Efficient Enzymatic Incorporation of Dehydroalanine Based on SAMDI-Assisted Identification of Optimized Tags for OspF/SpvC

Anming Yang, Huanyu Tao, Lindsey C. Szymbczak, Liang Lin, Junfeng Song, Yi Wang, Silei Bai, Justin Modica, Sheng-You Huang,* Milan Mrksich,* and Xinxin Feng*

ABSTRACT: Site-specific modification of proteins has important applications in biological research and drug development. Reactive tags such as azide, alkyne, and tetrazine have been used extensively to achieve the abovementioned goal. However, bulky side-chain “ligation scars” are often left after the labeling and may hinder the biological application of such engineered protein products. Conjugation chemistry via dehydroalanine (Dha) may provide an opportunity for “traceless” ligation because the activated alkene moiety on Dha can then serve as an electrophile to react with radicalophile, thiol/amine nucleophile, and reactive phosphine probe to introduce a minimal linker in the protein post-translational modifications. In this report, we present a mild and highly efficient enzymatic approach to incorporate Dha with phosphothreonine/serine lyases, OspF and SpvC. These lyases originally catalyze an irreversible elimination reaction that converts a doubly phosphorylated substrate with phosphothreonine (pT) or phosphoserine (pS) to dehydrobutyryl (Dhb) or Dha. To generate a simple monophosphorylated tag for these lyases, we conducted a systematic approach to profile the substrate specificity of OspF and SpvC using peptide arrays and self-assembled monolayers for matrix-assisted laser desorption/ionization mass spectrometry. The optimized tag, [F/Y/W]-pT/pS-[F/Y/W] (where [F/Y/W] indicates an aromatic residue), results in a ∼10-fold enhancement of the overall peptide labeling efficiency via Dha chemistry and enables the first demonstration of protein labeling as well as live cell labeling with a minimal ligation linker via enzyme-mediated incorporation of Dha.

INTRODUCTION

Proteins represent the majority of nature’s worker biomacromolecules. Post-expression site-specific modification of proteins diversify the structure of proteins, enabling their applications as biological tools as well as therapeutic molecules.1,2 Modifications such as PPant attachment, lipoylation, and biotinylation are widely used; but these tags do not harbor additional chemical reactivity for next-step chemical modification. Instead, a “tag-and-graft” approach can be generalized as a site-selective protein-labeling method, where a functional group is introduced by modifying the preinstalled atom (S, N, Se, or C) (Figure 1a).12–14 The active alkene moieties on these residues can then serve as electrophiles in Michael addition reactions with thiol and amine nucleophiles,13 with reactive phosphine probes,15 or even radicalophiles, to install a full spectrum of authentic post-translational modifications (PTMs) via a “chemical mutagenesis” approach.12,14,16–18 Therefore, Dha has been used to install multiple important PTMs, including glycosylation,21 phosphorylation,22 methylation, and acetylation.23 In addition, a few recent reports have demonstrated the viability of diastereoselective Michael addition, taking advantage of the chiral protein environment19 or a Rh-based catalyst.20 Thus, it may be possible to make stereoselective modification using this Dha strategy with some necessary improvements in the future.

To date, the insertion of the Dha tag into proteins mostly proceeds chemically via installation of precursors,13 as direct
insertion of unnatural amino acids Dha is not yet possible due to the interference of its reactive side chain to the protein translation machinery. Common chemical precursors of Dha include cysteine, seleno-cysteine/lysine, and pS were converted to Dha by 2,5-dibromohexanediamide, H₂O₂, and Ba (OH)₂, respectively. Lyases OspF and SpvC eliminate phosphate groups from phosphothreonine/serine residues in substrate proteins and convert them to the corresponding Dhb or Dha products. Our proposed strategy to identify a monophosphorylated tag for OspF and SpvC for the efficient enzymatic incorporation of Dha or Dhb.

Figure 1. Formation and reactions of Dha and Dhb. (a) “Tag-and-graft” approach is generally used as a site-selective protein-labeling method for protein of interest (POI). (b) Current methods of introducing Dha. Cysteine, seleno-cysteine/lysine, and pS were converted to Dha by 2,5-dibromohexanediamide, H₂O₂, and Ba (OH)₂, respectively. (c) Lyases OspF and SpvC eliminate phosphate groups from phosphothreonine/serine residues in substrate proteins and convert them to the corresponding Dhb or Dha products. (d) Our proposed strategy to identify a monophosphorylated substrate for OspF and SpvC to provide a practical enzymatic route for site-specific labeling of proteins.

In this paper, we report our approach to identify a monophosphorylated tag for OspF and SpvC via substrate specificity profiling using self-assembled monolayers for matrix-assisted laser-desorption−ionization mass spectrometry (SAMDI-MS) and peptide arrays. SAMDI-MS technique is well-suited for profiling enzyme activity and has been successfully used to characterize the activities of a broad range of enzymes, including phosphatases, proteases, deacetylases, acetylases, glycosyltransferases, and cytochrome P450. Here, we use SAMDI-MS to profile the activity of OspF and SpvC on peptide substrates that have a pS or pT residue flanked by all possible amino acids (except cysteine) at the +1 positions. This approach revealed monophosphorylated tags that display comparable activity to native di-phosphorylated substrates in generating Dha and Dhb. The structure−activity relationship analysis as well as molecular docking studies indicate that hydrophobic and aromatic interactions are the two major determinants for the observed substrate specificity. We also successfully demonstrate that, with the optimized monophosphorylated tag for OspF, the efficiency of the Dha-based peptide labeling is enhanced 10-fold. Finally, we report the first examples of...
protein and live bacteria labeling via enzyme-mediated Dha incorporation.

■ RESULTS AND DISCUSSION

Substrate Specificity Profiling for OspF and SpvC. We first expressed and purified OspF and SpvC (Figure S1) and confirmed their activities using the di-phosphorylated substrate (Erk2 analogue, P1 = CDHTGFL-pT-E-pY-VATR) and monophosphorylated substrate (P38 analogue, P2 = CDEM-pT-GYV) with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), where a mass shift of −98 Da confirm the formation of elimination products. Lyase activity is determined from characteristic mass shifts and relative peak intensities in the MS spectrum. We also performed a kinetic characterization with the EnzChek assay for native di-phosphorylated and monophosphorylated substrates (Table S2) and found kinetic parameters that were consistent with previous studies (Figure S4).

We next designed the peptide array needed for profiling the substrate specificity of the two enzymes. It is known that tyrosine on the +2 position is highly conserved and critical for activity among the MAPK family homologues. Also, the −1 and +1 positions of the peptide substrate are reported to be critical for enzymatic activity: for example, placement of alanine at the −1 position resulted in no detectable enzymatic activity, and substitutions at the +1 position had the greatest impact for substitutions at the −3 to +3 positions. Therefore, we reasoned that by profiling the specificity of ±1 positions flanking pT/pS, we could identify a more efficient monophosphorylated tag for enzymatic incorporation of Dhb or Dha by OspF/SpvC. We next prepared a peptide array having 361 peptides with the sequence Ac-GX−1-pT-X+1GRC, where X−1 and X+1 were one of 19 amino acids, with Cys being excluded to avoid interference with the immobilization reaction. A library with pT is selected, as the pT substrates give higher overall activities with OspF and thus better signal-to-noise ratios in the specificity study. The peptides (2 μL, 20 μM) were transferred onto array plates having 384 gold islands that were modified with a monolayer-presenting maleimide groups at a density of approximately 10% mixed with a tri(ethylene glycol)-terminated alkanethiolate. We then applied the OspF/SpvC enzyme (0.1 μM, 2 μL) to each island and allowed the reaction to proceed for 2 h. The array plates were rinsed, treated with matrix, and analyzed by MALDI-TOF to obtain a mass spectrum for each spot (Figure 2a). The lyase activity was calculated by measuring the relative intensity (I) of each molecular ion peak (M+) (Figures 2b and S5) from each spectrum using the following equation:

\[
\text{Conversion (\%)} = \frac{I(\text{Product})}{I(\text{Product}) + I(\text{Substrate})} \times 100
\]

Figure 2. Substrate specificity profiling for OsP and SpvC. (a) Overview of SAMDI-based activity characterization of OsP and SpvC. Peptide solutions from a Ac-GX−1-pT-X+1GRC library were individually transferred to monolayer array plates where the peptide substrates were immobilized to maleimide-terminated monolayers. These peptide-functionalized spots were then treated with OsP/SpvC to give products. In SAMDI-MS, the monolayer is irradiated with a laser that dissociates the thiolate-gold bond, resulting in the release of ionized alkanethiolate-peptide species from the surface. (b) Examples of SAMDI spectra for enzyme activity quantification. A mass shift of −98 Da indicates the formation of elimination products. Lyase activity is determined from characteristic mass shifts and relative peak intensities in the MS spectrum. (c) Heatmap of OsP activity profiling with the Ac-GX−1-pT-X+1GRC library. The red boxes represent X−1 and X+1 of the three natural substrates. The cyan box indicates that OsP prefers aromatic amino acids for the X−1 and X+1 positions. (d) Correlation of conversion between OsP and SpvC. Each data point represents one peptide sequence from the Ac-GX−1-pT-X+1GRC library. Numeric data for these experiments can be found in Figure S6.
As shown in Figures 2c and S6, yields for the lyase-mediated conversion of Ac-GX$_{-1}$-pT-X$_{+1}$GRC into Ac-GX$_{-1}$-Dhb-X$_{+1}$GRC ranged from 0% to nearly 90%, with OspF and SpvC exhibiting a similar pattern of substrate specificity on the array (Figure 2d). The average conversions for the native motif with M-pT-G (p38), L-pT-G (Erk1/2), and M-pT-P (JNK) were 23 and 33% for OspF and SpvC, respectively (boxed in red, Figures 2c and S6 for numeric data), indicating a suitable dynamic range for the lyase activity assay. Low conversion was observed for Pro (a conformationally constrained residue) at the X$_{-1}$ position, which is consistent with required conformational flexibility observed for the pT-X-pY motif in native substrates. In addition, inclusion of the basic residue Arg at the X$_{-1}$ and X$_{+1}$ positions also resulted in reduced activity, but this was not observed for the other two basic residues, His and Lys. When the 19 amino acids were grouped according to the hydrophobicity and charge of the side chain functional groups, we found that aromatic amino acids Phe, Tyr, and Trp were preferred at the X$_{-1}$ and X$_{+1}$ positions (Figure 2c, boxed in cyan). While one-side substitution with aromatic residues resulted in substantial increase of activity compared to the native motif (23% vs 69% for OspF and 80% vs 33% for SpvC; see Figure S7 for details of calculation).

Molecular Mechanisms of Substrate Specificity. To gain mechanistic insights into the observed specificity pattern, we applied a chemoinformatic approach to determine the factors contributing most to the activity of the lyases. We calculated 365 molecular descriptors for each of the 19 amino acids. Each of these descriptors was used to train a linear least squares regression model that predicts the average conversion efficiency for each amino acid residue. Through this analysis, it was revealed that SlogP (log of the octanol/water partition coefficient) and number of aromatic/hydrophobic atoms were the most significant parameters that correlated with the average conversion efficiency for the X$_{-1}$ and X$_{+1}$ position analogues, respectively (Figure 3a).

The results of our experiments were also consistent with molecular modeling studies. Figure 3b shows the binding mode of one of the best substrate peptides, GGY-pT-WGG, to the OspF protein, which was predicted by the hierarchical protein–peptide docking program HPEPDOCK. We see that the binding pocket accommodating the peptide consists of three distinct regions: one hydrophilic region in the middle and two hydrophobic regions on each side (Figure 3b). The middle hydrophilic region is formed by four residues (K102, R146, R211, R218, and H104) responsible for the binding of pT at the X$_{0}$ position of the peptide (Figure 3c). One of the two hydrophobic regions, which is formed by four hydrophobic residues (V83, F84, V147, and A151) is located...
at the X,1 position (Figure 3c). The other hydrophobic region, which is mainly formed by three hydrophobic residues (F98, Y156, and L214), is located at the X,1 position (Figure 3c). The presence of these two hydrophobic regions explains the observed preference of aromatic residues Phe/Tyr/Trp at both X,1 and X,1 positions, as they can form favorable hydrophobic interactions with the hydrophobic regions on the protein.

We also substituted the X,1 and/or X,1 positions with each of the 19 amino acids and calculated the binding energy scores between OspF and these peptides. We found that the calculated binding energy scores showed a good correlation with the experimental data. That is, peptides with hydrophobic residues at X,1 and X,1 positions generally give lower binding energy scores (Figure 3d, see Figure S8 for numeric data). Overall, the Pearson correlation coefficient between the experimental data and the calculated binding scores is 0.67 (Figure 3e). This result suggests the reliability of our molecular modeling and proposed binding mechanism.

Optimization of a Monophosphorylated Tag for OspF/SpvC. We next asked whether the optimal residues at the ±1 positions could be used to improve the activity of the monophosphorylated peptide sequence from the best native substrate protein p38, CDEM-pT-GYV (P2). We synthesized a peptide that substituted ±1 positions with Y/W to obtain CDEM-pT-WYV (P3). Similarly, we prepared the pS version of both P2 and P3, RDEM-pS-GYV (P2-pS) and RDEY-pS-WYV (P3-pS), respectively.

We then characterized the conversion of these four peptides by OspF and SpvC and obtained kinetic parameters by mono- and di-phosphorylated peptides. Molecular docking studies of these peptides (P2-pS: 0.053 s⁻¹; P3-pS: 0.059 s⁻¹). The data indicate that the substitution of Lys to Y/W significantly increased the binding affinity between the peptide and the lyase and serves as the major contribution to the observed increase in catalytic efficiency for the modified monophosphorylated peptides. Molecular docking studies of these protein—peptide interactions also provided an insight into the enhanced lyase-substrate binding: the binding energy of P3-OspF is 56.3 kcal/mol lower than that of P2-OspF (∼185.29 kcal/mol vs −128.96 kcal/mol). In short, the optimization of ±1 positions resulted in a 2.8−5.8-fold overall improvement of catalytic efficiency. This, therefore, enables the simplification of the di-phosphorylated X-pT/pS-X-pY motif to a monophosphorylated [F/Y/W]-pT/pS-[F/Y/W] sequence that shows comparable activity with the best di-phosphorylated substrate reported to date.

Enchancement of Peptide Labeling Efficiency with the Optimized Monophosphorylated Tag. The usefulness of the OspF/SpvC-mediated transformations stems from the versatility of the product residues, Dha/Dhb, which can be used as sites for covalent modifications of peptides or proteins through nucleophilic addition (thia-Michael addition and aza-Michael addition) as well as radical addition (Figure 1b). As a proof-of-concept demonstration of the application of this lyase-mediated α, β-unsaturated carbonyl insertion, we first investigated the reactivity of Dha and Dhb in the optimized peptide (R/CDEY-Dha/Dhb-WYV) toward four thiols [3-mercaptobutane-1,2-diol, ethyl-2-mercaptoacetate, N-acetylcysteine, and 2-(dimethylamino) ethanethiol]. The Dha/Dhb-containing peptide R/CDEY-Dha/Dhb-WYV was first generated by treating phosphopeptide (R/CDEY-pS/pT-WYV) with OspF for 30 min, and subsequent Michael addition reactions were carried out with 20 mM thiol at room temperature for 2 h (Figure 5a). Indeed, Dha showed good reactivity as a thiol-Michael acceptor with conversion of 75−92% with the four thiolis (Figure 5b), as characterized by MALDI-TOF MS and MS/MS (Figures 5c,d and S9−11). However, the reaction of Dhb with thiol compounds did not
result in any detectable product formation in thia-Michael addition, due presumably to the attenuated electrophilicity of Dhb as a result of steric hindrance and hyperconjugation of the additional methyl group (Figure S12). To increase the reactivity of Dhb, a recently developed Dhb-reactive phosphine probe may be coupled with the pT-based monophosphorylated tag for efficient labeling of protein targets. In addition, ATP–Cu(II) catalyst and Mn(I) complexes were also reported to efficiently catalyze Michael reaction for structures bearing a Dhb motif.

Having confirmed the reactivity of Dha, we then compared the labeling efficiency of the optimized monophosphorylated tag (P3-pS) and the unoptimized monophosphorylated tag (P2-pS). (f) Labeling rate of optimized monophosphorylated tag (P3-pS) is 10× fold higher than the unoptimized monophosphorylated tag (P2-pS) with 2-(dimethylamino) ethanethiol. (g) Labeling rate of the optimized monophosphorylated tag (P3-pS) is 6.5× fold higher than the unoptimized monophosphorylated tag (P2-pS) with N-acetylcysteamine.

Figure 5. Site-specific peptide labeling via OspF-mediated Dha insertion. (a) Scheme for stepwise OspF-mediated Dha insertion and thia-Michael addition reactions to enable site-specific peptide labeling. (b) Chemical structures for the four thiols and the corresponding yields of the reactions. (c) Product of 2-(dimethylamino) ethanethiol and P3-Dha was detected with MALDI-TOF. (d) Tandem MS analysis of reaction products confirming successful labeling of Dha with 2-(dimethylamino) ethanethiol. Fragment ions “b” are shown in red, and fragment ions “y” are shown in blue. (e) Scheme of the one-pot reaction to compare the labeling efficiency of the optimized monophosphorylated tag (P3-pS) and the unoptimized monophosphorylated tag (P2-pS). (f) Labeling rate of optimized monophosphorylated tag (P3-pS) is 10× fold higher than the unoptimized monophosphorylated tag (P2-pS) with 2-(dimethylamino) ethanethiol. (g) Labeling rate of the optimized monophosphorylated tag (P3-pS) is 6.5× fold higher than the unoptimized monophosphorylated tag (P2-pS) with N-acetylcysteamine.
modification group are all present. The results in Figure 5f,g show that the optimized P3-pS displayed 10-fold and 6.5-fold enhancement of conversion over the unoptimized P2-pS, with 2- (dimethylamino) ethanethiol and N-(2-sulfanylethyl) acetamide, respectively. These data demonstrate that optimization of the residues at the ±1 positions of the monophosphorylated substrate was effective at identifying a monophosphorylated substrate for modification by lyases and subsequent use in peptide/protein-labeling applications.

OspF-Mediated Labeling of GFP Protein and E. coli Cell with the Optimized Monophosphorylated Tag. We next evaluated the OspF-mediated Dha insertion and labeling method in bacterial cells. We first evaluated whether the optimized monophosphorylated tag could be used to introduce a biotin group on the model protein GFP (Figure 6a). For such in cellulo applications, the Dha precursor can be readily introduced via efficient genetic encoding of phosphoserine. Chin and co-workers evolved an orthogonal aminoacyl-tRNA synthetase/tRNA<sub>CUA</sub> pair that directs the efficient incorporation of pS into recombinant proteins in E. coli. This approach enables quantitative decoding of the amber stop codon as pS. They also discovered a phosphothreonyl-tRNA synthetase/tRNA<sub>CUA</sub> pair and created an entirely biosynthetic route to incorporate phosphothreonine in proteins. In addition, phosphorylation of Ser or Thr on target peptide can be made possible with a choice of kinase with suitable substrate specificity. Here, as a quick demonstration, we prepared C-terminus-tagged GFPs through ligation of GFP-LPETG and the monophosphorylated peptide (P4: GGGRDEYM-pS-WYV or P5: GGGRDEYM-pS-WYV) using the transpeptidase SrtA. The resulting proteins, GFP-GGGRDEYM-pS-WYV (GFP-P3-pS) and GFP-GGGRDEYM-pS-GYV (GFP-P2-pS), were used to add a biotin group with a poly(ethylene glycol) linker (Figure 6b). We treated GFP-P3-pS (30 μM) with OspF (S

Figure 6. Protein and live cell labeling via OspF-mediated enzymatic incorporation of Dha. (a) Scheme showing the production of GFP-P2/P3-Dha and their reactions with biotin-PEG-SH. (b) Western blot detects biotinylated GFP (GFP-P3-Biotin) only in the positive sample with all four components (GFP-LPETG, SrtA, OspF, and biotin-PEG-SH). (c) Labeling efficiency of biotinylated GFP-GGGRDEYM-pS-WYV (GFP-P3-pS) is significantly higher than that of GFP-GGGRDEYM-pS-GYV (GFP-P2-pS). (d) Scheme of outer membrane protein Lpp-OmpA modification of live bacterial cell surfaces via OspF-mediated Dha insertion. (e) Confocal microscopy of rhodamine-labeled E. coli showing that bacteria can be labeled via the surface-displayed Lpp-OmpA-GGGRDEYM-pS-WYV. (f) Flow cytometry of rhodamine-labeled E. coli. The labeling efficiency of bacteria with Lpp-OmpA-GGGRDEYM-pS-WYV is significantly higher than the control group (no overexpressed OmpA).
µM) for 4 h at room temperature and a thiol-PEG-biotin (50 mM) to generate a biotinylated GFP that was visualized by streptavidin-HRP using Western blot. We next compared the rate of labeling between the two GFPs with different tags (GFP-P2-pS: GGGRDDEM-pS-GGYV GFP-P3-pS-GGGRDDEY-pS-WYV). The results in Figure 6c show that within the time frame of the experiment, GFP with the optimized tag, GFP-P3-pS, was clearly labeled, but no detectable label was found in the GFP with the unoptimized tag, GFP-P2-pS.

Finally, we evaluated whether the optimized monophosphate tag could enable OspF-based labeling of live bacteria. In this study, we employed an E. coli strain that overexpresses a surface-exposed transmembrane protein Lpp-OmpA-LPETG.56,57 We then installed the OspF recognition motif from Dha and Dhb can be exploited for applications in which dual-site protein labeling is desired.58

Additionally, the di-monophosphorylated tag has further lowered technical barriers for such applications. Additionally, the di-monophosphorylated tag has further lowered technical barriers for such applications. Additionally, the di-monophosphorylated tag has further lowered technical barriers for such applications.

RESULTS

General. Unless otherwise stated, all chemical reagents are of analytical grade, obtained from commercial suppliers and used without further purification.

Materials. Genes and Cells. The Shigella flexneri osp gene was synthesized at Kingsray Biotechnology Co., Ltd. The spv, sra, and ospA genes were amplified from Salmonella typhimurium, Staphylococcus aureus, and E. coli, respectively. All competent cells, including shuffle T7, E. coli DE3, E. coli DH5α, and E. coli DH10B, were purchased from Thermo Fisher.

Chemicals, Enzymes, and Other Reagents. 2-amino-6-mercaptop-7-methylpurine riboside (MESG) was purchased from Cayman Co., Ltd., and purine nucleoside phosphorylase (PNP) was purchased from Thermo Fisher. Potassium chloride (KCl), sodium chloride (NaCl), Tris–HCl, glycine, nickel(II) sulfate (NiSO4), and sodium phosphate were purchased from Hefei Bomei Biotechnology Co., Ltd. Thiol compounds were purchased from Shanghai Macklin Biochemical Co., Ltd. The CloneExpress II One Step Cloning Kit was purchased from Novizan Biotechnology Co., Ltd. Restriction endonucleases and protease inhibitor cocktails were purchased from Thermo Fisher. Streptavidin was purchased from Sangon Biotech (Shanghai) Co., Ltd. The GFP antibody and secondary antibodies were purchased from Abmart Pharmaceutical Technology (Shanghai) Co., Ltd. 96-well filter plates (cat. no. AWFP-F20000, Arctic White LLC) and all fluorenlymethylocarbonyl (Fmoc) amino acids and Rink-amide resins were purchased from AnaSpec, Inc. All solvents, N,N-dimethylformamide (DMF), dichloromethane, trifluoroacetic acid (TFA), and piperidine were purchased from Thermo Fisher Scientific. Other chemical reagents used in peptide synthesis were purchased from Sigma-Aldrich unless otherwise noted.

Gene Cloning, Expression, and Purification. Cloning. All protein plasmids used in this study were generated by homologous recombination with standard protocols. Briefly, the target gene amplified with forward and backward primers, purified, and ligated into the corresponding vectors with recombinants. Information of vectors, restriction sites, and primers is given in Table S1.

Protein Expression. Plasmids were transformed to the corresponding competent cells (shufffe T7-K12 and E. coli DE3 for protein expression; E. coli DH5α and E. coli DH10B for plasmid amplification). Transformants were plated on a LB agar plate supplemented with ampicillin (100 µg/mL) or kanamycin (50 µg/mL) and incubated overnight at 30 °C. A single colony was inoculated into 20 mL of LB supplemented with ampicillin (100 µg/mL) or kanamycin (50 µg/mL) and agitated at 220 rpm at 30 °C. On the following day, overnight cell culture was diluted in 1000 mL LB supplemented with ampicillin (100 µg/mL) or kanamycin (50 µg/mL) and agitated at 220 rpm at 30 °C until the OD600 reached 0.6–0.8. Protein expression was induced with 1 mM IPTG at 18 °C for 20 h. Cells were collected by centrifugation at 8000 g for 40 min at 4 °C and stored at −80 °C.

Protein Purification. Cell pellets were resuspended in 25 mL lysis buffer (50 mM Tris–HCl, pH 8.0, 250 mM NaCl, lysozyme 1 mg/mL, DNase 0.1 mg/mL, and 1× protease inhibitor cocktails). The cell suspension was sonicated with Sonic Dismembrator (Ningbo Sciento Biotechnology Co., LTD, with 40% output, 40 min, 2 s off, 1 s on) in an ice-water bath, followed by centrifugation to remove the debris (18,000 g, 60 min, 4 °C). The supernatant was collected and filtered through a 0.45 µm filter. The filtrate was loaded onto a Poly-Prep Chromatography Column, washed with lysis buffer, and eluted by gradually increasing the eluent buffer (50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole) for 60 min. The eluate was concentrated and desalted against a buffer with no imidazole. Bacterial lysates and eluted fractions were characterized using SDS-PAGE stained with Coomassie brilliant blue. The GST tag was removed from the target protein by cleavage with HRV 3C at a concentration of 10 µg per mg target protein for an overnight at 4 °C. The digested protein mixture was purified again with a nickel column, and the eluates were collected and characterized with SDS-PAGE. The final protein product was concentrated and desalted into protein.
storage buffer (50 mM Tris−HCl, pH 7.4, and 150 mM NaCl, 25% v/v glycerol) at a concentration of 3.2 mg/mL, and stored at −80 °C.  

SAMDI-Based Substrate Specificity Profiling. Solid-Phase Synthesis of Peptide Arrays. Solid-phase peptide synthesis (SPPS) was performed as described previously.21 Briefly, SPPS was performed on Rink-amide resin (10 mg) housed in 96-well filter plates. N-terminal Fmoc protecting groups were deprotected with 20% piperidine in DMF at room temperature for 20 min. The resin was filtered and rinsed five times with DMF on a multiscreen vacuum manifold. Amino acids were coupled to the resin with PyBop and N-methylmorpholine (NMm) in DMF at a 4:4:8 molar excess for 20 min, twice. The deprotection and coupling were repeated for each residue, following deprotection of the final residue. Peptides were cleaved and deprotected in 95% TFA, 2.5% triethylsilane, and 2.5% water for 16 h. The cleavage solution was evaporated with N2 gas flow. The peptides were resuspended in 0.1% TFA in water, lyophilized, and resuspended again in 0.1% TFA in water to a final concentration of 200 μM and stored at −80 °C.  

Preparation of SAMDI Plates. Steel array plates evaporated with 384 gold spots (with a diameter of 3.0 mm) were soaked at 4 °C in a solution of 1 mM total disulphide with 0.8 mM tri(ethylene glycol)-terminated C11-alkaline disulphide and 0.2 mM C11-alkaline disulphide with one terminal tri(ethylene glycol) and one terminal maleimide in ethanol for 2 days to allow assembly of the monolayer (10% maleimide coverage). Peptides were diluted to a final concentration of 20 μM with 100 mM Tris−HCl buffer (pH 7.5) in a 384-well plate with 5 μL of TCEP beads (Thermo Scientific) in each well. Using a TECAN robotic liquid handler, 2 μL of peptide solution from each well was transferred on to the corresponding gold spot on the peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated at 37 °C for 1 h to allow for peptide immobilization.  

Profileing Conversion Efficiency of OspF/SpvC. OspF/SpvC (0.1 μM) was diluted in Tris−HCl buffer (100 mM Tris−HCl, pH 7.5, 50 mM NaCl), and 2 μL was applied to each gold spot on the peptide array plate with a Multidrop Combi (Thermo Scientific). The reactions were incubated at 37 °C in a humidified chamber for 1 h to allow for peptide immobilization.  

Enzyme Activity Assay and Kinetic Characterization. Product Characterization with MALDI. The OspF/SpvC reactions were performed using 0.1 μM enzyme and 250 μM peptide in Tris−HCl buffer (50 mM, pH 7.5) for 30 min at 25 °C. The reactions were quenched by heating at 95 °C for 10 min. The reaction solution was mixed with 4 μL of the CHCA matrix (5 mg/mL), and the reaction products were detected by a £800 MALDI-TOF/TOF (ABSciex) with reflective positive ion mode.  

Kinetic Characterization with EnzChek Assay. The steady-state kinetic analysis was performed according to the previously reported phosphate detection method.38 Briefly, OspF/SpvC was diluted in Tris−HCl buffer (50 mM, pH 7.5) containing 0.1 mg/mL MESG and 0.3 unit of PNP to a final concentration of 100 nM. Reactions were initiated by adding substrate peptide at different concentrations (0−300 μM) to each well of a 384-well plate to a final volume of 30 μL. The reaction was monitored at 360 nm for product formation for 60 min with a plate reader (Spark 10 M, TECAN). Initial velocity (V0) for product formation was calculated following the linear range of the reaction profile.  

\[ V_0 = \frac{\text{Absorbance (360 nm)}}{11,000 \text{ M}^{-1} \text{cm}^{-1} \text{X Time (s)}} \]

Initial velocities were then analyzed with the Michaelis–Menten kinetics equations by using GraphPad Prism 6 (GraphPad Software, San Diego) to obtain the kinetic parameters.  

Quantitative Analysis and Molecular Modeling. Chemo-informatics Analysis. Physiochemical descriptors of amino acids were calculated using Molecular Operation Environment (MOE, version 2014) after energy minimization of the structure with Amber 10 force field.84 Data with each descriptor were used to train a least linear regression model that predicts the average conversion for each amino acid residue.  

Molecular Modeling of OspF-Peptide Interactions. The OspF-peptide interactions were modeled using a template-based docking approach. Specifically, the three-dimensional (3D) OspF structure was modeled based on the high-homologous SpvC protein structure (PDB ID: 2Z8P) using the homology modeling program Modeler.62 To retain the peptide-bound conformation of OspF, the native peptide ligand from the SpvC structure was kept during the homology modeling process. Given the high sequence identity of 63.2% between OspF and SpvC (Figure S13a), the obtained OspF protein had a high-quality structure with an excellent TMscore of 0.998 (Figure S13b).59 For each sequence in the GGX3−PT-X3−GG library, the 3D conformation of this modeled peptide was constructed based on the native peptide ligand from the SpvC structure by MOPDEP program.44 Then, with the 3D structures of OspF and modeled peptides, the putative peptide binding modes were sampled using our hierarchical protein−peptide docking program HPEPDock,43,44,63 where the ITScorePP scoring function for protein−protein interactions was used to evaluate the binding score between the protein and the peptide.55 The top binding mode with the best binding score was chosen as the predicted OspF-peptide complex structure.  

OspF-Based Peptide Labeling. Generation of Dha/Dhb-Containing Peptide. 1 mM phosphorylated peptide was dissolved in 100 μL Tris buffer (100 mM, pH = 7.5), to which purified OspF was then added to a final concentration of 10 μM. The reaction was then incubated at 25 °C for 30 min before it was quenched by and analyzed by MALDI as mentioned above to ensure full conversion of phosphorylated peptide to Dha/Dhb-containing peptide.  

Coupling of the Dha/Dhb-Containing Peptide with Thiol Compounds. 50 μL of 20 mM thiol compounds in Tris−HCl buffer (100 mM, pH = 7.5) was added to the abovementioned Dha/Dhb-containing peptide (1 mM, 50 μL), and the reactions were carried out at room temperature for 2 h before it was quenched and analyzed by MALDI as mentioned above. The percentage conversions of the reaction were calculated using the following formula:

\[
\text{Conversion (\%)} = \frac{I(\text{product})}{I(\text{product}) + I(\text{substrate})} \times 100
\]

The thiol adduct peptide was also analyzed with Tandem MS/MS (Data Explorer V4.5).  

“One-Pot” Reaction for Comparison of Labeling Efficiency. The phosphorylated peptide, OspF, and thiol compound were dissolved in 50 μL of sodium phosphate buffer (50 mM, pH 8.0) at final concentrations of 250 μM, 0.1 μM, and 25 mM, respectively. The percentage conversion of each reaction was monitored at 1, 2, 5, 10, 15, and 20 min by quenching and analyzing with MALDI as mentioned above.  

Protein Labeling. Preparation of GFP-P2/P3-Dha. 30 μM GFP-LPETG, 60 μM StmA (D107A, G109E mutant), and 1 mM monophosphorylated peptide (P4: GGGRDEM-pS-GYV or P5: GGGRDEY-pS-WYV) were mixed in 50 μL reaction buffer (Tris−HCl 50 mM, 10 mM CaCl2, pH = 7.5). The transpeptidase reaction was allowed for 6 h at room temperature, and then monophosphorylated peptide and CaCl2 were removed by ultrafiltration with Tris−HCl buffer (50 mM, pH = 7.5). OspF was then added to a final concentration of 5 μM, and the mixture was incubated for 4 h at room temperature for Dha formation.
Labeling GFP-P2/P3-Dha with Biotin-PEG-SH. 50 mM biotin-PEG₄-SH was added to the abovementioned reaction mixture and reacted overnight at room temperature. The GFP-P3-Biotin was imaged and quantified by Western blot with either 1:1000 HRP-conjugated streptavidin to detect the biotin tag or with mouse monoclonal eGFP (1st)/Goat Anti-Mouse IgG HRP (2nd) to detect GFP. The GFP-P3-Biotin was identified as the band with concurrence on the biotin blot and GFP blot.

Comparison of Protein-Labeling Efficiency. Labeling efficiency of GFP-P2/P3-pS was compared similarly as mentioned above, with OspF-mediated Dha formation being terminated at different time points (1, 2, 3, and 4 h). The labeling efficiency presented in Figure 6c was quantified with ImageJ (version 1.8.0).

Live Bacterial Cell Surface Labeling. A truncated version of transmembrane protein OmpA was used in this experiment, and it was redirected to the bacterial cell surface by adding a signaling peptide Lpp.36,37 Bacteria harboring pET28a-Lpp-OmpA-LPETG were grown in 5 mL of LB medium supplemented with 1 mM IPTG and 50 μg/mL kanamycin for 20 h, harvested by centrifugation (1200 g), washed with a gentle agitation of 120 rpm. The cells were then washed with phosphate buffer (pH 7.4) at a final OD₆₀₀ of ~25 and a volume of 200 μL. SrtA and monophosphorylated peptide (GGGRDEY-pS-WYV) were added to final concentrations of 10 μM and 1 mM, respectively. The transpeptidase reaction was then proceeded at room temperature for 5 h with a gentle agitation of 120 rpm. The cells were then washed with phosphate buffer to remove SrtA and monophosphorylated peptide and resuspended in 150 μL of a 100 mM phosphate buffer (pH 7.4). OspF (5 μM) and SH-PEG-rhodamine (1 mM) were added to the bacteria suspension, and the eliminylation reaction and (pH 7.4) at a transpeptidase reaction was then proceeded at room temperature for 5 h with a gentle agitation of 120 rpm. The cells were then washed with phosphate buffer to remove SrtA and monophosphorylated peptide and resuspended in 150 μL of a 100 mM phosphate buffer (pH 7.4). OspF (5 μM) and SH-PEG-rhodamine (1 mM) were added to the bacteria suspension, and the eliminylation reaction and (pH 7.4) at a transpeptidase reaction was then proceeded at room temperature for 5 h with a gentle agitation of 120 rpm. The cells were then washed with phosphate buffer and subsequently analyzed with confocal imaging and flow cytometry. The bacteria sample without overexpressed Lpp-OmpA-LPETG was used as control.

Protein sequence of Lpp-OmpA: MKATKLVLGAVILGSTL-

GPTHEQLGAGAAGGGYQVNYPYFEMGDWLRMPYKGS-
VEGANYKAGQVQLTAKLINGITDDLDLITYTRLGGMV-
WRADTKSNVYGTKHNHTGVSPVFAGGVEAYITPEATGSGGL-
PETGG.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.1c00866.

OspF/SpvC expression, enzyme assays and kinetic profiles, raw results for phosphorylated peptide library screening with SAMDI-MS, additional characterizations of Michael addition reactions, all plasmid construction information, peptide sequence information, and kinase specificity table (PDF)

Modeled complex structure (PDB)

AUTHOR INFORMATION

Corresponding Authors

Xinxin Feng — Institute of Chemical Biology and Nanomedicine, State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan Provincial Key Laboratory of Biomacromolecular Chemical Biology, and Department of Chemistry, Hunan University, Changsha 410082, China; orcid.org/0000-0003-3142-5629; Email: xinxin_feng@hnu.edu.cn

Milan Mrksich — Departments of Chemistry and Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States; orcid.org/0000-0002-4964-796X; Email: milan.mrksich@northwestern.edu

Sheng-You Huang — School of Physics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China; orcid.org/0000-0002-4209-4565; Email: huangsy@hust.edu.cn

Authors

Anming Yang — Institute of Chemical Biology and Nanomedicine, State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan Provincial Key Laboratory of Biomacromolecular Chemical Biology, and Department of Chemistry, Hunan University, Changsha 410082, China

Huanyu Tao — School of Physics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

Lindsey C. Szmyczak — Departments of Chemistry and Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States

Liang Lin — State Key Laboratory of Bio-organic and Natural Products Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Junfeng Song — Institute of Chemical Biology and Nanomedicine, State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan Provincial Key Laboratory of Biomacromolecular Chemical Biology, and Department of Chemistry, Hunan University, Changsha 410082, China

Yi Wang — Institute of Chemical Biology and Nanomedicine, State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan Provincial Key Laboratory of Biomacromolecular Chemical Biology, and Department of Chemistry, Hunan University, Changsha 410082, China

Silei Bai — Institute of Chemical Biology and Nanomedicine, State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan Provincial Key Laboratory of Biomacromolecular Chemical Biology, and Department of Chemistry, Hunan University, Changsha 410082, China

Justin Modica — Departments of Chemistry and Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.1c00866

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (21807031 and 22177031 to X.F.), 602072199 and 31670724 to S.-Y.H.), Hunan Provincial Natural Science Foundation of China (grant No. 2020JJ4177 to X.F.), Open Funding Project of the State Key Laboratory of Biocatalysis and Enzyme Engineering (grant SKLBE2019003 to X.F.), Hunan Provincial Innovation Foundation For Postgraduate (grant No. CX20190264 to A.Y.), the Defense Threat Reduction Agency (grant No. HDTRA1-15-10052 to M.M.), and the Robert H. lurie Cancer Center’s Translational Bridge Program at Northwestern University. We thank N. Zhou in Hunan University for kindly providing the SrtA plasmid.

REFERENCES


