an isodichroic point close to 200 nm, which are characteristics of an α-helical conformation (29). The α-helix content of these peptides in aqueous TFE were estimated to be 57.3%, 36.0%, and 35.1% for ybbR11, ybbR12, and ybbR13, respectively (30). In contrast, the Sfp inactive peptides ybbR3, ybbR8, ybbR14, and ybbR15 exhibited a lesser extent of α-helical conformation (16.6%, 11.9%, 16.6%, and 7.7%, respectively), suggesting C-terminal residues ASKLA in ybbR11–13 were important for the propensity of these peptides to adopt α-helical conformation in solution.

Because peptides ybbR11–13 all had a strong tendency for α-helix formation in solution and were good substrates of Sfp, we tested whether short peptides with sequences flanking the Ppant-modified Ser and encompassing helix II in known PCPs and ACPs could be the substrates of Sfp as well. Based on the sequences of TycC3-ACP (24) and PksL-ACP (23), two peptides, TycC3 (GGH-ŠLKAMAVAAQVHREY) and PksL1 (GLNSSGILLEVETDISKI), were synthesized with the Ppant-modified Ser at the end of a long flexible loop, which would be a very desirable property for site-specific protein labeling using the ybbR tags.

Table 1. Kinetic parameters of Sfp-catalyzed ybbR tag and PCP modification

<table>
<thead>
<tr>
<th>Peptide sequence</th>
<th>Biotin-CoA with 500 μM peptide or 20 μM PCP</th>
<th>Peptide or PCP with 150 μM biotin CoA</th>
<th>Fluorescein-CoA with 500 μM peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( k_{cat} ), ( K_m ), ( k_{cat}/K_m )</td>
<td>( k_{cat} ), ( K_m ), ( k_{cat}/K_m )</td>
<td>( k_{cat} ), ( K_m ), ( k_{cat}/K_m )</td>
</tr>
<tr>
<td>ybbR11</td>
<td>12.1 142 0.085</td>
<td>9.3 128 0.073</td>
<td>19.1 69.9 0.273</td>
</tr>
<tr>
<td>ybbR12</td>
<td>14.7 60.8 0.242</td>
<td>11.2 123 0.091</td>
<td></td>
</tr>
<tr>
<td>ybbR13</td>
<td>10.3 32.4 0.318</td>
<td>10.3 4.1 2.51</td>
<td></td>
</tr>
<tr>
<td>GrsA-PCP</td>
<td>0.267 10.8 0.025</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Yin et al. / H9251

N-terminal residues of the target proteins, which might facilitate was at the N terminus, the presence of residues ASKLA in the ybbR was at the C terminus of the protein. In contrast, when the ybbR tag conformation and subsequent Sfp recognition when the ybbR tag (DSLEFIASKLA) were crucial for the formation of with biotin by Sfp-catalyzed biotin–CoA transfer (Fig. 6). This termini of the target proteins, all of the fusion proteins were labeled when the same peptide sequence was fused to the N EGFP, GST, or maltose-binding protein, none of the fusion protein (Fig. 6). When a six-residue peptide DSLEFI was fused to the C termini of could still be labeled by Sfp-catalyzed biotin–CoA transfer (Fig. 6). Fusions on the ybbR peptide (ybbR11 and ybbR12) did not affect the terminus, which is conserved in PCPs. In fact, N-terminal exten-sions on the ybbR peptide (ybbR13) with part of the sequence displayed library of the B. subtilis genome, a short 11-residue peptide tag ybbR13 (DSLEFIASKLA) with part of the sequence derived from the B. subtilis ORF ybbR, to be an efficient substrate of Sfp. The ybbR tag has the conserved DSL tripeptide sequence at the N terminus (Fig. 4a), matching the conserved sequence motif of a PCP, although it is missing the DXFxxLGG sequence at its N terminus, which is conserved in PCPs. In fact, N-terminal exten-sions on the ybbR peptide (ybbR11 and ybbR12) did not affect the activity of the ybbR peptide as a substrate of Sfp, suggesting the N terminus of the ybbR tag plays a minor role for interaction with Sfp. In contrast, when ybbR12 was truncated by three or five residues from the C terminus (ybbR14 and ybbR8), both peptides failed to be a substrate of Sfp, revealing the importance of the C-terminal sequence of the ybbR tag for Sfp recognition.

As shown by Western blot probed with streptavidin–horseradish peroxidase, and the control reactions with either biotin–CoA or Sfp excluded did not give any biotin labeling (Fig. 6) nor did the control reactions in which the target protein without the ybbR tag was incubated with biotin–CoA and Sfp (data not shown), suggesting the labeling reaction strictly depends on Sfp-catalyzed biotin Ppant transfer onto the ybbR tag.

Shorter ybbR tags were fused to the target proteins, and it was found when 11-residue peptide tag ybbR13 (DSLEFIASKLA) was fused to either the N or C terminus of EGFP, the fusion protein could still be labeled by Sfp-catalyzed biotin–CoA transfer (Fig. 6). When a six-residue peptide DSLEFI was fused to the C termini of EGFP, GST, or maltose-binding protein, none of the fusion proteins were labeled with biotin in the presence of Sfp and biotin-CoA. In contrast, when the same peptide sequence was fused to the N termini of the target proteins, all of the fusion proteins were labeled with biotin by Sfp-catalyzed biotin–CoA transfer (Fig. 6). This suggests that residues ASKLA at the C terminus of ybbR13 (DSLEFIASKLA) were crucial for the formation of α-helical conformation and subsequent Sfp recognition when the ybbR tag was at the C terminus of the protein. In contrast, when the ybbR tag was at the N terminus, the presence of residues ASKLA in the ybbR tag might not be necessary, because the tag was followed by the N-terminal residues of the target proteins, which might facilitate α-helical formation of the N-terminal ybbR tag to be recognized by Sfp.

Because the truncated ybbR protein was selected by phage display in which Sfp-catalyzed biotin labeling of the ybbR tag VLDSLEFIASKL was flanked by the truncated ybbR protein and phage capsid protein pIII at the N and C termini, respectively, we rationalized that the ybbR tag might still function as the substrate of Sfp when the tag was inserted into a flexible loop region in the middle of the target protein. To test this idea, ybbR tag with flanking glycine residues (GGTVDLSLEFIASKLAGG, 20 residues) was inserted in place of Tyr-145 of EGFP to give the construct EGFP-ybbR145, because it was previously reported that calmodulin or zinc finger proteins inserted at the same position in EGFP did not affect protein fluorescence (31). Biotin labeling of EGFP-ybbR145 by Sfp in cell lysates was confirmed by Western blot analysis (Fig. 6), and the EGFP-ybbR was also labeled site specifically with tetramethylrhodamine and Texas red upon Sfp-catalyzed fluorophore–CoA transfer onto the internal ybbR tag in EGFP-ybbR145 to give dual-colored GFP (see supporting information).

Discussion

Sfp phosphopantetheinyl transferase activates NRPS and PKS clusters in B. subtilis by posttranslational transfer of the Ppant group from CoA to a conserved Ser residue in each of the PCPs and ACPs embedded in the biosynthetic NRPS and PKS assembly lines (12). It was previously noted that the Ppant-modified Ser (underlined) in the PCPs is located in a conserved motif with the sequence DxxDxxLGG(H/D)G(32) (‘‘x’’ denotes any of the 20 proteinogenic amino acids, and ‘‘H/D’’ denotes either H or D at the same position) (32). However, a 19-residue peptide with the sequence GVTDNFFMIGGHSLKAMMM from the B. subtilis SrfB1-PCP encompassing the conserved motif was previously tested for Sfp-catalyzed Ppant modification and was found not to be a substrate of Sfp (33). Here we report the identification from a phage-displayed library of the B. subtilis genome, a short 11-residue peptide tag ybbR13 (DSLEFIASKLA) with part of the sequence from the B. subtilis ORF ybbR, to be an efficient substrate of Sfp. The ybbR tag has the conserved DSL tripeptide sequence at the N terminus (Fig. 4a), matching the conserved sequence motif of a PCP, although it is missing the DxxDxxLGG sequence at its N terminus, which is conserved in PCPs. In fact, N-terminal exten-sions on the ybbR peptide (ybbR11 and ybbR12) did not affect the activity of the ybbR peptide as a substrate of Sfp, suggesting the N terminus of the ybbR tag plays a minor role for interaction with Sfp. In contrast, when ybbR12 was truncated by three or five residues from the C terminus (ybbR14 and ybbR8), both peptides failed to be a substrate of Sfp, revealing the importance of the C-terminal sequence of the ybbR tag for Sfp recognition.

We also found that the activity of the ybbR tag as the substrate of Sfp-catalyzed Ppant modification was coupled to its tendency to adopt α-helical conformation in solution. All Sfp active peptides ybbR11–13 showed high contents of α-helical conformation in 30% 2,2,2-trifluoroethanol, as measured by CD spectra, whereas C-terminal-truncated peptides ybbR3, ybbR8, and ybbR14 showed poor α-helical formation and were not substrates of Sfp modification. Because in PCP and ACP domains, the Ppant-modified Ser is immediately followed by helix II, which has been reported as a key element for Sfp recognition (34–36), it is plausible to propose that the Sfp active ybbR peptides adopt an α-helical conformation for the peptide sequence SLESFIASKL with the Ppant-modified Ser at the tip of the α-helix upon its binding to Sfp. Following this notion, peptides TycC3 and PksL1 encompassing the Ppant-modified Ser and the helix II region of the known PCP and ACP were tested for Sfp-catalyzed biotin–CoA modification but were found not to be substrates of Sfp. Although these peptides also had
a tendency to form α-helices, as shown by their CD spectra. They were not recognized by Sfp suggests the peptide sequence flanking the Ppant-modified Ser in PCP or ACP needs to be presented to Sfp as part of the intact PCP or ACP domain, whereas the ybbR peptide tag selected by phage displayed some unique yet unidentified features in addition to a tendency for α-helix formation that facilitates binding between the ybbR tag and Sfp.

A helical wheel representation of the ybbR tag sequence SLEFIASKLA revealed that the proposed α-helix adopted by the ybbR tag has an amphiphilic distribution of the side chains with nonpolar residues Ile, Leu, and Ala on one side of the helix and charged or polar residues Lys, Glu, and Ser on the opposite side (see supporting information). The corresponding helical wheel plots of the helix II region of the TycC3-PCP and FrenN-ACP with known NMR structures showed no significant similarities with the ybbR tag in terms of distribution or alignment of certain type of residues on a specific side of the helix. We thus expect that structural studies of ybbR peptide substrates bound to Sfp will be required to decipher the recognition rules.

The shortest Sfp active peptide ybbR13 (DSLEFIASKL) is only 11 residues long with a similar $K_m$ for Sfp-catalyzed biotin-CoA labeling as that of the full length PCP and 30-fold higher $K_m$ than that of the ybbR tag. Thirty-minute incubation of the ybbR-tagged proteins with Sfp and biotin–CoA results in $>80\%$ of the protein being labeled with biotin. Sfp retained its substrate promiscuity for the ybbR tag modification with small-molecule probes conjugated to CoA and catalyzed fluorescein–CoA transfer to the ybbR tag at a similar rate as that of biotin–CoA. The ybbR peptide tag has been shown to be recognized by Sfp when fused to the N and C termini of various proteins, and the site-specific labeling reaction of the ybbR tagged proteins by Sfp can be carried out in cell lysates within a short time (10–20 min). Interestingly, we also showed that the ybbR tag inserted in a flexible loop in the middle of EGFP could still be labeled by Sfp with various fluorophore–CoA conjugates.

The phage selection on the B. subtilis proteome turns up a peptide substrate with part of the sequence from the ybbR ORF of an unknown function. It is yet to be tested whether full length ybbR is an authentic substrate for posttranslational phosphopantetheinylation in B. subtilis cells, although the fusion of ybbR fragments to the ybbR tag can be attached to phosphopantetheine via an acidic site specifically by Sfp-catalyzed small-molecule CoA modification.

**Conclusion**

Sfp-catalyzed ybbR tag posttranslational modification has the following advantages for versatile protein labeling:

(i) The size of the ybbR tag (11 residues) is much smaller compared with the size of PCP or ACP (75–80 residues).

(ii) The ybbR tag has an excellent portability for fusion to various proteins for protein labeling: the ybbR tag can be attached to the N or C terminus of the target protein or inserted into a flexible loop in the middle of the target protein and labeled site specifically by Sfp-catalyzed small-molecule CoA modification.

(iii) Sfp has broad substrate specificities with respect to the small-molecule probes conjugated to CoA, including sugars, affinity probes such as biotin, glutathione, fluorescent probes such as fluorescein, Alexa Fluor dyes, and redox probes such as porphyrin. Furthermore, Sfp-catalyzed ybbR tag labeling by various small-molecule probes is highly specific and efficient and can be accomplished in one step.

We thus expect Sfp-catalyzed ybbR tag posttranslational modification would have broad applications for expanding the structural and functional diversity of ybbR-tagged recombinant proteins in vitro and in vivo. At present, the in vivo use is limited by the inability of CoA derivatives to cross biological membranes, but a recent report on the uptake of a fluorescent pentetheine derivative by E. coli, in vivo conversion to the fluorescent CoA, and phosphopantetheinylation transfer to the VBCP indicates a way forward for intracellular labeling of ybbR-tagged proteins (37).

We thank Stephen C. Blacklow and Cherryl Sanchez-Irizarry for helping with the CD measurements. P.D.S. was supported by a National Science Foundation Postdoctoral Fellowship in Microbiology Biology (DBI-200307). This work was supported by Grants GM20011 (to C.T.W.), GM58213 (to R.K.), and HL32854 (to D.E.G.) from the National Institutes of Health.

and $z^*$ ions generated by Electron Capture Dissociation (ECD, ions shown in red and green) localized the modification to Ser268, Ser274 and Ser280 (according to the numbering in full length ybbR). Further dissociation with Infrared Multiphoton Dissociation (IRMPD) and Collisionally Activated Dissociation (CAD) produced a series of $y$-ions (shown in blue and green), which demonstrated that Ser280 was not covalently modified (Fig. 7f). However, another series of $y$-ions portrayed +80 Da mass shifts (shown in purple), consistent with the ejection of the pantetheinate and retention of phosphate (Fig. 7d and 7e). These ions indicate that the covalently modified serine is Ser274.

Supporting figure legends
Fig. 8. CD spectra of the peptides in 5 mM potassium phosphate, pH 7.5 with 30% TFE.

(a) ybbR11, ybbR12 and ybbR13 that were the substrates of Sfp. (b) ybbR3, ybbR8, ybbR14 and ybbR15 that were not the substrates of Sfp. (c) Peptides PksL1 and TycC3 with sequences that flank the Ppant modified Ser residue in PksL-PCP and TycC-PCP, respectively, were not the substrates of Sfp.
**Fig. 9.** (a) Inserting the ybbR tag in the middle of EGFP (PDB ID 1EMA) in the position of Tyr145 to make the construct EGFP-ybbR145. The Ppant modified Ser residue in the ybbR tag is underlined. (b) UV spectra of EGFP-ybbR145 labeled with tetramethylrhodamine (blue line) or Texas Red (orange line) by Sfp catalyzed ybbR labeling.
**Fig. 10.** Helical wheel plots of ybbR peptide and the helix II region of TycC3-PCP and FrenN-ACP starting with the Ppant modified Ser residue (underlined). Solvent exposed side of helix II in TycC3-PCP and FrenN-ACP was shaded in grey. Residues with nonpolar side chains were colored in yellow, uncharged polar side chains in green, negatively charged side chains in blue and positively charged side chains in red.

**Supporting methods**

**Chemical synthesis.** Synthesis of biotin-CoA (1) was previously reported (1). Fluorescein - CoA (2), tetramethylrhodamine – CoA (3) and Texas Red – CoA (4) (Fig. 1) were synthesized following a similar procedure. Briefly, to separate solutions of coenzyme A trilithium salt (Sigma) (16.5 mg, 22 µmol) in 2 mL sodium phosphate 100 mM, pH 7.0, was added fluorescein-5-maleimide (Molecular Probes, F150) (8.5 mg, 20 µmol) in 0.5 mL DMSO, or tetramethylrhodamine-5-maleimide (Molecular Probes, T6027) (5 mg, 10 µmol) in 0.5 mL DMSO, or Texas Red C₂ maleimide (Molecular Probes, T6008) (5 mg, 6.9 µmol) in 0.5 mL DMSO, respectively. The reaction mixture was stirred at room temperature for one hour in dark followed by purification with preparative HPLC on a reverse-phase C18 column with a gradient of 0-50% acetonitrile in 0.1% TFA/water over 30 minutes. The purified compound was lyophilized and the